

**AHFE**

**Volume 157**

# **International**

**Tareq Ahram, Luca Casarotto and  
Pietro Costa  
Editors**

## **Human Interaction & Emerging Technologies (IHiet 2024)**

**Proceedings of the 12th International Conference on Human  
Interaction & Emerging Technologies, August 26-28, 2024, IUAV  
University of Venice, Italy**

**Open Access Science in Human  
Factors Engineering and Human  
Centered Computing**

**Issue 157 2024**

# Autonomy at the Crossroads: Knowledge Workers Teamed With Intelligent Machines a Qualitative Systematic Review

Mahdis Smith<sup>1</sup>, Luke Houghton<sup>1</sup>, Carla Riverola<sup>1</sup>, and Ali Intezari<sup>2</sup>

<sup>1</sup>School of Business, Strategy, and Innovation, Griffith University, Brisbane, Australia

<sup>2</sup>University of Queensland Business School, Brisbane, Australia

## ABSTRACT

**Purpose** – This study aims to identify risks in adopting artificial intelligence (AI) for organizational decision-making by examining empirical studies. AI is increasingly applied to automate tasks and decisions which were traditionally made by humans, posing challenges to sense of autonomy.

**Design/methodology** – A total of 28 empirical studies were selected using predefined inclusion and exclusion criteria. To this end, this research systematically explored the processes of inquiry, identification, selection, critical appraisal, and the synthesis of empirical studies. This study is undertaken to address the following primary inquiries: (1) What is the direction of the observed effect? (2) What is the magnitude of the effect within the inclusion criteria? (3) Does the effect exhibit a consistent pattern across the spectrum of studies encompassed in the analysis? (4) What is the level of evidentiary robustness underlying the discovered effect?

**Findings** – This content analysis interpreted within task-technology fit (TTF) model revealed that AI adoption represents a promising outlook for the future of human-AI teams. Anchoring on reliable data, this qualitative systematic review informs knowledge workers and leaders on adoption of AI systems and how it positively influences their working processes.

**Contributions/value** – This research conducted a structured analysis to reveal the gap between the collective perception of AI adoption and what leaders and knowledge workers have experienced in relying on AI systems. AI tools are becoming more autonomous therefore a true representation of human-AI team interaction must be displayed. By uncovering the diverse approaches of leaders and the reactions of knowledge workers to AI integration, this paper contributes to a deeper understanding of the evolving landscape of working in the age of AI. The provided insights can assist organizations in harnessing the potential of AI while maintaining a healthy balance of autonomy within their domain.

**Keywords:** Leadership, AI-driven decisions, Problem-solving, Digital autonomy, Hybrid intelligence, Human-machine teams, Qualitative systematic literature review

## INTRODUCTION

This study investigates the balance of human-AI autonomy in leaders and knowledge workers, focusing on the impact of AI on critical thinking.

Through a qualitative systematic review of peer-reviewed literature, it examines concerns about AI integration into decision-making and problem-solving, emphasizing the need to understand its implications thoroughly.

## BACKGROUND

The ever-accelerating pace of technological advancement has left organizational leaders in a state of bewilderment (Dencik et al., 2023) as they wrestle with the consequences of embracing AI. *AI is the system's ability to decode data, learn from, and use that learning to accomplish defined tasks* (Kaplan & Haenlein, 2019a). Various categories of AI, including neural networks, swarm intelligence, genetic algorithms, and fuzzy logic, are applicable in addressing diverse real-world problems (Autor, 2015).

On the one hand, AI systems are autonomously learning to improve processing data (Lange et al., 2012) and as they grow to become more autonomous, so is the gap in understanding how and where they are being used in driving organisational tasks and decisions (Booyse & Scheepers, 2023). On the other hand, freedom and autonomy are pillars of human's liberal democracy, values, ethics, and dignity (Prunkl, 2022; Raz, 1986; Roessler, 2021). Anderson and Honneth reconceptualize autonomy, underscoring that individuals' ability to live autonomously is influenced by their social interactions. They present a recognitional framework where autonomy arises from socially embedded capacities, emphasizing the significance of self-respect, self-trust, and self-esteem in enabling individuals to pursue fulfilling lives (Anderson & Honneth, 2005). To embark on an exploration of how leaders and knowledge workers engage with AI systems, it is imperative to comprehend the implications of their reliance on AI tools. To achieve this, task-technology fit (TTF) (Goodhue & Thompson, 1995) provides the capacity to recognize how elements weigh on adoption of AI technology. This theory provides a platform for conceptualizing the thematic findings of this review.

## METHODS

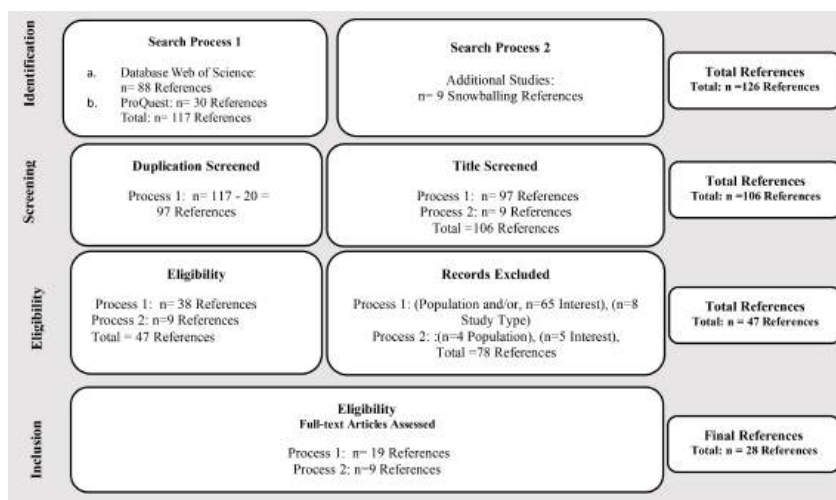
Qualitative systematic reviews are conducted to detect, handpick, critically evaluate, and synthesize data from empirical studies, with the aim of addressing the following inquiries: "(1) What is the direction of the observed effect? (2) What is the magnitude of the effect within the inclusion criteria? (3) Does the effect exhibit a consistent pattern across the spectrum of studies encompassed in the analysis? (4) What is the level of evidentiary robustness underlying the discovered effect? (Higgins, 2008; Paré et al., 2015; Popay et al., 2006)" A typical technique for aggregating research findings within the framework of qualitative systematic reviews is narrative synthesis. Narrative synthesis applies a non-statistical narrative approach for summarizing the outcomes of studies during the synthesis process (Petticrew et al., 2009). Additionally PICO model (Population; Interest/Intervention; Comparison; Outcome) has exhibited utility in this investigation, not only as a search strategy tool but also in the formulation of search terms (Considine

et al., 2017; Eriksen & Frandsen, 2018, p. 69). Within the scope of this investigation, the PICO framework has been defined as follows: (1) Population of leaders and knowledge workers, (2) Interest in adoption of AI systems (tasks/decisions); (3) Comparison of positive and negative impacts of relying on AI-driven tasks or decisions; and (4) Outcome of each individual record.

**Table 1.** Inclusion and exclusion criteria – (ProQuest and web of science).

Inclusion Criteria	Exclusion Criteria
(1) “Leader*” AND “Artificial Intelligence”; (2) “Manage*” AND “Artificial Intelligence”; (3) “Autonomy*” AND “Artificial Intelligence”; (4) “Choice*” AND “Artificial Intelligence”; (5) “Custom* Instruct*” AND “Artificial Intelligence”. Academic Journals (Peer Reviewed), and Reviews English Years 2017 – 2022	Articles not responding to the research question. Journal article that does not relate to “Leadership”, AND “Decision-Make*”, AND “Decision Make*”, OR “Problem Solve*”, OR “Problem- Solve*”, AND “Artificial Intelligence”, OR “AI”. Book Papers, Conference papers, Thesis, dissertation, and non-academic material Non-English Documents, and academic papers

A systematic database search was executed in ProQuest and Web of Science, adhering to the predefined criteria (see Figure 1). This analysis yielded meaningful insights about the phenomenon (Elo & Kyngäs, 2008) demonstrated in detail (see the Appendix).

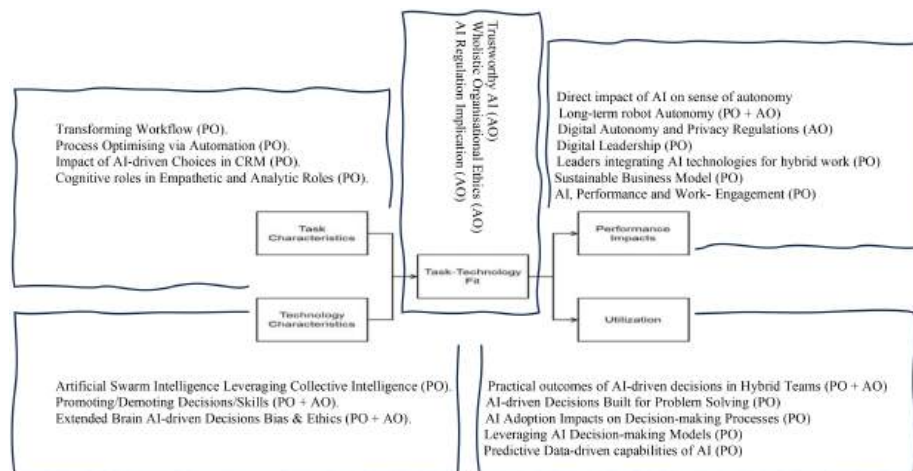


**Figure 1:** CCDAN PRISMA flowchart (Moher et al., 2009).

## FINDINGS

This qualitative systematic review presents a positive overarching outcome with a magnitude weight of 68% from the total included studies. Records reporting a positive outlook shown as (PO) in the Appendix endorse a future where human-AI collaborations shape sustainable business models, mount hybrid intelligence, and optimize workflows. As reflected in Figure 2, Nineteen records recognised AI tools as a catalyst for organizational development, provoking sustainable business models, and transforming workflow and work engagement. AI-driven decision models emphasize pragmatic outcomes through predictive data-driven capabilities, transforming organizational landscapes and optimizing efficiency. Integration of AI systems enhances cognitive abilities of knowledge workers, promoting organizational skills (Hao et al., 2020).

Despite the predominantly positive findings regarding AI integration in this review, 11% of the records exhibited diverse impacts. Organizational AI adoption yielded mixed effects, as indicated in the records. These effects, categorized as Positive Outlook (PO) and Alarming Outlook (AO) in the appendix.



**Figure 2:** Qualitative content analysis categorisation integrated with “task-technology-fit approach (Goodhue and Thompson, 1995, p. 220)”.

Despite the predominantly positive findings regarding AI integration in this review, 11% of the records exhibited diverse impacts. Organizational AI adoption yielded mixed effects, as indicated in the records. These effects, categorized as Positive Outlook (PO) and Alarming Outlook (AO) in the Appendix, highlighted themes suggesting that AI usage may lead to less authentic decision-making (Hao et al., 2020), where others such as Wang (2021) positively defined AI-driven decisions as an extension to brain. These studies emphasize the crucial need for organizations to maintain a balance between skill promotion and demotion. These findings underscore

the emergence of a bidirectional human-AI autonomy interface, where AI systems can learn from human feedback and vice versa.

Floridi and Cows (2019) link autonomy to humans' inclination to delegate decisions to AI, while Wang (2021) explores how human moral judgment can enhance AI-assisted decision-making. Recognizing the significance of human-machine interdependence offers insights into decision-making AI systems, revealing critical challenges and opportunities. Independent analysis of records and exploration of emerging themes reveal discernible patterns, as depicted in Figure 2.

Remainder records, comprising 21% of reviewed studies, highlight structural deficiencies in organizational AI adoption. These studies scrutinize ethical regulations and autonomy balancing, emphasizing the intricate boundaries of AI adoption within organizational contexts. Negative impacts of AI adoption involves various dimensions, particularly concerning fear, trust, and communication within human-machine teams (Abbass, 2019; Baum, 2020). Establishing trustworthy AI is crucial for achieving sustainable autonomy balance (Jones, 2018). Organizational ethics regulations are pivotal in safeguarding digital autonomy and privacy (Baum, 2020; Dobbe et al., 2021). In the absence of regulated guidelines for digital autonomy, human privacy may be compromised (Jarrahi, 2018; Laacke et al., 2021).

## DISCUSSION

Exploiting the depth of each study within the task-technology fit (TTF) model several themes were identified. Evidence is provided on the extent to which the observed magnitudes have continued through the scope of this review and beyond.

This narrative synthesis provides several key insights: first, leaders and knowledge workers exhibited predominantly positive workflow transformation and task optimization experiences. An imperative discovery was made that AI integration leads to an increase in workforce demand, as when it is utilized in original contributions it results in increased efficiency (i.e., speed and quality of task completion) Das and Granados (2022). Additionally, Hao's research exhibited that Authentic leadership displays as an effective facilitator in progressing technological capabilities of followers (Hao et al., 2020). Change leadership played a positive moderating role in the relationship between AI adoption and employee work engagement (Wijayati et al., 2022).

Second prospect in the context of educational leadership was using AI systems as an extended brain during decision-making processes Wang (2021). However, according to (Jarrahi, 2018) if AI is to benefit humanity, it needs to respect human autonomy (Jarrahi, 2018). Furthermore, explains the effects of digital experiences on human autonomy are complex and inconsistent, leading to our third insight.

Advancements in replication of human cognitive abilities is continuing to enable autonomous agents to intercommunicate and exchange knowledge following certain protocols (Gonzalez-Rodriguez & Hernandez-Carrion, 2018). An imperative discovery that AI integration leads to an increase in

workforce demand, as when it is utilized in original contributions, it results in increased efficiency (i.e., speed and quality of task completion) (Das et al., 2022).

Fourth, a five-year analysis of scientific advancements elucidates the absence of trust in the Human-AI relationship by comparing the interactions between human and artificial cognitive intelligence (Abbass, 2019). Baum (2020) confirms Abbass's outcomes by adding three decision-making challenges in the design of AI systems based on social choice facets: 1) determining whose ethical views are included; 2) how to identify and measure these views; and 3) how to aggregate individual perspectives into a cohesive view that pilots AI conduct.

Our findings have implications for the following domains. Within the scope of this review, leaders and knowledge workers are faced with ongoing effects of AI adoption experiences. These experiences have complex and inconsistent impacts on their sense of autonomy (Jarrahi, 2018). Laacke et al. (2021), for instance, showcased the use of AI depression detectors (AIDDs) that can analyze data from social media to detect signs of mental disorders like depression. AIDDs can identify individuals who may be at risk of depression before they seek professional help. The ethical considerations of applying AIDDs on personal data of users posed challenges in extending this concept due to a breach of digital autonomy.

Overall, AI integrations have a predominantly positive impact on choices and workflow transformation. Concerns about AI's effect on choice authenticity are countered by its efficiency gains and the synergy between human and AI decision-making capabilities. This synergy, termed "hybrid decision-making," demonstrates the effective fusion of AI capabilities with human critical thinking, as evidenced by compelling cases. Hybrid decision-making, emphasized for its significance in achieving long-term autonomy balance, entails the integration of AI planning for task sequencing, daily task generation, and action prioritization (Kunze et al., 2018). Baum (2020) underscores the necessity for meticulous organizational decision-making during AI design, as delegating such decisions solely to AI is impractical.

## APPENDIX

### Content Assessment

No.	Study	Content Validation
1	(Abbass, 2019) Qualitative Australia	Human-Machine teams & Trustworthy AI AO
2	(Baum, 2020) Qualitative UK	Wholistic Organizational Ethics Regulation AO
3	(Bilan et al., 2022) Mixed Methods Lithuania	AI and Organizational Change PO
4	(Das et al., 2022) Quantitative India	Impact of AI driven choices in CRM PO

No.	Study	Content Validation
5	(Dobbe et al., 2021) The Netherlands-USA	Ethics Implication AO
6	(Bankins & Formosa, 2021) Qualitative Australia	Promote/Demote /Decisions/Skills PO & AO
7	(Fedorets et al., 2022) Germany	Direct impact of AI on Autonomy PO
8	(Hao et al., 2020) Quantitative Phuket	AI adoption and Performance: Individuals vs Team PO & AO Individuals –PO Teams– AO
9	(Jones, 2018) Qualitative USA	Long Term Autonomy Balance AO
10	(Kunze et al., 2018) Quantitative	Long-term robot Autonomy PO
11	(Sebastian Laacke et al., 2021) Quantitative USA	Digital autonomy and Privacy regulations AO
12	(Lawless et al., 2019) Quantitative USA	Human Machine teams PO
13	(Lin & Zhu, 2021) Qualitative China	Leveraging AI decision-making models PO
14	(Pescetelli, 2021) Qualitative USA	Bias-proved Hybrid Intelligence PO
15	(Ploug et al., 2021) Mixed Methods Denmark	Outcomes of AI-driven decisions in hybrid teams PO
16	(Smith & Green, 2018) Qualitative USA	Digital Leadership PO
17	(Unhelkar & Gonsalves, 2020) Qualitative USA	Predictive Data-driven capabilities of AI PO
18	(Yinying Wang, 2021) Qualitative USA	(Extended Brain) AI driven Decisions biases & Ethics PO & AO
19	(Wijayati et al., 2022) Quantitative Indonesia	AI, Performance & Work Engagement PO
20	(Di Vaio et al., 2020) Malaysia	Sustainable Business model PO
21	(Metcalf et al., 2019) Qualitative USA	Artificial swarm intelligence leveraging collective intelligence PO
22	(Barro & Davenport, 2019) Qualitative UK	Transforming Workflow PO
23	(Huang et al., 2019) Qualitative USA	Changing roles in Empathetic and Analytic roles PO
24	(Kaplan & Haenlein, 2019b) Qualitative USA	Leaders integrating AI technologies for hybrid work PO
25	(Haseeb et al., 2019) Quantitative Thailand	Process Optimising via Automation PO



No.	Study	Content Validation
26	(Duan et al., 2019) Qualitative UK	Problem Solving use of AI- Decision making systems PO
27	(Schneider & Leyer, 2019) Quantitative Germany	AI Adoption Impacting Decision-making Processes PO
28	(Jarrahi, 2018) Qualitative USA	Autonomy in digital environments AO

## REFERENCES

- Abbass, H. A. (2019). Social Integration of Artificial Intelligence: Functions, Automation Allocation Logic and Human-Autonomy Trust. *Cognitive Computation*, 11(2), 159–171. <https://doi.org/10.1007/s12559-018-9619-0>
- Anderson, J. H., & Honneth, A. (2005). Autonomy, vulnerability, recognition, and justice.
- Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30.
- Bankins, S., & Formosa, P. (2021). Ethical AI at work: The social contract for artificial intelligence and its implications for the workplace psychological contract. In *Redefining the Psychological Contract in the Digital Era* (pp. 55–72). Springer.
- Barro, S., & Davenport, T. H. (2019). People and machines: Partners in innovation. *MIT Sloan Management Review*, 60(4), 22–28.
- Baum, S. D. (2020). Social choice ethics in artificial intelligence. In (Vol. 35, pp. 165–176): Springer London.
- Bilan, S., Šuleř, P., Skrynnyk, O., Krajňáková, E., & Vasilyeva, T. (2022). Systematic bibliometric review of artificial intelligence technology in organizational management, development, change and culture [Article]. *Business: Theory and Practice*, 23(1), 1–13. <https://doi.org/10.3846/btp.2022.13204>
- Booyse, D., & Scheepers, C. B. (2023). Barriers to adopting automated organisational decision-making through the use of artificial intelligence. *Management Research Review*.
- Considine, J., Shaban, R. Z., Fry, M., & Curtis, K. (2017). Evidence based emergency nursing: designing a research question and searching the literature. *International Emergency Nursing*, 32, 78–82.
- Das, A. K., & Granados, C. (2022). FP-intuitionistic multi fuzzy N-soft set and its induced FP-Hesitant N soft set in decision-making [Article]. *Decision Making: Applications in Management and Engineering*, 5(1), 67–89. <https://doi.org/10.31181/dmame181221045d>
- Das, P., Perera, S., Senaratne, S., & Osei-Kyei, R. (2022). Paving the way for industry 4.0 maturity of construction enterprises: A state of the art review. *Engineering, Construction and Architectural Management* (ahead-of-print).
- Dencik, J., Goehring, B., & Marshall, A. (2023). Managing the emerging role of generative AI in next-generation business. *Strategy & Leadership*, 51(6), 30–36.
- Di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. (2020). Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, 121, 283–314.
- Dobbe, R., Thomas Krendl Gilbert, & Yonatan Mintz. (2021). Hard choices in artificial intelligence. *Artificial Intelligence*. 300. <https://doi.org/10.1016/j.artint.2021.103555>
- Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda [Article]. *International Journal of Information Management*, 48, 63–71. <https://doi.org/10.1016/j.ijinfomgt.2019.01.021>

- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, 62(1), 107–115.
- Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *Journal of the Medical Library Association: JMLA*, 106(4), 420.
- Fedorets, A., Kirchner, S., Adriaans, J., & Giering, O. (2022). Data on Digital Transformation in the German Socio-Economic Panel. *Jahrbücher für Nationalökonomie und Statistik*, 242(5–6), 691–705. <https://doi.org/doi:10.1515/jbnst-2021-0056>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Sci Rev. In*.
- Gonzalez-Rodriguez, D., & Hernandez-Carrion, J. R. (2018). Self-Organized Linguistic Systems: From traditional AI to bottom-up generative processes. *Futures*, 103, 27–34.
- Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, 19(2), 213–236. <https://doi.org/10.2307/249689>
- Hao, M., Lv, W., & Du, B. (2020). The influence mechanism of authentic leadership in artificial intelligence team on employees' performance. *Journal of Physics: Conference Series*.
- Haseeb, M., Hussain, H. I., Ślusarczyk, B., & Jermsittiparsert, K. (2019). Industry 4.0: A solution towards technology challenges of sustainable business performance. *Social Sciences*, 8(5), 154.
- Higgins, J. (2008). *Cochrane Handbook for Systematic Reviews of Interventions 5.0. 0*; Higgins J, Green S, editors. In: Chichester, UK: John Wiley & Sons Ltd.
- Huang, M.-H., Rust, R., & Maksimovic, V. (2019). The feeling economy: Managing in the next generation of artificial intelligence (AI). *California Management Review*, 61(4), 43–65.
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making [Article]. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>
- Jones, W. A. (2018). Artificial Intelligence And Leadership: A Few Thoughts, A Few Questions. *Journal of Leadership Studies*, 12(3), 60–62. <https://doi.org/10.1002/jls.21597>
- Kaplan, A., & Haenlein, M. (2019a). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence [Article]. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Kaplan, A., & Haenlein, M. (2019b). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25.
- Kunze, L., Hawes, N., Duckett, T., Hanheide, M., & Krajník, T. (2018). Artificial Intelligence for Long-Term Robot Autonomy: A Survey. *IEEE Robotics and Automation Letters*, 3(4), 4023–4030. <https://doi.org/10.1109/lra.2018.2860628>
- Laacke, S., Mueller, R., Schomerus, G., & Salloch, S. (2021). Artificial Intelligence, Social Media and Depression. A New Concept of Health-Related Digital Autonomy. *American Journal of Bioethics*, 21(7), 4–20. <https://doi.org/10.1080/15265161.2020.1863515>
- Lawless, W. F., Mittu, R., Sofge, D., & Hiatt, L. (2019). Artificial intelligence, autonomy, and human-machine teams—interdependence, context, and explainable AI. *Ai Magazine*, 40(3), 5–13.

- Lin, S., & Zhu, W. (2021). How subjective information with AI for digital revolution [Article]. *Journal of Intelligent & Fuzzy Systems*, 1–11. <https://doi.org/10.3233/jifs-211624>
- Metcalfe, L., Askay, D. A., & Rosenberg, L. B. (2019). Keeping Humans in the Loop: Pooling Knowledge through Artificial Swarm Intelligence to Improve Business Decision Making [Article]. *California Management Review*, 61(4), 84–109. <https://doi.org/10.1177/0008125619862256>
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199.
- Pescetelli, N. (2021). A brief taxonomy of hybrid intelligence. *Forecasting*, 3(3), 633–643.
- Petticrew, M., Arai, L., Roberts, H., Britten, N., & Popay, J. (2009). Testing methodological guidance on the conduct of narrative synthesis in systematic reviews. *Evaluation*, 15(1), 49–73.
- Ploug, T., Sundby, A., Moeslund, T. B., & Holm, S. (2021). Population preferences for performance and explainability of artificial intelligence in health care: choice-based conjoint survey. *Journal of Medical Internet Research*, 23(12), e26611.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version*, 1(1), b92.
- Prunkl, C. (2022). Human autonomy in the age of artificial intelligence. *Nature Machine Intelligence*, 4(2), 99–101.
- Raz, J. (1986). *The morality of freedom*. Clarendon Press.
- Roessler, B. (2021). *Autonomy: An Essay on the Life Well-Lived*. John Wiley & Sons.
- Schneider, S., & Leyer, M. (2019). Me or information technology? Adoption of artificial intelligence in the delegation of personal strategic decisions [Article]. *Managerial & Decision Economics*, 40(3), 223–231. <https://doi.org/10.1002/mede.2982>
- Smith, A. M., & Green, M. (2018). Artificial intelligence and the role of leadership. *Journal of Leadership Studies*, 12(3), 85–87.
- Unhelkar, B., & Gonsalves, T. (2020). Enhancing artificial intelligence decision making frameworks to support leadership during business disruptions. *IT Professional*, 22(6), 59–66.
- Wang, Y. (2021). Artificial intelligence in educational leadership: a symbiotic role of human-artificial intelligence decision-making. *Journal of Educational Administration*, 59(3), 256–270.
- Wang, Y. (2021). Probabilistic Modeling of Information Dynamics in Networked Cyber-Physical-Social Systems [Article]. *IEEE Internet of Things Journal*, 8(19), 14934–14947. <https://doi.org/10.1109/JIOT.2021.3072893>
- Wijayati, D. T., Rahman, Z., Fahrullah, A., Rahman, M. F. W., Arifah, I. D. C., & Kautsar, A. (2022). A study of artificial intelligence on employee performance and work engagement: The moderating role of change leadership. *International Journal of Manpower*, 43(2), 486–512. <https://doi.org/10.1108/IJM-07-2021-0423>

# Ergonomics and Collaborative Robotics: The Synergy to Prevent Workload in Industrial Assembly Tasks

Ana Colim<sup>1,2</sup>, Estela Bicho<sup>2</sup>, André Cardoso<sup>2</sup>, Carla Alves<sup>2</sup>,  
Amin Salimi<sup>1</sup>, Duarte Fernandes<sup>1</sup>, Nuno Sousa<sup>2</sup>, Nélson Costa<sup>2</sup>,  
Paula Carneiro<sup>2</sup>, Luís Louro<sup>2</sup>, Sérgio Monteiro<sup>2</sup>,  
João Gaspar Cunha<sup>2</sup>, and Pedro Arezes<sup>2</sup>

<sup>1</sup>DTx – Associação Laboratório Colaborativo em Transformação Digital, 4800–058  
Guimarães, Portugal

<sup>2</sup>Centro Algoritmi, School of Engineering, University of Minho, 4800–058 Guimarães,  
Portugal

## ABSTRACT

The synergy between Ergonomics and Collaborative Robotics represents a crucial and emergent framework to achieve innovative and human-centered assembly workstations. Among the various I4.0 technologies, Human-Robot Collaboration (HRC) has emerged as a cutting-edge solution to address ergonomic challenges while enhancing manufacturing productivity. Despite this potential, there is a significant gap in empirical studies on the implementation of these technologies, particularly in the context of integrating ergonomic requirements. This study aims to present real-world applications of HRC in assembly tasks performed by workers with musculoskeletal complaints, as well as to indicate future research challenges in this field. Our previous work involved the digital transformation of two real-industry workstations, mitigating musculoskeletal risk factors by the HRC implementation according to a human-centered design. During this work, ergonomic assessments of the workstations before and after the introduction of HRC were performed (e.g. the application of Inertial Measurement Units to quantify the Rapid Upper Limb Assessment score). Key performance indicators, such as production rates, were measured through time studies and direct observation. To gather additional insights into workers' well-being, questionnaires were applied. This multi-method approach revealed that the workstations with HRC resulted in (i) reduced production costs, (ii) improved ergonomic conditions, and (iii) enhanced worker well-being. Based on these previous findings, our current research is focused on creating a flexible integrated robotic system capable of executing a shared human-robot task plan and reasoning about the working conditions. This system will integrate a real-time ergonomics assessment system to track and evaluate postures and other physical indicators and extrapolate corrective measures. In our research, we intend to highlight that this synergy between Ergonomics and Collaborative Robotics underscores the continuous improvement of industrial processes while ensuring sustainable productivity and worker well-being.

**Keywords:** Ergonomics, Human-robot collaboration, Human-centered design, Assembly tasks

## INTRODUCTION

The current competition in the global labor market requires companies to implement flexible, efficient, and sustainable workstations. However, in the manufacturing industry, there are still several problems, exposing workers to cognitive and physical overload, awkward postures, monotonous and repetitive work (Beuß et al., 2019; Ding et al., 2023; Guimarães et al., 2015). These risk factors can negatively affect workers' health and productivity, compromising their well-being and representing a barrier to the creation of sustainable jobs. For these reasons, there is a pressing need for further research focused on the digital transformation of workstations for adaptive production systems in the manufacturing industry, which promotes inclusion and well-being at work (Nahavandi, 2019; Xu et al., 2021). This strategy aligns with the Horizon Europe 2021–2027 agenda, focusing on the challenges listed by the European Commission: “*Industry 5.0- Towards a sustainable, human-centric and resilient European industry*” (Breque et al., 2021).

In the manufacturing industry of the future, Human-Robot Collaboration (HRC) is identified as one of the main technologies to facilitate industrial evolution (John et al., 2020). Foreseeing the reduction of human workload, the Work-related Musculoskeletal Disorders (WMSD) risk mitigation, as well as enhancing safety, quality and productivity, collaborative solutions with human-robot interaction have been proposed (Gualtieri et al., 2021; Patil et al., 2023). In contrast to conventional and non-collaborative industrial robotics, in the HRC scenarios, robots are not confined to certain restricted areas, potentiating the opportunity to combine the benefits of automation with human cognitive flexibility, manipulation dexterity and other skills. However, safety and human well-being are the main challenges in this context, still requiring deep research, mainly in industrial environments, to test and validate the implementation of industrial HRC (Cardoso et al., 2021).

Therefore, the inclusion of the ergonomics scientific area in the design of innovative workstations in the industry has assumed an emerging relevance. This area focuses on the design and assessment of tasks, workstations, machines, tools, and systems, aiming to achieve safe, effective, and productive human activities, considering human capabilities, limitations, and behaviors (IEA and ILO, 2021). One should note that attributing low importance to human factors is associated with decreased productivity, increased absenteeism, worsening occupational accidents, and WMSD rates, which negatively impact the economies of industrialized countries. Currently, WMSD is an important occupational health problem, and its occurrence is often associated with physical factors (such as awkward postures and repeated efforts) (Chen et al., 2018), but also with other factors (e.g., individual limitations and age) (Peruzzini and Pellicciari, 2017).

Motivated by this, the current article aims to present the work developed by our research team focused on the human-centered design of novel assembly workstations with HRC. Our research has been grounded in the synergy between ergonomics and collaborative robotics, and it could be exemplified in the three use cases presented in this article. We believe that

this dissemination can significantly contribute to the body of knowledge for researchers and industry practitioners dedicated to integrating HRC into the workstations, to develop collaborative assembly tasks that prioritize ergonomics and human factors.

## METHODOLOGY

Our research team is composed of researchers from DTx Colab (*Associação Laboratório Colaborativo em Transformação Digital*) and the University of Minho – Portugal, comprising the scientific areas of Ergonomics, Robotics and Computer Vision. The current article summarized the digital transformation of three real-industry assembly workstations (use cases posteriorly presented), mitigating musculoskeletal risk factors by the HRC implementation. During this work (over the last four years), ergonomic assessments of the workstations before and after the introduction of HRC were performed, applying Inertial Measurement Units (IMU) to quantify the Rapid Upper Limb Assessment score (Colim et al., 2021a; McAtamney and Corlett, 1993). Key performance indicators, such as production rates and spending on consumables (e.g. glue), were measured through time studies and direct observation. To gather additional insights into workers' well-being, questionnaires were applied to estimate the physical and cognitive workload (Borg, 1990; Hart, 2006), as well as the human trust (Jian et al., 2000). This multi-method approach enabled us to achieve a human-centered design that also took into account the industry's real needs. Moreover, all stakeholders involved in these use cases, including the assembly workers, were engaged from the early design phases through the adoption of a participatory ergonomics intervention (Guimarães et al., 2015). Throughout our research, all participants signed an Informed Consent Term in agreement with the Committee of Ethics for Research in Social and Human Sciences of the University of Minho (approval number CEICSH 038/2020), respecting the Declaration of Helsinki.

## RESULTS AND ANALYSIS

### Use Case 1

The first use case was focused on an assembly workstation (Figure 1), characterized by the performance of repetitive manual tasks, in which preforms are produced, which are elements that will compose, for example, table tops. As a result of the ergonomic assessment, it was identified that the task of applying glue was the most critical from the point of view of effort and repetitiveness, especially for the hand-wrist system. In addition to the existing musculoskeletal problems and complaints, accidental burns with hot glue were important factors to be corrected, motivating the transformation of this workstation.

Therefore, for the future work cycle, the requirements of work organization were defined, selecting the task of applying glue to be replaced by the robotic system. For the design of the new workstation, anthropometric data were considered and a complete review of standards and legal documents was carried out at this stage, to define a set of requirements

in terms of safety. This design phase and requirements definition are present in the paper Ana Colim et al. (2021). As we demonstrated in Colim et al. (2021b), compared to the previous manual workstation, the improvements are significant in terms of: improvement in working conditions and workers' satisfaction; an increase of the workers' trust and confidence to work with cobots; improvement in productivity (7% less cycle time) and the reduction of glue consumption (29% less glue); a significant reduction in musculoskeletal risk (24% to 39% reduction of RULA scores).



**Figure 1:** Assembly workstation before (on the left) and after (on the right) HRC implementation.

## Use Case 2

The second use case was also developed and implemented in the same furniture manufacturing company as the first. The main goal of use case 2 was creating a robotic system to replace a repetitive, monotonous, and manual pick-and-place and palletization of “preforms” (pieces that incorporate the structure of tabletops or shelves, represented in Figure 2). This new system performs three important roles, namely: autonomous palletization (by a UR10e); quality control (throughout a computer vision system); and human-aware navigation (including an autonomous mobile robot MiR 200 with a mobile cart to move the UR10e) between different workstations according to production needs.

Relatively to the human-aware navigation, previously to the shopfloor implementation, several courtesy cues (such as stopping, decelerating, retreating, and retreating and moving aside) were tested and the findings published (Alves et al., 2022). These robotic cues could affect human trust and behavior, being this understanding crucial for improving human-robot collaboration in shared workspaces. The robot used in the study was the MiR 200, equipped with various sensors and programmed to perform the specified courtesy cues at a controlled industrial intersection. Participants' trust and perception of the robot's cues were measured using an adapted Human Trust in Automation (HTA) questionnaire (Jian et al., 2000), and their hesitation behaviors were analyzed through video recordings.

Key findings revealed no significant difference in participants' trust levels across different courtesy cues. However, the legibility of the robot's actions varied significantly. The retreating cue was the most legible from the forward view, while the decelerate cue led to the least hesitation among participants. The study suggests that while trust in the robot remained stable, the clarity of

the robot's behaviors and the reduction of human hesitation were influenced by the specific courtesy cues. These insights underscore the importance of designing HRC scenarios with clear and predictable movement patterns to enhance human-robot interaction in industrial environments.



**Figure 2:** Representation of a "preform" and the workstation before (on the left) and after (on the right) CR implementation.

### Use Case 3

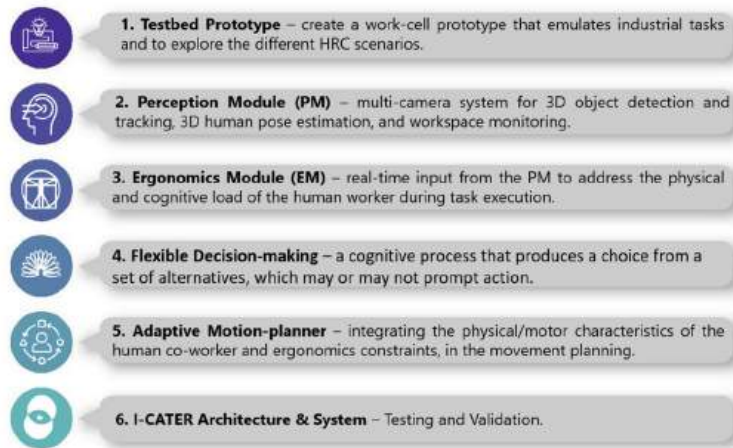
Finally, the third use case consists in an ongoing research project, with the title: *I-CATER – Intelligent Robotic Coworker Assistant for Industrial Tasks with Ergonomics Rationale*. With the current research, we intend to achieve the following outcomes:

- 1) Real-time ergonomics assessment system to track and evaluate, postures as well as cognitive variables, and extrapolate corrective measures;
- 2) Reinforcement learning system to map the corrective measures to robotic system inputs;
- 3) A flexible integrated robotic system capable of executing a shared human-robot task plan and reasoning about the working conditions and human characteristics.

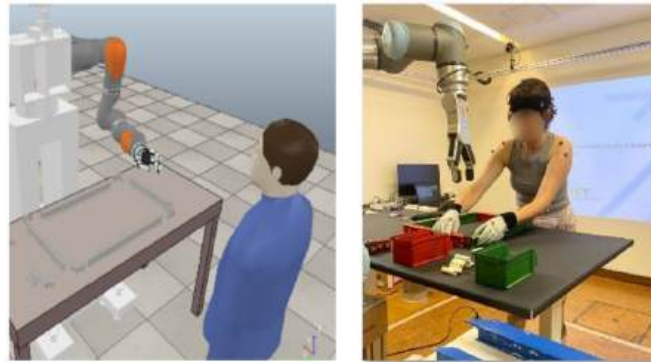
To achieve these goals, we have six main activities as represented in Figure 3.

The testbed prototype was inspired by a real-industry scenario of assembly frames for windows, where an ergonomic assessment was performed, including the application of the NASA Task Load Index (NASA-TLX) (Hart, 2006) and RULA method (McAtamney and Corlett, 1993). The ergonomic assessment demonstrated that this assembly task involves teamwork (two workers), a high difficulty level in terms of work content, and a significant prevalence of WMSD risk factors (e.g. awkward postures and repetitive movements). Before the testbed prototype creation, in a virtual context, the future collaborative workstation was tested, and a set of safety and ergonomic requirements (including anthropometric data) were defined for the design. The simulation scenario includes a Kuka collaborative robotic arm attached to a humanoid torso, with anthropomorphic dimensions, to make collaborations more human-like and fluent. Then, in our laboratory, a real-world prototype was created (Figure 4). This is a key premise for the development and testing of the collaborative robotic system.





**Figure 3:** Research activities of the use case 3.



**Figure 4:** Simulation in CoppeliaSim (on the left), and testbed prototype (on the right).

To evolve the state of the art of current collaborative robotics in the manufacturing industry, we propose to design and implement several modules/components, which, as a whole, will constitute a robotic system and platform for natural and efficient human-robot joint action. The high-level modules that compose the proposal and the robot system are:

1. **Flexible and human-aware decision-making:** to be a useful and efficient partner the robotic co-worker will be endowed with decision-making at different levels – task level, intention inference, and action selection. Because the robot will have to master various assembly tasks and interaction scenarios with different human partners, such decision-making processes will be intertwined with learning and adaptation processes supported by a knowledge base and a semantic environment. The innovations will be: the integration of ordinal and temporal information that will allow the robot to decide “*what sub-task is to be done next*” and “*when to do it*” in each given context; the integration of error communication and implementation of backup actions as

corrective measures; learn how to adapt behavior to different human co-workers. These challenges are addressed in Activity 4 (previously identified in Figure 3).

2. Human-aware motion planning (Activity 5): in human-robot shared tasks, when observing the robot's movement, the human should be able to understand the robot's actions, infer the underlying motor intention, and predict its outcome, effortlessly. Hence, the robot's decision to perform a specific complementary action will be translated into movements that are safe, fluent, smooth, and human-like/human-aware. Fundamental principles of human-motor control can be used to adopt human-like arm-hand motion planning strategies, that are meant to improve human comfort and enhance human-robot interaction and collaboration. The decisions at the motor level, and the generated movements are key elements to improve ergonomics. They should be tuned to the characteristics of the human partner and underlying ergonomics constraints.
3. Real-time assessment of human workload (Activity 3): to make decisions and actions that enhance human safety, well-being, and productivity during human-robot joint action, the robot needs to receive real-time measures of human activity. To this end, we will develop an ergonomics module for real-time assessment of the workload to which the human co-worker is exposed. New, will be: creating an assessment framework appropriate to the HRC tasks, that will integrate physical and cognitive ergonomic criteria; mapping of identified risk factors to ergonomic corrective measures; mapping of corrective actions to adequate robot behaviors or constraints, based on the ergonomic criteria and assessments.
4. General real-time perception module (Activity 2): HRC in shared tasks usually takes place in very dynamic and unconstrained scenarios. Thus, we will develop a perception module that allows, in real-time, 3D detection and tracking of task-related objects of interest, 3D human pose estimation, and workspace monitoring. The research focus will be on the generalization of the developed components, to add to the versatility and flexibility of the solution.

Summarily, our current research aims to create the first intelligent and human-aware ergonomic cobot. This next generation of robotic co-workers will enhance human capabilities, ergonomic criteria, and efficiency in assembly industrial tasks.

## CONCLUSION

The synergy between Ergonomics and Collaborative Robotics presents a promising avenue for the development of human-centered assembly workstations. This article highlights the transformative potential of integrating HRC in industrial assembly tasks, according to human-centered design, significant improvements in ergonomic conditions and workers' well-being were observed.

Our research underscored the importance of a multi-method approach that combines ergonomic assessments, real-time data collection, and workers' engagement to create adaptive and safe work environments. The three use cases detailed in this study demonstrated the practical benefits of HRC in mitigating musculoskeletal risk factors and enhancing overall work conditions. Furthermore, the ongoing research project, I-CATER, aims to advance this field by developing an intelligent and flexible robotic coworker capable of real-time ergonomics assessment and adaptive task planning. This research seeks to create a next-generation ergonomic cobot that enhances human capabilities and ensures sustainable productivity.

## ACKNOWLEDGMENT

This work was supported by FCT within Project I-CATER: Intelligent Robotic Coworker Assistant for industrial Tasks with an Ergonomics Rationale, reference PTDC/EEI-ROB/3488/2021.

## REFERENCES

- Alves, C., Cardoso, A., Colim, A., Bicho, E., Braga, A. C., Cunha, J., Faria, C., Rocha, L. A., 2022. Human–Robot Interaction in Industrial Settings: Perception of Multiple Participants at a Crossroad Intersection Scenario with Different Courtesy Cues. *Robotics* 11. <https://doi.org/10.3390/robotics11030059>
- Beuß, F., Sender, J., Flügge, W., 2019. Ergonomics simulation in aircraft manufacturing – Methods and potentials. *Procedia CIRP* 81, 742–746. <https://doi.org/10.1016/j.procir.2019.03.187>
- Borg, G., 1990. Psychophysical scaling with applications in physical work and the perception of exertion. *Scand J Work Environ Health* 16, 55–58. <https://doi.org/10.5271/sjweh.1815>
- Breque, M., De Nul, L., Petridis, A., 2021. Industry 5.0: Towards a sustainable, human-centric and resilient European industry. <https://doi.org/10.2777/308407>
- Cardoso, A., Colim, A., Bicho, E., Braga, A. C., Menozzi, M., Arezes, P., 2021. Ergonomics and human factors as a requirement to implement safer collaborative robotic workstations: A literature review. *Safety* 7. <https://doi.org/10.3390/safety7040071>
- Chen, D., Cai, Y., Cui, J., Chen, J., Jiang, H., Huang, M., 2018. Risk factors identification and visualization for work-related musculoskeletal disorders with wearable and connected gait analytics system and kinect skeleton models. *Smart Health* 7–8, 60–77. <https://doi.org/10.1016/j.smhl.2018.05.003>
- Colim, A., Cardoso, A., Arezes, P., Braga, A. C., Peixoto, A. C., Peixoto, V., Wolbert, F., Carneiro, P., Costa, N., Sousa, N., 2021a. Digitalization of musculoskeletal risk assessment in a robotic-assisted assembly workstation. *Safety* 7. <https://doi.org/10.3390/safety7040074>
- Colim, Ana, Faria, C., Cunha, J., Oliveira, J., Sousa, N., Rocha, A., 2021. Physical Ergonomic Improvement and Safe Design of an Assembly Workstation through Collaborative Robotics. *Safety* 7, 2–18. <https://doi.org/10.3390/safety7010014>
- Colim, A., Morgado, R., Carneiro, P., Costa, N., Faria, C., Sousa, N., Rocha, L. A., Arezes, P., 2021b. Lean Manufacturing and Ergonomics Integration: Defining Productivity and Wellbeing Indicators in a Human – Robot Workstation 13, 1931. <https://doi.org/https://doi.org/10.3390/su13041931>

- Ding, X., Guan, Z., Liu, N., Bi, M., Ji, F., Wang, H., Zhang, X., Liu, B., Niu, D., Lan, T., Xie, T., Li, J., Yan, T., 2023. Prevalence and risk factors of work-related musculoskeletal disorders among emerging manufacturing workers in Beijing, China. *Front Med (Lausanne)* 10. <https://doi.org/10.3389/fmed.2023.1289046>
- Gualtieri, L., Rauch, E., Vidoni, R., 2021. Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review. *Robot Comput Integr Manuf* 67, 101998. <https://doi.org/10.1016/j.rcim.2020.101998>
- Guimarães, L., Anzanello, M. J., Ribeiro, J. L. D., Saurin, T. A., 2015. Participatory ergonomics intervention for improving human and production outcomes of a Brazilian furniture company. *Int J Ind Ergon* 49, 97–107. <https://doi.org/10.1016/j.ergon.2015.02.002>
- Hart, S. G., 2006. NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society* 904–908. <https://doi.org/10.1177/154193120605000909>
- IEA, ILO, 2021. Principles and guidelines for human factors/ergonomics (HFE) design and management of work systems. International Labour Organization publications, Geneva, Switzerland.
- Jian, J.-Y., Bisantz, A. M., Drury, C. G., 2000. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *Int J Cogn Ergon* 4, 53–71. [https://doi.org/10.1207/s15327566ijce0401\\_04](https://doi.org/10.1207/s15327566ijce0401_04)
- John, K. K., Adarsh, S. N., Pattali, V., 2020. Workers to super workers: A brief discussion on important technologies for industry 5.0 manufacturing systems. *AIP Conf Proc* 2311, 0–6. <https://doi.org/10.1063/5.0034521>
- McAtamney, L., Corlett, N., 1993. RULA: A survey method for the investigation of work-related upper limb disorders. *Appl Ergon* 24, 91–99.
- Nahavandi, S., 2019. Industry 5.0 — A Human-Centric Solution. *Sustainability* 11, 4371. <https://doi.org/10.3390/su11164371>
- Patil, S., Vasu, V., Srinadh, K. V. S., 2023. Advances and perspectives in collaborative robotics: A review of key technologies and emerging trends. *Discover Mechanical Engineering* 2. <https://doi.org/10.1007/s44245-023-00021-8>
- Peruzzini, M., Pellicciari, M., 2017. Advanced Engineering Informatics A framework to design a human-centred adaptive manufacturing system for aging workers. *Advanced Engineering Informatics* 33, 330–349. <https://doi.org/10.1016/j.aei.2017.02.003>
- Xu, X., Lu, Y., Vogel-Heuser, B., Wang, L., 2021. Industry 4.0 and Industry 5.0—Inception, conception and perception. *J Manuf Syst* 61, 530–535. <https://doi.org/10.1016/j.jmsy.2021.10.006>

# How Many Robots Is Too Many? Findings About Single-Human Multiple-Robot Systems

**Estefany Rey-Becerra and Sascha Wischniewski**

Federal Institute for Occupational Safety and Health, Dortmund, Germany

## ABSTRACT

Mobile robots are increasingly being used to perform tasks that are difficult for humans to reach. Due to their high degree of autonomy, a human can control multiple robots through a graphical user interface, which is called one-human multiple-robot (SHMR) system. However, information about them is scarce. The current literature review synthesizes the features of SHMR systems related to safety and health at work for the operator and system performance. A total of 658 records were identified through an exploratory search since 2000, and 35 were selected to meet the inclusion criteria. The characteristics were consolidated and provide valuable insights about field of application, team composition, and reported outcomes. Future research can focus on exploring these systems in industries that have not yet been studied, or it can examine the impact of individual operator characteristics on these systems.

**Keywords:** Human-robot interaction, Human-computer interaction, Multiple robots

## INTRODUCTION

*Mobile robots* are automatic machines with locomotion mechanisms and a sense of position that enables to navigate a space with minimal intervention (Rubio et al., 2019, Siegwart and Nourbakhsh, 2004). As an example are zoomorphic robots that can be used for remote inspection (BostonDynamics, 2023). These robots are particularly useful in areas that are difficult for humans to reach, such as in a natural disaster (Casper and Murphy, 2003), or for progress monitoring, such as in construction (Kim et al., 2022), among other fields. In addition, these robots do not require a human operator on board to perform their intended tasks and can be remotely controlled (Huang, 2004). Depending on their purpose, they can navigate on the ground (Unmanned Ground Vehicle – UGV), in the water (Unmanned Surface/Underwater Vehicles – USV/UUV), or in the air (Unmanned Aerial Vehicle – UAV, also known as drone) (Balestrieri et al., 2021). A group of these robots is called a swarm, which typically follows a single command autonomously (Wilson et al., 2023).

Autonomy in robots is required to support effective performance (sense an environment, plan based on it, and act in it to reach a goal) without external control (Beer et al., 2014). However, even with the current level of artificial intelligence, these robots are not fully autonomous and still require humans to evaluate tasks, modify goals, or intervene when the system

fails (Patel et al., 2022). With a high degree of autonomy, an operator can coordinate multiple robots, which is referred to as *Single-Human Multiple-Robots (SHMR) systems* (Wong et al., 2011). In general, SHMR systems are used for resource optimization, e.g., reduce labor costs, and improve system performance, e.g., more goals in less time (Wong and Seet, 2017). Among the levels of shared interaction, we focus on the collective structure on supervising multiple robots, where each robot has a task assigned directly by the operator (Yanco and Drury, 2002).

In order for humans to interact with multiple robots, a graphical user interface (GUI) is used as an intermediary (Lewis, 2013). This means that the relationship between humans and multiple robots is established through one or more screens. Therefore, it is important to consider the field of human-computer interaction (HCI) (Kawamura et al., 2003), especially to understand how this interaction affects human performance (Lewis, 2013). The use of robots interacting with human workers raises new safety and health concerns that must be addressed in order to ensure the safe use of robotic systems in the workplace (Murashov et al., 2016). In the case of SHMR systems, some scholars have found that the human workload may increase with each additional robot (Adams, 2009, Trouvain and Wolf, 2002, Velagapudi et al., 2008), and that switching attention from one robot to another may reduce supervisor's awareness (Prewett et al., 2009), leading to a greater tendency to fail to complete critical tasks (Wickens, 2008).

Regarding system effectiveness, some scholars have proposed equations to estimate the number of robots an operator can effectively control. Olsen Jr. & Wood (2004) and Crandall et al. (2005) developed the fan-out equation, which is the ratio between the effective time a robot can operate without human intervention and the time required by the operator to interact with a robot. Crandall & Cummings (2007) and Cummings & Mitchell (2008) complemented the equation with the waiting times caused by the operator, including interaction with other robot, human decision-making queue, and loss of situation awareness. Later, Breslow et al. (2014) incorporated the time available for humans to solve a robot problem before a damage considering that several robots are likely to require human attention at the same time. Parallel, Boussemart and Cummings (2011) used a model with Markov chains to create a predictor of operator actions in the area of human-monitored automated systems. Understanding the threshold at which the number of robots becomes too large for a single operator to effectively manage is critical to the design and implementation of these systems. To this end, it is essential to examine the limits of the system considering the relevant user states, which include workload, engagement, situational awareness, attention, fatigue, and emotional state (Schwarz et al., 2014), in particular by interacting with the robots through a GUI.

Some reviews have been made so far. Prewett et al. (2009) and Lewis (2013) performed literature search that involved remote human supervision on multiple unmanned vehicles, its performances indices and human effects. However, the literature is outdated. A more recent meta-analysis is done by Hocraffer and Nam (2017), who evaluated the state of research of mobile robots, but they focus on unmanned aerial vehicle swarm only. One more specific is done by Moud et al. (2018) in the construction sector. And one

more broadly is done by Chen and Barnes (2021), who cover the human-robot interaction in a detailed and comprehensive manner, without focusing on the SHMR systems of interest. Therefore, this literature review seeks to comprehend the features of SHMR systems that researchers have reported in any task or industry, excluding the military sector. This manuscript provides a synthesis of the main characteristics such as the field of application, the number and type of robots and outcome measurements related to the operator and system performance. The findings will help practitioners to develop effective SHMR systems.

## METHODOLOGY

We conducted a literature review in September 2023 following the “Methodology for developing evidence-informed management knowledge” by Tranfield et al. (2003), divided into three stages. In the first stage, planning the review, we focused on the research question: *What are the characteristics of SHMR systems?* Here, characteristics refer to field of application, team composition and reported outcomes. Inclusion criteria were systems with one human and multiple highly automated robots, mobile robots coexisting with the human, and the human as operator or supervisor. Multiple humans, industrial manipulators, swarms with a common goal, interactions without a GUI, technical studies focused on algorithms, and the military sector were excluded.

In the second stage of conducting the review, we searched on the Ebsco Discovery Service databases (EBSCO, 2023) using the following keywords within our scope: (*“Highly automated robot” OR “mobile robots” OR “Remotely operated vehicle” OR “Autonomous underwater vehicle” OR “Unmanned ground vehicle” OR “Unmanned aerial vehicle”*) AND (*“Human-robot interaction” OR “Human-robot interface” OR “Human robot interaction” OR “Human robot interface”*) NOT (*Militar\**). Only articles and conference proceedings published in English in peer-reviewed journals from the year 2000 onward that included experiments with more than two participants were included. We did not include short reports, conference posters, or abstracts.

A total of 444 records were found in the literature search directly from the EBSCO database. The references of the papers were checked and a further 214 records were identified. After the first scan, 134 documents were excluded as duplicates or irrelevant. After screening, 86 reviews on robotics, 32 articles on the military field, and 245 on other topics were discarded. This left 161 eligible manuscripts. After reading all the papers, 28 were excluded because they dealt with a single robot, 47 because they were technical, and 50 because they did not meet the criteria, leaving 16 journal articles and 19 proceedings to be included. Data from 35 documents were extracted and classified according to the following aspects:

- Field of application: industry where the SHMR system was tested;
- Team composition: number and type of robots (e.g. UGV, USV/UUV, UAV).
- Reported outcomes: operator and system performance (Steinfeld et al., 2006).

Finally, in the third, reporting and dissemination stage, we answer the research question in the next section, which provides a descriptive analysis of the findings, and then we critically evaluate the sources to identify and discuss strengths and weaknesses. Later, we consolidate the results of the literature review and derive a checklist for developing effective SHMRS. In the end, our conclusions provide insights into how humans can effectively manage multiple robots simultaneously.

## RESULTS

### Team Composition

Table 1 shows the team composition that the different authors have explored, including the number and type of robots. Around 63% of the studies used UGV only, 17% drones only, and 11.4% a combination of both. Just two authors were found to used different mobile robots: Glas et al. (2012) who tested social robots for monitoring guides in a shopping mall and Sellner et al. (2006a), (2006b) who used a combination of Roving eye, UGV and a Crane robot to simulate remote assembly in construction sites.

About 49% of the records investigated the influence of the operator-robot ratio on system performance. Adams (2009), Humphrey et al. (2007), Velagapudi et al. (2008) evaluated robot team composition to see the effects on task and operator performance. Chandarana et al. (2021) used four combinations to simulate costs in different scenarios. Crandall and Cummings (2007) tested the usefulness and appropriateness of different metrics for SHMR. Other authors conducted mixed designs to evaluate performance of the SHMR system, varying the number of robots and different factors such as coordination techniques (Glas et al., 2012), exploration modes (Vilela et al., 2013, Wang et al., 2009b), operation modes (Hong et al., 2019), different task conditions (Lewis et al., 2010), different environments (Trouvain et al., 2003, Trouvain and Wolf, 2002), interaction modes (Villani et al., 2020), video displays (Wang et al., 2009a), automation and latency level (Khasawneh et al., 2019), and automation reliability and scheduling discipline (Chien et al., 2013).

**Table 1.** Team composition in the experiments per each author.

ID	Author	Number of robots	Type of robots
[1]	Adams (2009)	1H: 1,2,4R	UGVs
[2]	Brooks et al. (2011)	1H:12R	UGVs
[3]	Chandarana et al. (2021)	1H: 5,8,10,15R	UGVs
[4]	Chien et al. (2013)	1H:3,6R	UGVs
[5]	Crandall and Cummings (2007)	1H:2,4,6,8R	UGVs + UAVs

(Continued)



**Table 1.** Continued

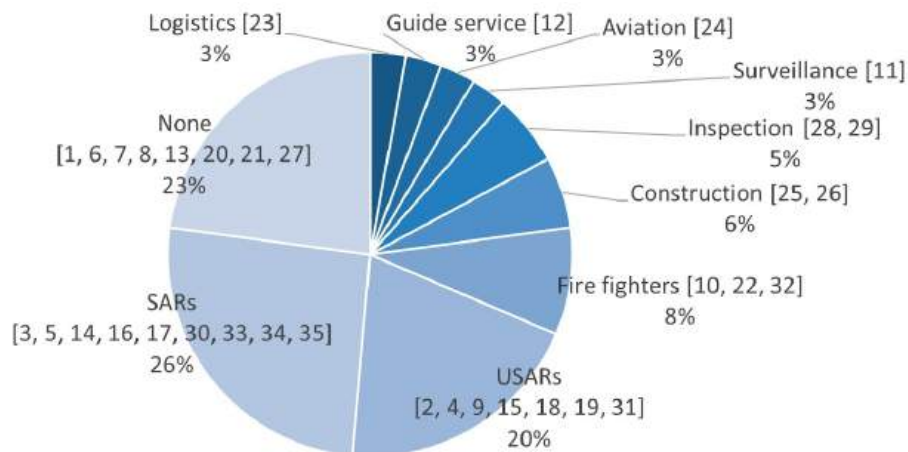
ID	Author	Number of robots	Type of robots
[6]	Crandall et al. (2005)	1H:3R	UGVs
[7]	Dietz et al. (2017)	1H:10R	UGVs
[8]	Donmez et al. (2010)	1H:5R	UGVs + UAVs
[9]	Fooladi Mahani et al. (2021)	1H:3R	UAVs
[10]	Frische and Ludtke (2013)	1H:3R	UAVs
[11]	Fuchs et al. (2014)	1H:4R	UAVs
[12]	Glas et al. (2012)	1H:2,3,4R	Social robots
[13]	Goodrich et al. (2007)	1H:3R	UGVs
[14]	Hong et al. (2017)	1H:4R	UAVs
[15]	Hong et al. (2019)	1H:5,10,15,20R	UGVs
[16]	Humphrey et al. (2007)	1H:6,9R	UGVs
[17]	Khasawneh et al. (2019)	1H:1,2R	UGVs
[18]	Lewis et al. (2010)	1H:4,8,12R	UGVs
[19]	Lewis et al. (2014)	1H:6R	UGVs
[20]	Olsen Jr and Wood (2004)	1H:18R	UGVs
[21]	Ratwani et al. (2010)	1H:5R	UAVs
[22]	Roldán et al. (2017)	1H:2R	UGVs + UAVs
[23]	Rosenfeld et al. (2017)	1H:3R	UGVs
[24]	Ruiz et al. (2015)	1H:3R	UAVs
[25]	Sellner et al. (2006a)	1H:3R	Roving eye + UGV + Crane
[26]	Sellner et al. (2006b)	1H:3R	Roving eye + UGV + Crane
[27]	Setter et al. (2015)	1H:5R	UGVs
[28]	Trouvain and Wolf (2002)	1H:2,4,8R	UGVs
[29]	Trouvain et al. (2003)	1H:1,2,4R	UGVs
[30]	Velagapudi et al. (2008)	1H:4,8,12R	UGVs
[31]	Vilela et al. (2013)	1H:1,2,3,4R	UGVs
[32]	Villani et al. (2020)	1H:2,3,12R	UGVs + UAVs
[33]	Wang and Lewis (2007)	1H:3R	UGVs
[34]	Wang et al. (2009a)	1H:4,8,12R	UGVs
[35]	Wang et al. (2009b)	1H:4,8,12R	UGVs

### Fields of Application

Most studies focused on search and rescue (SARs) and urban search and rescue (USARs), in particular for victim detection. Others authors researched on fire fighters, assembly in construction, inspections, surveillance, aviation, guide service and logistics. Eight studies did not specify the application of domain (see also Figure 1).

### Reported Outcomes

According to Steinfeld et al. (2006), common metrics for human-robot interaction assess the performance rather for the operator, the system or the robot. From a human factor perspective, operator performance in human-robot interaction is influenced by a combination of cognitive, physical, and psychological factors (Hopko et al., 2022). The outcomes are listed in Table 2.



**Figure 1:** Distribution of field of domains [IDs] can be found in Table 1.

As expected, the most common operator performance measures were workload and situational awareness (SA). Of the 22 authors who measured workload, 18 used NASA-TLX, while others measured workload with robot attention demand, the operator utilization rate, the ZEIS rating method by Pitrella (1989), or the heart rate. In the case of SA, some scholars used standardized methods such as the SAGAT by Endsley (1988) or the 3D SART by Taylor (1990, 2017), while others used eye movement or tracking, navigation or perception metrics, and operator response or waiting times. In addition, attention was measured in three studies using vehicle wait times due to operator attentional inefficiencies. Other studies investigated the trust in automation or in the robots. Only one author considered emotional state using the SAM questionnaire by Bradley and Lang (1994).

**Table 2.** Operator performance outcomes.

Outcome	Measuring instrument	IDs
Workload	NASA-TLX	[1, 2, 3, 4, 13, 16, 17, 18, 19, 22, 23, 24, 25, 27, 29, 30, 34, 35]
Situational awareness	Other instruments	[6, 8, 28, 32]
	SAGAT and 3D SART	[16, 22, 24]
	Eye movement or tracking	[10, 21]
Attention	Other instruments	[5, 14, 25, 26]
Trust	Waiting times	[8, 13, 33]
	Trust in automation	[4, 17]
Emotional state	Trust in the system	[9]
	SAM questionnaire	[7]

System performance measures, on the other hand, are those that assess how well the human and robots perform as a team (Steinfeld et al., 2006). The compilation of these outcomes is listed in Table 3. In summary, in search

and rescue, including fire fighters, most authors considered the number of victims found (performance rate) or missed or (error rate) as an effectiveness measure, and the time to complete the mission and the area covered or missed by the robots as efficiency measures. Just three authors considered the mean fan-out depending on the workload or SA.

In other fields, scholars focused more on number of successful task completions or task completion time, number of errors or false alarms, waiting time or interaction time. Other outcomes included percentage of collisions or robot damage, robot's detection accuracy, progress, speed, or synchronization, maneuverability, force mapping, path and perceived usefulness.

**Table 3.** System performance outcomes.

Field	Outcome	IDs
Search and rescue (including fire fighters)	Victim detection	[2, 4, 10, 14, 15, 17, 18, 19, 33, 34, 35]
	Task completion time	[3, 10, 15, 16, 19, 30]
	Area explored	[3, 4, 15, 19, 31, 33, 34, 35]
	Fan-out	[5, 30, 35]
	Others	[9, 14, 15, 22]
Other fields	Task success	[1, 8, 12, 21, 23, 25, 28]
	Task completion time	[1, 6, 8, 13, 20, 25, 28, 29]
	Errors or false alarms	[1, 26]
	Others	[7, 11, 27, 29]

## DISCUSSION

The objective of this literature review was to comprehend the characteristics of SHMR systems and provide a synopsis of the most important features. We found that SHMRs are mainly used for search and rescue operations. In this domain, robots assist the operator during natural disasters by entering a hard-to-reach disaster area and understanding the environment in order to save lives (Kruijff-Korbayová et al., 2015). About 80% of the identified records used only one type of mobile robot. Combining different types of robots might bring advantages (e.g., increased area coverage), especially in emergency situations where the landscape is mixed (Magid et al., 2022), or in construction sites where the structure is constantly changing (Kim et al., 2022).

Studies have primarily focused on evaluating mental workload using the NASA TLX, and to a lesser extent, situational awareness using various techniques. Therefore, cognitive factors have been studied but with questionnaires that seem to show less sensitivity to task changes (Barajas-Bustillos et al., 2023, Rubio et al., 2004). Besides, when interacting with robots through a GUI, it is necessary to consider all multidimensional user states (Schwarz et al., 2014), including motivation, which has not been studied. Hence, there is a need to explore physical and psychological factors of the operator when interacting with multiple robots using one or multiple screens.

Among the system performance outcomes, some studies have shown that the more robots are used, the more problems may occur, such as robot

damage or low task performance. This is due to increased operator workload (Adams, 2009) or decreased operator's attentional capacity (Crandall and Cummings, 2007). Finally, researchers have tested different robot team combinations, but few authors have explored the robot threshold at which system performance is no longer improved by adding additional robots. Wang et al. (2009b) and Velagapudi et al. (2008) proposed a fan-out of 9 robots, while Crandall and Cummings (2007) explained that there seems to be a plateau at 6 robots. Regarding robot automation, some authors argued that it might reduce operator workload by managing attention, reducing robot demands, and improving task performance (Goodrich et al., 2007, Hong et al., 2019). On the contrary, others explained that a higher level of automation does not necessarily improve system performance (Chien et al., 2013, Khasawneh et al., 2019).

Although we conducted an exploratory search, it is likely that we did not find more documents related to the topic of interest given our keywords, and therefore we obtained more records from other authors' references than directly from the database. Nevertheless, the synthesis presented here offers an overview of the current state of the systems. In summary, SHMR systems have the potential to enhance task performance, but their effect on the operator has not been sufficiently explored. Future research could focus on exploring these systems in industries that have not yet been studied (e.g., energy industry, gas or electricity), and analyzing the impact of individual characteristics, such as age or gender, on these systems.

## FUNDING

The author(s) declare that they have received financial support for the research, writing, and/or publication of this article. This work is part of the project F2557 "Human Centered Technologies for a Safer and Greener European Construction Industry" (HumanTech), which is funded by the European Union's research and innovation program Horizon Europe under grant agreement N° 101058236) and the project F2569 "Optimized Work Design for Network Control Centers of Critical Infrastructure" (Beautiful), which is funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) (funding code: 03SF0694) and is coordinated by the Project Management Agency Jülich (PTJ).

## REFERENCES

- Adams, J. A. 2009. Multiple robot / single human interaction: effects on perceived workload. *Behaviour & Information Technology*, 28, 183–198.
- Balestrieri, E., Daponte, P., De Vito, L. & Lamonaca, F. 2021. Sensors and Measurements for Unmanned Systems: An Overview. *Sensors* [Online], 21.
- Barajas-Bustillos, M. A., Maldonado-Macías, A. A., Serrano-Rosa, M. A., Hernandez-Arellano, J. L., Llamas-Alonso, L. & Balderrama-Armendariz, O. 2023. Impact of experience on the sensitivity, acceptability, and intrusiveness of two subjective mental workload techniques: The NASA TLX and workload profile. *Work*, 75, 1265–1275.

- Beer, J. M., Fisk, A. D. & Rogers, W. A. 2014. Toward a framework for levels of robot autonomy in human-robot interaction. *J Hum Robot Interact*, 3, 74–99.
- Bostondynamics. 2023. *Spot - Innovating work across industries* [Online]. Available: <https://bostondynamics.com/industry/> [Accessed 9. April 2024].
- Boussemart, Y. & Cummings, M. L. 2011. Predictive models of human supervisory control behavioral patterns using hidden semi-Markov models. *Engineering Applications of Artificial Intelligence*, 24, 1252–1262.
- Bradley, M. M. & Lang, P. J. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49–59.
- Breslow, L. A., Gartenberg, D., Mccurry, J. M. & Trafton, J. G. 2014. Dynamic Operator Overload: A Model for Predicting Workload During Supervisory Control. *IEEE Transactions on Human-Machine Systems*, 44, 30–40.
- Brooks, N., Scerri, P., Sycara, K., Wang, H., Chien, S.-Y. & Lewis, M. 2011. Asynchronous Control with ATR for Large Robot Teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55, 444–448.
- Casper, J. & Murphy, R. R. 2003. Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, IEEE Trans. Syst., Man, Cybern. B*, 33, 367–385.
- Chandarana, M., Hughes, D., Lewis, M., Sycara, K. & Scherer, S. 2021. Planning and Monitoring Multi-Job Type Swarm Search and Service Missions. *Journal of Intelligent & Robotic Systems*, 101, 1–14.
- Chen, J. Y. C. & Barnes, M. J. 2021. Human–Robot Interaction. *Handbook of Human Factors and Ergonomics*.
- Chien, S.-Y., Lewis, M., Mehrotra, S. & Sycara, K. 2013. Imperfect Automation in Scheduling Operator Attention on Control of Multi-Robots. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57, 1169–1173.
- Crandall, J. W. & Cummings, M. L. 2007. Developing performance metrics for the supervisory control of multiple robots. IEEE.
- Crandall, J. W., Goodrich, M. A., Olsen, J. D. R. & Nielsen, C. W. 2005. Validating Human-Robot Interaction Schemes in Multitasking Environments. *IEEE Transactions on Systems, Man & Cybernetics: Part A*, 35, 438–449.
- Cummings, M. L. & Mitchell, P. J. 2008. Predicting Controller Capacity in Supervisory Control of Multiple UAVs. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, IEEE Trans. Syst., Man, Cybern. A*, 38, 451–460.
- Dietz, G., Kim, L. H., Jane, L. E., Follmer, S. & Washington, P. Human perception of swarm robot motion. 2017. Association for Computing Machinery, 2520–2527.
- Donmez, B., Nehme, C. & Cummings, M. L. 2010. Modeling Workload Impact in Multiple Unmanned Vehicle Supervisory Control. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, IEEE Trans. Syst., Man, Cybern. A*, 40, 1180–1190.
- Ebsco, C. 2023. *What content is included in my EBSCO Discovery Service (EDS) profile?* [Online]. Available: [https://connect.ebsco.com/s/article/What-content-is-included-in-my-EBSCO-Discovery-Service-EDS-profile?language=en\\_US](https://connect.ebsco.com/s/article/What-content-is-included-in-my-EBSCO-Discovery-Service-EDS-profile?language=en_US) [Accessed April 2024].

- Endsley, M. R. Situation awareness global assessment technique (SAGAT). Proceedings of the IEEE 1988 national aerospace and electronics conference, 1988. IEEE, 789–795.
- Fooladi Mahani, M., Jiang, L. & Wang, Y. 2021. A Bayesian Trust Inference Model for Human-Multi-Robot Teams. *International Journal of Social Robotics*, 13, 1951–1965.
- Frische, F. & Ludtke, A. 2013. SA-Tracer: A tool for assessment of UAV swarm operator SA during mission execution. IEEE.
- Fuchs, C., Borst, C., De Croon, G. C. H. E., Van Paassen, M. M. & Mulder, M. 2014. An Ecological Approach to the Supervisory Control of UAV Swarms. *International Journal of Micro Air Vehicles*, 6, 211–229.
- Glas, D. F., Kanda, T., Ishiguro, H. & Hagita, N. 2012. Teleoperation of Multiple Social Robots. *IEEE Transactions on Systems, Man & Cybernetics: Part A*, 42, 530–544.
- Goodrich, M. A., Mclain, T. W., Anderson, J. D., Sun, J. & Crandall, J. W. 2007. Managing autonomy in robot teams: Observations from four experiments. IEEE.
- Hocraffer, A. & Nam, C. S. 2017. A meta-analysis of human-system interfaces in unmanned aerial vehicle (UAV) swarm management. *Appl Ergon*, 58, 66–80.
- Hong, A., Igharoro, O., Liu, Y., Niroui, F., Nejat, G. & Benhabib, B. 2019. Investigating Human-Robot Teams for Learning-Based Semi-autonomous Control in Urban Search and Rescue Environments. *Journal of Intelligent & Robotic Systems*, 94, 669–686.
- Hong, A., Lee, D., Bülthoff, H. & Son, H. 2017. Multimodal feedback for teleoperation of multiple mobile robots in an outdoor environment. *Journal on Multimodal User Interfaces*, 11, 67–80.
- Hopko, S., Wang, J. & Mehta, R. 2022. Human Factors Considerations and Metrics in Shared Space Human-Robot Collaboration: A Systematic Review. *Front Robot AI*, 9, 799522.
- Huang, H.-M. 2004. *Autonomy Levels for Unmanned Systems (ALFUS) Framework - Volume I: Terminology*. National Institute of Standards and Technology.
- Humphrey, C. M., Henk, C., Sewell, G., Williams, B. W. & Adams, J. A. 2007. Assessing the scalability of a multiple robot interface. IEEE.
- Izadi Moud, H., Shojaei, A. & Flood, I. Current and Future Applications of Unmanned Surface, Underwater, and Ground Vehicles in Construction. *Construction Research Congress 2018*, 106–115.
- Kawamura, K., Nilas, P., Muguruma, K., Adams, J. A. & Chen, Z. An agent-based architecture for an adaptive human-robot interface. 36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the, 6–9 Jan. 2003, p. 8.
- Khasawneh, A., Rogers, H., Bertrand, J., Madathil, K. C. & Gramopadhye, A. 2019. Human adaptation to latency in teleoperated multi-robot human-agent search and rescue teams. *Automation in Construction*, 99, 265–277.
- Kim, Y., Kim, H., Murphy, R., Lee, S. & Ahn, C. R. 2022. Delegation or Collaboration: Understanding Different Construction Stakeholders' Perceptions of Robotization. *Journal of Management in Engineering*, 38, 1–12.
- Kruijff-Korbayová, I., Colas, F., Gianni, M., Pirri, F., Greeff, J., Hindriks, K., Neerincx, M., Ögren, P., Svoboda, T. & Worst, R. 2015. TRADR Project: Long-Term Human-Robot Teaming for Robot Assisted Disaster Response. *KI: Künstliche Intelligenz*, 29, 193–201.

- Lewis, M. 2013. Human Interaction With Multiple Remote Robots. *Reviews of Human Factors and Ergonomics*, 9, 131–174.
- Lewis, M., Chien, S.-Y., Mehorthra, S., Chakraborty, N. & Sycara, K. 2014. Task Switching and Single vs. Multiple Alarms for Supervisory Control of Multiple Robots. In: Harris, D. (ed.) *International Conference on Engineering Psychology and Cognitive Ergonomics*.
- Lewis, M., Huadong, W., Shih Yi, C., Velagapudi, P., Scerri, P. & Sycara, K. 2010. Choosing Autonomy Modes for Multirobot Search. *Human Factors*, 52, 225–233.
- Magid, E., Matsuno, F., Suthakorn, J., Svinin, M., Bai, Y., Tsoy, T., Safin, R., Lavrenov, R., Zakiev, A., Nakanishi, H., Hatayama, M. & Endo, T. 2022. e-ASIA Joint Research Program: development of an international collaborative informational system for emergency situations management of flood and land slide disaster areas. *Artificial Life & Robotics*, 27, 613–623.
- Murashov, V., Hearl, F. & Howard, J. 2016. Working safely with robot workers: Recommendations for the new workplace. *J Occup Environ Hyg*, 13, D61–71.
- Olsen J. R., D. R. & Wood, S. B. Fan-out: Measuring human control of multiple robots. 2004. 231–238.
- Patel, J., Sonar, P. & Pinciroli, C. 2022. On multi-human multi-robot remote interaction: A study of transparency, inter-human communication, and information loss in remote interaction. *Swarm Intelligence*, 16, 107–142.
- Pitrella, F.D. 1989. *A Cognitive Model of the Internal Rating Process*, Forschungsges. für Angewandte Naturwiss.
- Prewett, M. S., Saboe, K. N., Johnson, R. C., Coovert, M. D. & Elliott, L. R. 2009. Workload in Human-Robot Interaction: A Review of Manipulations and Outcomes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53, 1393–1397.
- Ratwani, R. M., Mccurry, J. M. & Trafton, J. G. 2010. Single operator, multiple robots: An eye movement based theoretic model of operator situation awareness.
- Roldán, J. J., Peña-Tapia, E., Martín-Barrio, A., Olivares-Méndez, M. A., Del Cerro, J. & Barrientos, A. 2017. Multi-Robot Interfaces and Operator Situational Awareness: Study of the Impact of Immersion and Prediction. *Sensors (14248220)*, 17, 1720.
- Rosenfeld, A., Agmon, N., Maksimov, O. & Kraus, S. 2017. Intelligent agent supporting human–multi-robot team collaboration. *Artificial Intelligence*, 252, 211–231.
- Rubio, F., Valero, F. & Llopis-Albert, C. 2019. A review of mobile robots: Concepts, methods, theoretical framework, and applications. *International Journal of Advanced Robotic Systems*, 16.
- Rubio, S., Díaz, E., Martín, J. & Puente, J. M. 2004. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology*, 53, 61–86.
- Ruiz, J. J., Martinez-De-Dios, J. R., Ollero, A. & Viguria, A. Immersive displays for building spatial knowledge in multi-UAV operations. 2015. Institute of Electrical and Electronics Engineers Inc., 1043–1048.
- Schwarz, J., Fuchs, S. & Flemisch, F. Towards a more holistic view on user state assessment in adaptive human-computer interaction. 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 5–8 Oct. 2014, 1228–1234.
- Sellner, B., Heger, F. W., Hiatt, L. M., Simmons, R. & Singh, S. 2006a. Coordinated Multiagent Teams and Sliding Autonomy for Large-Scale Assembly. *Proceedings of the IEEE*, 94, 1425–1444.

- Sellner, B. P., Simmons, R., Singh, S. & Hiatt, L. M. Attaining situational awareness for sliding autonomy. 2006b. Association for Computing Machinery (ACM), 80–87.
- Setter, T., Fouraker, A., Kawashima, H. & Egerstedt, M. 2015. Haptic Interactions With Multi-Robot Swarms Using Manipulability. *Journal of Human-Robot Interaction*, 4.
- Siegwart, R. & Nourbakhsh, I. R. 2004. *Introduction to autonomous mobile robots*, Cambridge, Mass. [u.a, MIT Press].
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. & Goodrich, M. 2006. Common metrics for human-robot interaction. *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. Salt Lake City, Utah, USA: Association for Computing Machinery.
- Taylor, R. M. Situational Awareness Rating Technique (SART): The development of a toll for aircrew systems design. In: AGARD Conference Proceedings No. 478, Situational Awareness in Aerospace Operations. Aerospace Medical Panel Symposium. Copenhagen, 2nd–6th October, 1989, 1990.
- Taylor, R. M. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. *Situational awareness*. Routledge.
- Tranfield, D., Denyer, D. & Smart, P. 2003. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14, 207–222.
- Trouvain, B., Schlick, C. & Mevert, M. Comparison of a map-vs. camera-based user interface in a multi-robot navigation task. SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483), 2003. IEEE, 3224–3231.
- Trouvain, B. & Wolf, H. L. 2002. Evaluation of multi-robot control and monitoring performance. Piscataway, NJ, USA, USA: IEEE.
- Velagapudi, P., Scerri, P., Sycara, K., Huadong, W., Lewis, M. & Jijun, W. 2008. Scaling effects in multi-robot control. IEEE.
- Vilela, J., Yugang, L. & Nejat, G. 2013. Semi-autonomous exploration with robot teams in urban search and rescue. IEEE.
- Villani, V., Capelli, B., Secchi, C., Fantuzzi, C. & Sabattini, L. 2020. Humans interacting with multi-robot systems: A natural affect-based approach. *Autonomous Robots*, 44, 601–616.
- Wang, H., Lewis, M., Chien, S.-H. & Velagapudi, P. 2009a. Scaling Effects for Synchronous vs. Asynchronous Video in Multi-robot Search. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53, 364–368.
- Wang, H., Lewis, M., Velagapudi, P., Scerri, P. & Sycara, K. 2009b. How search and its subtasks scale in N robots. *ACM/IEEE International Conference on Human-Robot Interaction*, 141–148.
- Wang, J. & Lewis, M. 2007. Human control for cooperating robot teams. IEEE.
- Wickens, C. D. 2008. Multiple Resources and Mental Workload. *Human Factors*, 50, 449–455.
- Wilson, J., Chance, G., Winter, P., Lee, S., Milner, E., Abeywickrama, D., Windsor, S., Downer, J., Eder, K., Ives, J. & Hauert, S. 2023. Trustworthy Swarms. *Proceedings of the First International Symposium on Trustworthy Autonomous Systems*.



- 
- Wong, C. Y. & Seet, G. 2017. Workload, awareness and automation in multiple-robot supervision. *International Journal of Advanced Robotic Systems*, 14.
- Wong, C. Y., Seet, G. & Sim, S. K. 2011. Multiple-robot systems for USAR: Key design attributes and deployment issues. *International Journal of Advanced Robotic Systems*, 8, 85–101.
- Yanco, H. A. & Drury, J. L. A Taxonomy for Human-Robot Interaction. AAAI Fall Symposium on Human-Robot Interaction, 2002. 111–119.

# Robotization of Work - What Are the Experiences Among Employees in Automotive Industry Company in the Czech Republic

Vladimira Lipsova<sup>1,2</sup>, Viola Pirova Hubacova<sup>1</sup>, Jana Murza<sup>1</sup>,  
and Ivana Libalova<sup>1</sup>

<sup>1</sup>National Institute of Public Health, Srobarova 48, 100 00 Prague 10, Czech Republic

<sup>2</sup>Department of Occupational Medicine, First Faculty of Medicine, Charles University, Prague, Czech Republic

## ABSTRACT

New technologies, such as advanced robotic systems, have led to a renewed debate on the impact of automation on occupational safety and health. The present research aimed to map how employees perceive robotics in their work activities and what impact the introduction of robots has on their mental health. Due to the nature of the study, it is an exploratory research with a qualitative design. Content analysis was used to analyze the data. The main method of data collection was unstructured and semi-structured interviews with managers and other employees of the automotive industry company. Unstructured interviews were first held with 3 managers (technical manager, health and safety manager, and HR manager) and then semi-structured interviews were held with 6 production operators, 1 shift leader, and 1 technical worker. The interviews were conducted with 7 women and 1 man in the age range of 35–60 years. The analysis of the interviews yielded 9 outcome themes: general perceptions of robotization; change; problems; support at work at the beginning of robot introduction and current support; social relations and communication; employee characteristics; physical and mental health. The results of the interviews showed the importance and relevance of the topic for both the management and employees. Initially, the introduction of robots is associated with several problems, but these are gradually turned into benefits and make employees' work easier. The greatest psychological stress is associated with the work of technical staff and managers. Findings from this study suggest the importance of support in terms of adequate and easy-to-understand training and also supportive leadership. The characteristics of the employee who is collaborating and cooperating with robots in the company studied were unique – not highly educated but very skilled and proactive. It also seems that company culture and managerial support are necessary to create a healthy and successful organization.

**Keywords:** Robotisation, Employee mental health, Corporate culture

## INTRODUCTION

The appearance of new technologies, such as advanced robotic systems that can closely interact with humans, has led to a revival of the debate on

the automation potential of jobs and tasks as well as their consequences on occupational safety and health (EU-OSHA, 2018). There are concerns that the nature and organization of work, as well as employment status, will change considerably. There could be a significant loss in medium-skilled jobs and a significant increase in higher-skilled jobs. The workforce is expected to become more diverse and dispersed, changing jobs frequently and working online rather than in person. All of this will bring challenges and opportunities, including in the area of OSH (EU-OSHA, 2018). An increase in robot exposure at work reduces work-related injuries, but there could be an increase in mental health problems (Gihleb, Giuntella, Stella, & Wang, 2022). Some evidence shows lowering effects on job physical intensity and disability, but no evidence of significant effects on mental health and work and life satisfaction (Gihleb, Giuntella, Stella, & Wang, 2022). Also inconclusive are data about either replacing workers with robots or making their working tasks more responsible and causing more psychosocial stress (Leso, Fontana, & Iavicoli, 2018). Also, the phenomenon of digital stress has arisen, its main components are stress and digital technologies. Digital stress can negatively affect individual physiological well-being, user satisfaction, or individual performance at work (Fischer, 2021).

Mental health is also affected by other factors: the error rate of robots (O'Driscoll, Brough, Timms, & Sawang, 2010); technical problems such as breakdowns or slowdowns (Körner et al., 2019); increasing requirements for employee competencies (Cascio & Montealegre, 2016); increased need to react quickly to emerging problems (Dvash & Mannheim, 2001) and, last but not least, wary about one's job (De Vries, Gentile, & Wacker, 2020). Thus, employees may experience increased stress, anxiety, frustration, increased cognitive load (Gihleb, Giuntella, Stella, & Wang, 2022; Dvash & Mannheim, 2001), or an increase in the use of addictive substances (Gihleb, Giuntella, Stella, & Wang, 2022).

Possible interventions that can be implemented to enhance both technological engagement and workers' well-being consist primarily of training, supporting, and developing an organizational culture and climate that fosters technological change, at the same time recognizing the needs of individual workers and the importance of maintaining their well-being (O'Driscoll, Brough, Timms, & Sawang, 2010).

## **METHODS**

Considering the research objective, qualitative research was chosen using the methods of unstructured and semi-structured interviews. Data collection took place in the automotive industry company. Collaborative robots (cobots) were used in this company. The robotization process started in 2017 and the company saw a reduction in the number of employees from 1,600 to 600. Robotization helped with lowering the risks at work (chemical hazards and physical load). Unstructured interviews were first held with 3 managers (technical manager, health and safety manager, and HR manager), and then semi-structured interviews were held with 6 production operators, 1 shift

leader, and 1 technical worker. The interviews took place with 7 women and 1 man in the age range of 35–60 years.

The interviews took place in reserved rooms during their working hours. Due to operational limitations, it was not possible to record the interviews, so they were written down in real-time. An informed consent was always signed before the actual data collection.

Transcripts from individual interviews were used for data analysis. These texts were repeatedly read by analyzers, who highlighted important sections of the texts. Subsequently, open and axial coding was used. As part of open coding, text sequences were divided into units to which codes were assigned. Within the axial coding, mutual connections and relationships were sought between the codes until the resulting coding system was created. The emerging system of codes was regularly discussed at analyzer meetings. After analyzing the individual interviews, common themes were identified. In the end, the obtained data were interpreted.

Data collection took place in August 2023.

## **RESULTS**

### **Unstructured Interviews**

Managers agree that the introduction of robots affects the culture of the entire company. The company must get ready for the robotization process in advance. Concerning operators, they describe that the recruitment process is more focused on committed, flexible, and technically educated employees. Among existing employees, proactivity and responsibility are also key characteristics.

### **Semi-Structured Interviews**

By analyzing the interviews, 9 resulting themes were identified. In the beginning, we present a list of basic topics and their subtopics:

- General perception
- Change
- Fear
- Problems
- Support (training and current support)
- Social relations and communication
- Characteristics of the employee
- Physical health
- Mental health (stress, fatigue, cognitive functions, memory and attention, abuse).

### **General Perception**

The general perception of robotization in companies is positive for multi-member employees. Employees evaluate robots positively - they make their work easier, especially physically.

## **Change**

The introduction of robotization in the company is a big change in all directions. The employees mainly saw a change in the reduction of physical demands.

## **Fear**

Most employees are not afraid of working with robots. Everyone is trained and knows exactly what they can and cannot do with the robots. Some respondents spoke of a healthy respect for machines and new technology.

## **Problems**

The biggest problems with robotization are often connected with their introduction into the process. Consequently, the most common problems are connected with the supplied material rather than with the work of the robots themselves.

## **Support**

Employees rate the training they received as sufficient, simple, and effective. The most important for production operators is to know the basic principles of working with robots, where they can and cannot intervene. It is also useful for them to know what the most common problems are and how to solve them.

## **Social Relations and Communication**

Almost all employees agreed that social relations and communication have not changed significantly due to robotization.

## **Characteristics of the Employee**

The employees more or less agree that no specific skills or education are needed to work as a production operator. In interviews, it is often mentioned that an employee suitable for working with robots should first of all want to work with robots and also be willing to learn new things.

## **Physical Health**

Participants often stated that the introduction of robots significantly reduced physical demands and therefore pain and strain on their arms and hands.

## **Mental Health**

Most respondents generally do not consider their work stressful. They only feel a greater degree of stress in situations where problems with robots occur. Also, most respondents perceive working with robots as less tiring for them. The interviewees did not mention increased demands on memory processes. They did not observe the effect of robotization on memory, attention, sleep patterns, or addictive behavior.

## DISCUSSION

The introduction of robotization in companies brings with it significant changes that affect health and safety. Protection of workers from excessive physical load or other work hazards is the main positive outcome of our research and similar other research (Gihleb, Giuntella, Stella & Wang, 2022; Ishida, Tanaka, Taniguchi & Moriizumi, 2006).

The results regarding the impact on psychological health are not that clear-cut. According to this study, employees experience the highest level of stress when introducing new robots, which entails many changes, technical problems, or learning new procedures. According to other authors, these factors are also, significant stressors when working with robots (O'Driscoll, Brough, Timms & Sawang, 2010; Körner et al., 2019; Dvash & Mannheim, 2001). However, after the initial problems, most of the employees are satisfied and perceive the involvement of robots in the production process positively. Managers, who communicate with employees and constantly explain the positive effects for themselves, have a lot to do with this.

The production operators themselves perceived that the training was trouble-free and simple, and the support they receive is adequate and helpful. This can be the result of an already set up system of education and training at the given workplace, optimally set communication, and satisfactory organization of work. This finding is in line with other authors (O'Driscoll, Brough, Timms & Sawang, 2010).

Some research (Lu, Xie, Wang, Li & Xu, 2022) also speaks of increased psychological stress that results from the fear of robots or the possible unpredictability of the situation when working with them. Despite this, most of the probands in the presented research do not fear or worry about robots. They are well-trained and follow safety guidelines to protect their health. Rather than fear, they spoke of a healthy respect for robots.

The presented study came to the conclusion that managers are the most psychologically threatened by robotization. They have to communicate with employees about the change and process and constantly explain the desired positive effects. Technical workers or shift leaders are those, whose training is very demanding and requires significantly more time in comparison with the training of operators.

From the present research study, as well as other similar studies (Murashov, Hearl & Howard, 2016; Adem, Çakit & Dağdeviren, 2020), the recommendation is to change corporate culture and work with human resources.

The current research is the first of its kind in the Czech Republic and brings with it many limitations. The first of them is the relatively small sample of respondents, which consisted of employees who still work with robots. We do not have any information from people who may have been negatively affected by robotization, *e.g.*, by losing their jobs. The company joined the project voluntarily and was motivated to lead good practice. The biggest technical limitation was the impossibility of recording individual interviews. Some other important information may have been lost in the real-time transcription.

## CONCLUSION

Advanced robotic systems at work have led to a renewed debate on the impact of automation on occupational safety and health. The present research aimed to map how employees perceive robotics in their work activities and what impact the introduction of robots has on their mental health.

The current research is the first of its kind in the Czech Republic and brings with it many limitations. The conclusions show that employees experience the highest level of stress when introducing new robots, which entails many changes, technical problems, or learning new procedures. However, after the initial problems, most of the employees are satisfied and perceive the involvement of robots in the production process positively.

Production operators in the research did not experience any significant changes in stress, work pressure, frustration, or cognitive load. All of them said that working with robots does not affect their sleep or addictive behaviors.

Findings from this study suggest the importance of support in terms of adequate and easy-to-understand training and also supportive leadership. The characteristics of the employee who is collaborating and cooperating with robots in the studied company were unique – not highly educated but very skilled and proactive. It also seems that company culture and managerial support are necessary to create a healthy and successful organization.

It is therefore important to continue with further research and professional discussion to better understand the impact of robotization on companies and their employees and thus be able to take full advantage of its benefits and minimize negative impacts.

## ACKNOWLEDGMENT

Supported by Ministry of Health, Czech rep. - RVO (NIPH, 75010330).

## REFERENCES

- Adem, A., Çakit, E., & Dağdeviren, M. (2020). Occupational health and safety risk assessment in the domain of Industry 4.0. May 2020, *SN Applied Sciences* 2(5), doi: 10.1007/s42452-020-2817-x.
- Cascio, W. F., & Montealegre, R. (2016). How technology is changing work and organizations. *Annual Review of Organizational Psychology and Organizational Behavior* 3(1):349–375, doi: 10.1146/annurev-orgpsych-041015-062352.
- De Vries, G. J., Gentile, E., M. S., & Wacker, K. M. (2020). The rise of robots and the fall of routine jobs. July 2020, *Labour Economics* 66:101885, doi: 10.1016/j.labeco.2020.101885.
- Dvash, A., & Mannheim, B. (2001). Technological coupling, job characteristics and operators' well-being as moderated by desirability of control, May 2001, *Behaviour and Information Technology* 20(3):225–236, doi: 10.1080/01449290110044768.
- Euosha (2018). Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025: (2018). European Risk Observatory Report. European Agency for Safety and Health at Work.

- Fischer T., (2021) The Digital Stressors Scale, *Front. Psychol.*, 12 March 2021, Sec. Personality and Social Psychology, Volume 12 - 2021 | <https://doi.org/10.3389/fpsyg.2021.607598>.
- Gihleb, R., Giuntella, O., Stella, L., & Wang, T. (2022). Industrial robots, workers' safety, and health. *Labour Economics*, 2022, 78, 102205.
- Ishida, H., Tanaka, H., Taniguchi, H., & Moriizumi, T. (2006). Mobile robot navigation using vision and olfaction to search for a gas/odor source., June 2006, *Autonomous Robots* 20(3):231–238, doi: 10.1109/IROS.2004.1389370.
- Körner U, Müller-Thur K, Lunau T, Dragano N, Angerer P, Buchner A. (2019). Perceived stress in human-machine interaction in modern manufacturing environments-Results of a qualitative interview study. *Stress Health*. 2019 Apr;35(2): 187–199. doi: 10.1002/smi.2853. Epub 2019 Jan 30. PMID: 30609231.
- Leso V, Fontana L, Iavicoli I. (2018) The occupational health and safety dimension of Industry 4.0. *Med Lav*. 2018 Oct 29;110(5): 327–338. doi: 10.23749/mdl.v110i5.7282. PMID: 30378585; PMCID: PMC7682172.
- Lu L, Xie Z, Wang H, Li L, Xu X. Mental stress and safety awareness during human-robot collaboration - Review. *Appl Ergon*. 2022 Nov;105:103832. doi: 10.1016/j.apergo.2022.103832. Epub 2022 Jun 27. PMID: 35772289.
- Murashov V, Hearl F, Howard J. Working safely with robot workers: Recommendations for the new workplace. *J Occup Environ Hyg*. 2016;13(3): D61–71. doi: 10.1080/15459624.2015.1116700. PMID: 26554511; PMCID: PMC4779796.
- O'Driscoll, M. P., Brough, P., Timms, C., & Sawang, S. (2010). Engagement with information and communication technology and psychological well-being., *Research in Occupational Stress and Well Being* 8, doi: 10.1108/S1479-3555(2010)0000008010.



# Empirical Analysis of Social Implications During the Development of Automated Driving

Nuria Brüggemann, Sebastian Preis, and Arnd Engeln

Stuttgart Media University, Nobelstraße 10, 70569 Stuttgart, Germany

## ABSTRACT

This article presents and discusses studies and their first results on the development of an empirical method for the investigation of social implications caused by automated driving. The basic research project KARLI is funded by the German Federal Ministry for Economic Affairs and Climate Action. One of the objectives of KARLI is to develop a methodological approach for the empirical identification and evaluation of social implications expected in different phases of the user-centred development of automated driving. In KARLI, the empirical analysis of social implications is integrated in different studies throughout an iterative user-centred development process of interaction concepts for automated driving (SAE 0-4). In our definition, social implications address the consequences of a technological development for social structures or processes, as well as the development-related prerequisites necessary for a desired social target state. The empirical identification of possible social implications is conducted in two studies, using a qualitative survey (N = 12) and open-ended questions as part of an online survey (N = 35). Based on the results, a first draft of a questionnaire with closed questions is designed to assess the social implications previously identified. In a forthcoming third study, a VR simulation to test three different concepts to promote level-compliant driver behaviour, the social implications are assessed using the designed questionnaire. In addition, further expected social implications are elicited through open-ended questions. The study runs from the end of March to May 2024 with participants between 35 and 45 years of age (N = 95). According to the results, the statements about probable social implications are well answerable by car users, but the validity of the forecast given by the users remains unclear. That for, for a valid estimation of social implications a combination of users and expert's perspectives seems to be helpful.

**Keywords:** Social implications, Automated driving, Method development

## INTRODUCTION

This article bases on the collaborative project KARLI (Artificial Intelligence for Adaptive, Responsive and Level-Compliant Interaction in Future Vehicles), funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK; Diederichs et al., 2022). The main aim of our research is to promote level-compliant driver behaviour across SAE levels 0–4 (SAE International, 2021).

An additional research aim of the Stuttgart media university within KARLI is to find out how social implications expected by innovations of automated driving can be empirically evaluated. Particularly, methods for identifying and evaluating social implications are implemented and evaluated within the iterative, user-centered development (DIN EN ISO 9241-210:2020-03, 2020) and evaluation of HMI prototypes to promote level-compliant driver behaviour.

In a literature review of 15 project reports (not published yet), no definition of “social implication” sufficient for our work can be found. So we derive one by ourselves as synthesis from Schubert & Klein, 2021; Lehner, 2011 and suggest it as follows: The term “social” here is understood as the change in society or parts of society with regard to the effects of the innovations to be created on the order, organization and actions of the people acting within it. According to Kemmer (2020), an implication describes a link between a prerequisite A and a consequence B. In formal notation, an implication is represented as follows:  $A \rightarrow B$  and is read as “If A then B”. In relation to KARLI, social implications can mean, for example, the causal or correlative consequences for society or parts of society (B) of an innovation created by KARLI (A). E.g.: If automated vehicles with driver status recognition become established on the market (A), then mobility in our society will change fundamentally (B) (Schweiker et al., 2023).

In our definition, social implications address the consequences of a technological development for social structures or processes, as well as the necessary development-related prerequisites for a desired social target state. In the 15 project reports analysed with the topic of social implications, various methods are used to record social implications. Workshops (Tirschmann & Brukamp, 2021), qualitative studies (Zerth et al., 2017), social media analyses (Niehaves, 2018) and design for all approaches are used (Reinboth & Witczak, 2012).

The aim of this work report is to develop an initial proposal for items that can measure social implications possibly emerging from automated vehicles. Based on the current results, this work report discusses the possible next steps.

## **METHOD**

### **Method and Results of the Previous First and the Second Study**

First study (from March to May 2023): The initial data processing stage is involved in a qualitative evaluation of low-fidelity prototypes for automated driving (which are presented in user narratives, a formulation of ideas in form of a story of use) in Germany. The qualitative online interviews are conducted and take approximately 90 minutes, each. 12 participants (three heavy commuters with level 2 experience, three young individuals between the age 18 and 25, three individuals aged 65 or older and three individuals who are responsible for childcare) are included. Within the interviews they are asked to consider the following question: “Imagine that such an automated driving system is established on the market. What do you think the impact will be on society and different social groups?”.

In data analysis, the quotes of the participants are matched to 24 different statements in German language. All statements were formulated in the same way (... , as...; e.g. “Fewer accidents, *as* the automated system drives purely reasonably”).

Second study (from August to September 2023): In an HMI-evaluation of preference for concepts to motivate adequate driver behaviour in different levels of automation, the statements of social implication from the first study are integrated as closed questions and rated by the participants the first time. 35 participants rate 21 statements (three statements from study 1 are removed as result of the pre-test) according to probability of occurrence from 0 to 100% in a standardized online questionnaire. In addition, to search for further statements, they are asked openly: “Do you have any other ideas about the impact the system could have on society once automated driving systems have established themselves on the market?”.

The standardised results of study 2 show a wide range of the expected probability of occurrence from 0 to 100% for nearly all statements. The mean of the ratings is between  $m = 23,77\%$  and  $m = 77,46\%$ , and the standard deviation is between  $SD = 20.62$  and  $SD = 34.83$ . Table 1 gives six examples of the results (highest and lowest mean of all statements, four in between).

**Table 1.** Examples of the results of rating social implication items in evaluation 2.

Statements of social implication	M(SD)
Fewer accidents, as the automated system drives purely reasonably.	77,46 (22.25)
The job of cab driver becomes obsolete, as automated vehicles no longer require an active driver.	59,00 (33.76)
Older people can participate more in life as they become more mobile.	54,74 (29.84)
Environment is protected as energy consumption of cars is reduced.	47,83 (31.99)
Some people are burdened by automated driving psychologically, as they are reachable even during the journey and have the opportunity to work.	28,86 (25.70)
German automotive industry suffers as driving pleasure gets lost.	23,77 (20.62)

The high variance of the mean values from standardized measurement can be seen as a first hint for a content-related validity of the items.

By the open question, 31 additional statements on social implications are collected. After revision, 19 with regard to content new and well-defined statements are added to the 24 statements from before. These results are published in Brüggemann et al. (2023). The high number of 19 additional statements found in the second study indicate an incompleteness of the first and maybe also of the current collection.

### Method of the Third Study

In a subsequent third evaluation within a VR simulation study to test against three different concepts for motivating adequate driver behaviour in different levels of automation, the social implication statements are estimated

according probability of occurrence, again. The evaluation runs from March 2024 to May 2024. At the end of the VR simulation, the participants are now asked to rate a total of 43 statements (the old 24 statements from study one and the additional 19 statements from study two) by a probability scale from 0 to 100%. In addition, they are asked again: “Do you have any other ideas about the impact it could have on society once automated driving systems are established on the market?”.

A total of 95 participants (48 females included) take part in the study. The average age is 39.55 years ( $SD = 3.13$ , ranging from 35 to 45 years).

For the evaluation of the items integrated in the third study, the mean values and the standard deviations are looked at. For the statistical verification of the item distributions, the Kolmogorov-Smirnov test for normal distribution is then carried out and the skewness and kurtosis of the individual items are considered (Bortz & Döring, 2006).

## RESULTS OF THE THIRD STUDY

Table 2 shows the six examples from Table 1 and in addition the extrema found in the sample according to M, Skewness, Kurtosis and Kolmogorov-Smirnov Test as examples of the results.

**Table 2.** Examples of the results of rating social implication items in evaluation 3.

Statements of social implications	M(SD)	Skewness	Kurtosis	Kolmogorov-Smirnov
Automated systems lead to discontent in society, as only financially privileged individuals can afford them.	77,11 (23.39)	-1.249	1.031	0.000
Novice drivers never learn manual driving properly, as the automated system takes on so many tasks for them.	74,54 (39.00)	-1.148	0.024	0.000
Fewer accidents, as the automated system drives purely rationally.	68,60 (28.58)	-0.849	-0.357	0.000
Many are unsettled by automated vehicles, as they are controlled by machines and not by humans.	65,59 (24.86)	-0.684	-0.084	0.006
The job of cab driver becomes obsolete, as automated vehicles no longer require an active driver.	56,06 (33.67)	-0.237	-1.303	0.005
Older people can participate more in life as they become more mobile.	53,86 (33.99)	-0.016	-1.496	0.001
Environment is protected as energy consumption of cars is reduced.	51,55 (33.65)	-0.034	-1.424	0.003
Some people are burdened by automated driving psychologically, as they are reachable even during the journey and have the opportunity to work.	51,92 (30.28)	-0.136	-1.162	0.066
German automotive industry suffers as driving pleasure gets lost.	45,07 (30.34)	0.123	-1.085	0.041
With fully automated vehicles, one would no longer need a driver's license, as the vehicle can drive without human intervention.	15,45 (23.04)	2.005	3.556	0.000

The mean of the estimations of the 43 statements for probability of occurrence is between  $m = 15,45\%$  and  $m = 77,11\%$ . The standard deviations of the answers given to the statements is between  $SD = 23.04$  and  $SD = 39.00$ . The results of kurtosis are almost all negative, as are the results of skewness. Most data thus show a left skewness and a slightly flattened distribution. Consequently, by a value below 0.05 the Kolmogorov-Smirnov test confirms the non-normally distributed values for 42 of the 43 items, tested in the study.

By the open question, about 100 new statements on social implications are collected. They are not finally edited, yet. Preliminary, we extracted 55 new statements.

## DISCUSSION

The variance of the mean values and the scatter in the quantitative scores can be seen as an indicator of a given content responsiveness of the quantitatively tested items. In addition, from a purely descriptive point of view, the mean values and standard deviations of Study 2 and Study 3 are mostly very close together, which can also be seen as hint for a good answerability in terms of content.

A reason for the deviations from the normal distribution could be the specific characteristic of the test sample. The range of answers given by the sample aged 35–45 could be limited. In fact, all people, who move in street traffic, so from almost all age groups are affected by the social implications of automated vehicles. Moreover, we have to note that a non-normal distribution does not mean invalid items, normal distribution is just a prerequisite for some statistical tests.

In sum, the items included as well as the probability rating scale applied here could be a suitable response scale, but are not adequate for some statistical tests.

On current state, the single statements measured in the study cannot be summed up to overarching variables because they refer to different expectations which are independent to each other. To prove this hypothesis in future work, all indicators will be included in an explorative factor analysis to empirically determine the number and characteristics of the factors underlying the measured indicators and the correlations between the indicators (DeVellis, 2003). Unfortunately, the sample size of study three only allows a first approach to a factor analysis. For a well-founded factor analysis with such a high number of items a sample size over  $N = 300$  is needed. At this point, however, it is necessary to see whether meaningful groups can be found or whether, each item should stand on its own.

The forecast validity of the data should also be reflected critically. What does it mean if a future scenario is considered probable or not probable in the vote? In fact, it shows a democratic assessment of the sample of respondents. Currently we do not know, if the forecast by many users is more valid than e.g. the qualitative estimation of a small number of experts for social implications. Perhaps, it is more worth to investigate the range of possible

social implications with a big number of users than to do the estimation of probability of occurrence with them.

For a multi perspective evaluation of the probability of occurrence of the statements, additional qualitative interviews could be helpful in order to gain more insights into the response behaviour of the test subjects. Or, as in similar projects, a mix of methods could be used, such as additionally rating the statements with experts.

The 55 additional implications from study three are not finally analysed yet. The fact that 55 additional potential social implications are found in addition to the 43 items already developed in the two preliminary studies indicate the impossibility to define a complete set of social implications for empirical evaluation in automated driving. Therefore, in addition to the further development and evaluation of quantitative scoring procedures, the possibility of adding new social implications should be integrated in a method for estimating social implications.

## **CONCLUSION**

The present study focuses on the development of methods for an empirical based prediction of social implications of automated driving. The study uses a combination of qualitative and quantitative methods to identify and evaluate potential social impacts by doing research with users and other people directly or non-directly affected by the automation. While the study provides valuable qualitative insights, a final set of relevant social implications in general cannot be defined. Moreover, the validity of user predictions of probability of implications can be questioned. It is therefore recommended to develop a kind of Delphi method for prediction of social implications: To predict social implications of a specific prototype, it could be useful to firstly generate possible social implications for this specific prototype by qualitative research with users (and experts). In a second step, the implications generated could be evaluated on standardised level according to their probability of occurrence by experts for those social implications. The KARLI project concludes in September 2024. Further work to this topic depends on additional funding.

## **REFERENCES**

- Bortz, J. Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- Brüggemann, N. Preis, S. Pagenkopf, A. Engeln, A. (2023) *Empirical analysis of social implications during the development of automated driving*. Hawaii: AHFE (Poster).
- Deutsches Institut für Normung e. V. (2020). "DIN EN ISO 9241-210:2020-03, Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019)", Deutsche Fassung EN ISO 9241-210:2019: Beuth, Berlin.
- DeVellis, R. F. (2003). *Scale development. Theory and applications*. Thousand Oaks: Sage Publications.

- Diederichs, F. Wannemacher, C. Faller, F. Mikolajewski, M. Martin, M. Voit, M. Widlroither, H. Schmidt, E. Engelhardt, D. Rittger, L. Hashemi, V. Sahakyan, M. Romanelli, M. Kiefer, B. Fäßler, V. Rößler, T. Großerüschkamp, M. Kurbos, A. Bottesch, M. Immoor, P. Engeln, A. Fleischmann, M. Schweiker, M. Pagenkopf, A. Mathis, L.-A. and Piechnik, D. (2022). “Artificial intelligence for adaptive, responsive, and level-compliant interaction in the vehicle of the future (KARLI)”, Conference 2022, International Conference on Human-Computer Interaction pp. 164–171. Springer International Publishing, Cham.
- Kemmer, D. (2020).  $A \rightarrow B$  ( $\gg$ Wenn A, dann B.  $\ll$ ): Zur zentralen Rolle von Implikationen für die Konzeption und Praxis einer Integrierten Forschung. Wiesbaden: Springer.
- Lehner, F. (2011). Sozialwissenschaft. Wiesbaden: VS.
- Niehaves, B. (2018). INEMAS: Grundlagen interaktions- und emotionssensitiver Assistenzsysteme. Siegen: Universität Siegen (Report).
- Reinboth, C. Witczak, U. (2012). Tecla - Autonomie, mobile Beratung zu technischen Assistenzsystemen zur Förderung des selbständigen Wohnens in der zweiten Lebenshälfte. (Report).
- SAE International (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. SAE Mobilus. [https://www.sae.org/standards/content/j3016\\_202104/preview/](https://www.sae.org/standards/content/j3016_202104/preview/)
- Schubert, K. Klein, A. (2021). Das Politiklexikon: Begriffe, Fakten, Zusammenhänge. Bonn: Dietz.
- Schweiker, M., Pagenkopf, A., Engeln, A. (2023). Integration von ELSI-Fragestellungen in KARLI. Babenhausen: halftime presentation KARLI (Poster).
- Tirschmann, F. Brukamp, K. (2021). Ethische und soziale Aspekte der Robotikassistenz bei Amyotropher Lateralsklerose (ALS). Ludwigsburg: Evangelische Hochschule Ludwigsburg (Report).
- Zerth, J. Müller, S. Heidl, C. (2017). SeLe- Sensoren für verbesserte Lebensqualität; Teilvorhaben: Nutzer- und Akzeptanzforschung sowie sozioökonomische Wirkungsanalyse. Fürth: Forschungsinstitut IDC, Wilhelm Löhe Hochschule für angewandte Wissenschaften (Report).

# The Best Fit Framework for Human Computer Interaction Research – Is it Possible?

**Benjamin Stephanus Botha and Lizette de Wet**

Department of Computer Science and Informatics, University of the Free State,  
Bloemfontein, South Africa

## ABSTRACT

The Best Fit Framework, originally proposed by Carroll et al. (2013) to synthesize qualitative data has been successful to conduct a review of the literature to produce models or frameworks for decision making and health behaviours. While successful in health behaviours, it has not been implemented within Human-Computer Interaction before. This paper aims to convey knowledge, experiences, and recommendations towards the use of the Best Fit Framework to synthesize data in the field of Human-Computer Interaction. The Best Fit Framework involves various stages. The first two stages run simultaneously and involve identifying relevant frameworks, models, or theories, using the BeHEMoTh (Behaviour of Interests; Health Context; Exclusions; Models, Theories, Frameworks) search technique, and to identify relevant primary research studies with qualitative evidence, using the SPIDER (Setting/Population; Phenomenon of Interest; Design, Evaluation, Research) search technique. The selected theories, frameworks, or models are then reduced to key elements and used as themes in the new framework. These themes are interpreted and compared to new or similar types of themes across the literature, found with the SPIDER technique. New identified themes are incorporated to create an updated framework. After the framework is created, it is tested as a final part of the synthesis process. To apply the Best Fit Framework in the field of Human-Computer Interaction, the researchers expanded the context of BeHEMoTh. The researchers sought to also include the prevention or minimization of Cybersickness. The inclusion of quantitative primary research studies as part of the SPIDER technique was added, as the original SPIDER technique focused on qualitative studies which assisted in expanding the pool of primary research studies. The last change addressed how the framework synthesis was tested. Rather than only revisiting evidence to create and explore relationships, the researchers evaluated as part of the newly created CyPVICS framework in a real-world case study to determine validity. The case study compared the usability and user experience of two immersive Virtual Reality navigation methods, namely touch controllers vs. omnidirectional treadmill, in the training of nursing students. In conclusion, the Best Fit Framework proved to be adaptable and useful in Human-Computer Interaction research.

**Keywords:** Best fit framework, Human-computer interaction, Cybersickness, Virtual reality



## INTRODUCTION

The Best Fit Framework was originally proposed by Carroll *et al.* (2013) as a refined method to synthesize qualitative data for creating frameworks. While the framework has been successful as a synthesis methodology to conduct reviews of the literature to produce models or frameworks for decision making and health behaviours, it has not been implemented within Human-Computer Interaction before.

This paper aims to convey knowledge, experiences, and recommendations towards the use of the Best Fit Framework to synthesize data in the field of Human-Computer Interaction. The process conveyed in this paper was part of a larger study and was used to create the CyPVICS framework (Botha and De Wet, 2024) to reduce or eliminate cybersickness (CS) during immersive virtual simulation.

## THE BEST FIT FRAMEWORK

The Best Fit Framework (Carroll *et al.*, 2013) uses the current developments in the qualitative data synthesis methodology to conduct a review of the literature to produce models or frameworks (Carroll, Booth and Cooper, 2011). This framework also provides a way to test, strengthen or improve existing models, for example, applying it to a different population. To use the framework, a relevant framework, theory, or model must be identified. The selected theory, framework or model is then reduced to certain key elements which are used as themes in the new framework. The themes are then interpreted by the reviewer and compared to new or similar types of themes across the literature (Carroll *et al.*, 2013).

The Best Fit Framework (Carroll *et al.*, 2013) involves various stages. Each stage depicts the application of a certain strategy or methodology. The first two stages run simultaneously. They include identifying relevant frameworks, conceptual models, or theory publications, using the BeHEMOTH technique, and to identify relevant research studies with qualitative evidence, using the SPIDER technique. These stages of the Best Fit Framework will be discussed in detail in the following sections, starting with the BeHEMOTH technique.

### BeHEMOTH Technique

The BeHEMOTH was created as a structured way to identify and specify relevant models or theories for a literature review. The BeHEMOTH consists of four parts, namely behaviour of interest (Be), health context (H), exclusions (E) and models or theories (MoTh) (Booth and Carroll, 2015).

Within the BeHEMOTH, there are various steps that must be adhered to so as to create an accurate priori framework, theory, or model. Once all publications have been identified, a priori framework must be constructed by means of thematic analysis. Thematic analysis is a method to analyse qualitative data sets by identifying recurring themes in different literature or data (Nowell *et al.*, 2017). The themes that are identified from the BeHEMOTH are combined and integrated to create the priori framework, theory or model (Carroll *et al.*, 2013). The next stage of the Best Fit Framework (Carroll *et al.*, 2013), which runs concurrently with the BeHEMOTH, is the SPIDER technique.

### **SPIDER Technique**

The SPIDER technique was created in response to limitations of the PICO (population/problem, intervention/exposure, comparison, and outcome). The SPIDER technique was meant to generate search terms easier than the PICO, while the results were also easier to manage (Cooke, Smith and Booth, 2012).

SPIDER involves five components, namely the setting/population (S), the phenomenon of interest (PI), the design (D), evaluation (E) and the research type (R). The setting/population refers to the sample or population and to the environment or setting to which the research is bound, for example, undergraduate nursing students in South Africa. The phenomenon of interest refers to the phenomenon or event being investigated or researched, for example, cybersickness (CS). The design, evaluation and research type refer to the research design, the methods of evaluation and the type of research, for example, case studies using mixed methods, and usability and UX evaluation (Carroll *et al.*, 2013).

To assist with the validity of the themes and elements in the priori framework, theory or model, the research studies obtained from the SPIDER technique must be appraised to determine their quality. This is done by synthesising the models, theories, and frameworks. Once the BeHEMOTH and SPIDER analyses are completed and the results are synthesised, the results from the research studies must be coded to that of the framework, theory or model to determine the validity, along with the themes and elements (Carroll *et al.*, 2013).

With the coding completed, the last steps of the Best Fit Framework (Carroll *et al.*, 2013) must be applied, namely, to create and add new themes based on aspects that could not be added in the initial creation, and to finally test the framework, theory or model.

### **Finalising of Framework, Theory or Model**

After the initial coding and thematic analysis, another analysis is necessary to include the results which could not be coded against the framework, theory, or model. These newly identified themes then need to be incorporated into the framework, theory, or model to produce an updated version. This version must finally be analysed to explore relationships between themes. It is then tested by means of exploring dissonance and the impact of variability. By testing the framework, theory or model, biases are reduced, for example, certain themes might be absent, which might actually be very important and relevant (Carroll *et al.*, 2013). It is important to note that evaluation or testing of a framework generally is a cyclical process and can be done on a continuous basis to improve the created framework (Hevner, 2007; Gregor and Hevner, 2017).

## **THE BEST FIT FRAMEWORK IN HUMAN COMPUTER INTERACTION**

To apply the Best Fit Framework in the field of Human-Computer Interaction, the researchers expanded the context of BeHEMOTH. While the original technique included the health context only, the researchers

sought to also include the prevention or minimization of Cybersickness (CS) (as the context of our study). The second change was the inclusion of quantitative primary research studies as part of the SPIDER technique. While the original SPIDER technique focused on qualitative studies only, the researchers included tried and tested primary studies in both qualitative and quantitative research. This assisted in expanding the pool of primary research studies, while also increasing the validity of a created framework. The third change addressed how the framework synthesis was tested. Rather than only revisiting evidence to create and explore relationships, the researchers tested (evaluated) a part of the newly created CyPVICS framework (i.e. sensory mismatch/conflict theory and improved models of interaction) in a real-world case study to determine validity. The case study involved comparing the usability and user experience (UX) of two immersive Virtual Reality navigation methods, namely touch controllers vs. omnidirectional treadmill, in the training nursing students (Botha and de Wet, 2024).

### Question and Search Strings

The first step in the Best Fit Framework (Carroll et al., 2013) was to determine the terms that would form part of the search strings for both the BeHEMoTh and the SPIDER techniques, from the main research question, namely: “*Which determinants should form part of a framework for designing immersive virtual clinical simulations to prevent or minimise cybersickness?*”. Although the main research question was the same for both the BeHEMoTh and SPIDER techniques, the way in which the search terms were extracted from this question, differed for these two techniques. The BeHEMoTh and the SPIDER techniques were also linked to the subsidiary research questions, namely:

- (BeHEMoTh): “*What are the common and unique constructs in existing frameworks, theories and models to prevent/minimise cybersickness in immersive virtual clinical simulation?*”
- (SPIDER): “*Which aspects affect/influence cybersickness during immersive virtual clinical simulation?*”

From the main research question, the first term that was identified for use in both the BeHEMoTh and the SPIDER techniques was *immersive virtual clinical simulation*, as it represented the application area of this study. Since alternative terms are used in literature for ‘immersive’ virtual clinical simulation (VCS), for example, simply VCS, or VRS (when in fact the authors are referring to immersive VCS), all related terms were used. Thus, all the research papers that focused on immersive VCS were included.

The second term was *cybersickness* (or CS, as it is often referred to), as it was the phenomenon in question for this study. The third term was determinants, which related to either constructs (subsidiary research question one) or aspects (subsidiary research question two). For all the terms, synonyms were also sourced from various literature studies where the study focused on immersive virtual reality in the clinical context. The synonyms assisted in obtaining the best possible results and to include as many literature

sources as possible related to the three identified terms (see Table 1 for all synonyms used in the search strings).

Once the applicable terms and synonyms were identified, the technique-specific approach could follow. The terms not indicated in Table 1 were excluded since they were not applicable to this study.

**Table 1.** Search terms and synonyms for BeHEMOTH and SPIDER techniques.

Question Extract and Applicable Technique	Search Terms
Immersive Virtual clinical simulation (SPIDER and BeHEMOTH)	Virtual clinical simulation (VCS) (Foronda, Godsall and Trybalski, 2013) Virtual Reality Simulation (VRS) (Jenson and Forsyth, 2012; Dubovi, Levy and Dagan, 2017) Virtual Simulation (VS) (Aebersold, Tschannen and Bathish, 2012) Clinical Virtual Simulation (CVS) (Padilha <i>et al.</i> , 2018) active HMD-based virtual reality (Arcioni <i>et al.</i> , 2019) virtual reality experiences with head-mounted displays (Arttu, 2018) Virtual Reality (Elwardy <i>et al.</i> , 2020) HMD-based virtual reality (Chang, Kim and Yoo, 2021) immersive content (Melo, Vasconcelos-Raposo and Bessa, 2018) VR simulation (Servotte <i>et al.</i> , 2020)
Cybersickness (SPIDER and BeHEMOTH)	Cybersickness (CS) (Wang <i>et al.</i> , 2019; Weech <i>et al.</i> , 2019) Virtual reality induced motion sickness (VRIMS) (Li <i>et al.</i> , 2021) Virtual reality induced symptoms and effect (VRISE) (Kemeny <i>et al.</i> , 2017) Visually induced motion sickness (VIMS) (Wang <i>et al.</i> , 2019; Weech <i>et al.</i> , 2019) Simulator Sickness (Duzmanska, Strojny and Strojny, 2018) Motion Sickness (Joseph, Browning and Jiang, 2020) Virtual Reality Sickness (Chang, Kim and Yoo, 2020)
Aspects (Relates to subsidiary research question 1) (SPIDER) All terms relate to well-known methods and techniques in primary research studies	Quantitative Qualitative Mixed method Case Study Interview Views Attitudes Focus group Experiment Opinions
Constructs (Relates to subsidiary research question 2) (BeHEMOTH)	Models, Theories, Frameworks (Carroll <i>et al.</i> , 2013)

The search strings for the BeHEMOTH and the SPIDER techniques were sent to an information specialist at the researcher's institutional library to conduct the search and return the results. The information specialist used the following databases to search for the literature: Academic Search Ultimate, Africa-Wide Information, APA PsycArticles, APA PsycInfo, CAB Abstracts, CINAHL with Full Text, ERIC, GreenFILE, Health Source - Consumer Edition, Health Source: Nursing/Academic Edition, Humanities Source Ultimate, MEDLINE, OpenDissertations, SPORTDiscus with Full

Text, Scopus, WoS. The technique-applicable search strings and their results will be discussed in the sections to follow, starting with the BeHEMOTH technique.

### BeHEMOTH Search String and Results

Two separate searches were done to obtain literature on models, theories or frameworks that could assist in compiling the CyPVICS Framework (Botha and De Wet, 2024). Both search strings had no date delimiter. The first search string for the BeHEMOTH technique was ((Be AND H AND MoTh) NOT E) and can be seen in Table 2.

**Table 2.** First BeHEMOTH search string used for this study.

First BeHEMOTH search string	
Be	Cybersickness OR Virtual reality induced motion sickness OR Virtual reality induced symptoms and effect OR visually induced motion sickness OR Simulator Sickness OR Motion Sickness OR Virtual Reality Sickness
AND	
H	Virtual clinical simulation OR Virtual Reality Simulation OR Virtual Simulation OR Clinical Virtual Simulation OR active HMD-based virtual reality OR virtual reality experiences with head-mounted displays OR Virtual Reality OR HMD-based virtual reality OR immersive content OR VR simulation
AND	
MoTh	Model OR Theory OR Framework
NOT	
E	Case Study OR Interview OR Views OR Attitudes OR Focus group OR Experiment OR Opinions OR Animal Model

From this search string six results were found; however, none were classified as a model theory or framework. Because the literature was lacking in the context of VCS, the search string was broadened by reducing it to (Be and MoTh), as can be seen in Table 3. More models, theories, and frameworks were obtained, although they were not necessarily in a VCS context. Even though these models, theories or frameworks were from different disciplines, CS was still applicable here. Consequently, they were evaluated for possible inclusion into the CyPVICS Framework (Botha and De Wet, 2024).

**Table 3.** Broadened search string for BeHEMOTH applicable to this study.

Second BeHEMOTH search string	
Be	Cybersickness OR Virtual reality induced motion sickness OR Virtual reality induced symptoms and effect OR visually induced motion sickness OR Simulator Sickness OR Motion Sickness OR Virtual Reality Sickness
AND	
MoTh	Model OR Theory OR Framework

From the second BeHEMOTH search string, a total of 1057 results were obtained after automatic deduplication. Once the search was completed, all the abstracts and titles were evaluated to determine whether they conformed for inclusion into this framework. The literature was evaluated based on the criteria in Table 3.

During the process of abstract and title evaluation, a total of 1050 titles and abstracts were excluded due to not being applicable to the terms identified for the BeHEMOTH technique. Once the abstract and title evaluations were completed, the full papers were sourced and analysed to determine their eligibility. From the seven papers, references were analysed to determine whether there were more articles that could be included as part of the ancestry search. After the ancestry search, a total of nine titles and abstracts were included, which brought the total full text articles that had to be reviewed for inclusion in the construction of the initial CyPVICS Framework for this study, to 12.

### SPIDER Search String and Results

The SPIDER technique was used to find primary research studies without limiting the date when the study was published. The search strategy for the first SPIDER technique (S AND PI AND DER) and the search string can be seen in Table 4. From the SPIDER technique, a total of 417 results were obtained after automatic deduplication. Once the search was completed, all the abstracts and titles were evaluated to determine whether they conformed to the research question from which the SPIDER technique search terms were derived.

During abstract and title evaluation, a total of 350 titles and abstracts were excluded as they were not applicable to the terms identified for the SPIDER technique in Table 4. Once the abstract and title evaluations were completed, the full papers were sourced and analysed to determine their eligibility. From the 57 papers, references were analysed to determine whether there were articles which could be included as part of the ancestry search. After an ancestry search, 13 titles and abstracts were included, which brought the article count to 70 full text articles that were reviewed for possible inclusion as primary research studies. In total, 39 studies conformed to the requirements. The data for the primary research studies were extracted into a table format. The raw data can be viewed table at <https://doi.org/10.38140/ufs.22955402.v1>.

**Table 4.** SPIDER search string used for this study.

SPIDER search string	
S	Virtual clinical simulation OR Virtual Reality Simulation OR Virtual Simulation OR Clinical Virtual Simulation OR active HMD-based virtual reality OR virtual reality experiences with head-mounted displays OR Virtual Reality OR HMD-based virtual reality OR immersive content OR VR simulation
AND	
PI	Cybersickness OR Virtual reality induced motion sickness OR Virtual reality induced symptoms and effect OR visually induced motion sickness OR Simulator Sickness OR Motion Sickness OR Virtual Reality Sickness
AND	
DER	Quantitative OR Qualitative OR Mixed method OR Case Study OR Interview OR Views OR Attitudes OR Focus group OR Experiment OR Opinions

## CONCLUSION

In conclusion, the Best Fit Framework proved to be adaptable and useful in Human-Computer Interaction research. This involved changes to the original Best Fit Framework, which included expanding the context, incorporating quantitative research studies, and implementing a usability and UX evaluation in a real-world case study.

## REFERENCES

- Aebersold, M., Tschannen, D. and Bathish, M. (2012) 'Innovative Simulation Strategies in Education', *Nursing Research and Practice*, 2012, pp. 1–7. Available at: <https://doi.org/10.1155/2012/765212>.
- Arcioni, B. *et al.* (2019) 'Postural stability predicts the likelihood of cybersickness in active HMD-based virtual reality', *Displays*, 58(April 2018), pp. 3–11. Available at: <https://doi.org/10.1016/j.displa.2018.07.001>.
- Arttu, T. (2018) *Effect of Visual Realism on Cybersickness in Virtual Reality*. University of Oulu. Available at: <https://jultika.oulu.fi/files/nbnfioulu-201802091218.pdf>.
- Booth, A. and Carroll, C. (2015) 'Systematic searching for theory to inform systematic reviews: Is it feasible? Is it desirable?', *Health Information and Libraries Journal*, 32(3), pp. 220–235. Available at: <https://doi.org/10.1111/hir.12108>.
- Botha, B. S. and De Wet, L. (2024) 'CyPVICS: A framework to prevent or minimise cybersickness in immersive virtual clinical simulation', *Heliyon*, 10(8), p. e29595. Available at: <https://doi.org/10.1016/j.heliyon.2024.e29595>.
- Carroll, C. *et al.* (2013) "'best fit" framework synthesis: Refining the method', *BMC Medical Research Methodology*, 13(1). Available at: <https://doi.org/10.1186/1471-2288-13-37>.
- Carroll, C., Booth, A. and Cooper, K. (2011) 'A worked example of "best fit" framework synthesis: A systematic review of views concerning the taking of some potential chemopreventive agents', *BMC Medical Research Methodology*, 11. Available at: <https://doi.org/10.1186/1471-2288-11-29>.

- Chang, E., Kim, H. T. and Yoo, B. (2020) ‘Virtual Reality Sickness: A Review of Causes and Measurements’, *International Journal of Human-Computer Interaction*, pp. 1658–1682. Available at: <https://doi.org/10.1080/10447318.2020.1778351>.
- Chang, E., Kim, H. T. and Yoo, B. (2021) ‘Predicting cybersickness based on user’s gaze behaviors in HMD-based virtual reality’, *Journal of Computational Design and Engineering*, 8(February), pp. 728–739. Available at: <https://doi.org/10.1093/jcde/qwab010>.
- Cooke, A., Smith, D. and Booth, A. (2012) ‘Beyond PICO: The SPIDER tool for qualitative evidence synthesis’, *Qualitative Health Research*, 22(10), pp. 1435–1443. Available at: <https://doi.org/10.1177/1049732312452938>.
- Dubovi, I., Levy, S. T. and Dagan, E. (2017) ‘Now I know how! The learning process of medication administration among nursing students with non-immersive desktop virtual reality simulation’, *Computers and Education*, 113, pp. 16–27. Available at: <https://doi.org/10.1016/j.compedu.2017.05.009>.
- Duzmanska, N., Strojny, P. and Strojny, A. (2018) ‘Can simulator sickness be avoided? A review on temporal aspects of simulator sickness’, *Frontiers in Psychology*, 9(NOV). Available at: <https://doi.org/10.3389/fpsyg.2018.02132>.
- Elwardy, M. *et al.* (2020) ‘Evaluation of Simulator Sickness for 360° Videos on an HMD Subject to Participants’ Experience with Virtual Reality’, in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, pp. 477–484. Available at: <https://ieeexplore.ieee.org/document/9090447/>.
- Foronda, C., Godsall, L. and Trybulski, J. A. (2013) ‘Virtual Clinical Simulation: The State of the Science’, *Clinical Simulation in Nursing*. Available at: <https://doi.org/10.1016/j.ecns.2012.05.005>.
- Gregor, S. and Hevner, A. R. (2017) ‘Positioning and Presenting Design Science Research for Maximum Impact’, *MIS Quarterly* [Preprint]. Available at: <https://doi.org/10.2753/MIS0742-1222240302>.
- Hevner, A. R. (2007) ‘A Three Cycle View of Design Science Research’, *Scandinavian Journal of Information Systems*, 19(2), pp. 87–92.
- Jenson, C. E. and Forsyth, D. M. (2012) ‘Virtual reality simulation: Using three-dimensional technology to teach nursing students’, *CIN - Computers Informatics Nursing*, 30(6), pp. 312–318. Available at: <https://doi.org/10.1097/NXN.0b013e31824af6ae>.
- Joseph, A., Browning, M. H. E. M. and Jiang, S. (2020) ‘Using Immersive Virtual Environments (IVEs) to Conduct Environmental Design Research: A Primer and Decision Framework’, *Health Environments Research and Design Journal*, 13(3), pp. 11–25. Available at: <https://doi.org/10.1177/1937586720924787>.
- Kemeny, A. *et al.* (2017) ‘New VR navigation techniques to reduce cybersickness’, *IS and T International Symposium on Electronic Imaging Science and Technology*, pp. 48–53. Available at: <https://doi.org/10.2352/issn.2470-1173.2017.3.ervr-097>.
- Li, H. *et al.* (2021) ‘The effect of navigation method and visual display on distance perception in a large-scale virtual building’, *Cognitive Processing* [Preprint], (0123456789). Available at: <https://doi.org/10.1007/s10339-020-01011-4>.
- Melo, M., Vasconcelos-Raposo, J. and Bessa, M. (2018) ‘Presence and cybersickness in immersive content: Effects of content type, exposure time and gender’, *Computers and Graphics (Pergamon)*, 71, pp. 159–165. Available at: <https://doi.org/10.1016/j.cag.2017.11.007>.



- Nowell, L. S. *et al.* (2017) 'Thematic Analysis: Striving to Meet the Trustworthiness Criteria', *International Journal of Qualitative Methods*, 16(1), pp. 1–13. Available at: <https://doi.org/10.1177/1609406917733847>.
- Padilha, J. M. *et al.* (2018) 'Clinical Virtual Simulation in Nursing Education', *Clinical Simulation in Nursing*, 15, pp. 13–18. Available at: <https://doi.org/10.1016/j.ecns.2017.09.005>.
- Servotte, J. C. *et al.* (2020) 'Virtual Reality Experience: Immersion, Sense of Presence, and Cybersickness', *Clinical Simulation in Nursing*, 38, pp. 35–43. Available at: <https://doi.org/10.1016/j.ecns.2019.09.006>.
- Wang, Y., Chardonnet, J. R. and Merienne, F. (2019) 'VR sickness prediction for navigation in immersive virtual environments using a deep long short term memory model', *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, pp. 1874–1881. Available at: <https://doi.org/10.1109/VR.2019.8798213>.
- Weech, S., Kenny, S. and Barnett-Cowan, M. (2019) 'Presence and cybersickness in virtual reality are negatively related: A review', *Frontiers in Psychology*, 10(FEB), pp. 1–19. Available at: <https://doi.org/10.3389/fpsyg.2019.00158>.

# A Human Centric Design Approach for Future Human-AI Teaming in Aviation

Barry Kirwan<sup>1</sup>, Roberto Venditti<sup>2</sup>, Nikolas Giampaolo<sup>2</sup>, and Miguel Villegas Sanchez<sup>3</sup>

<sup>1</sup>EUROCONTROL Innovation Hub, Bretigny-sur-Orge, France

<sup>2</sup>Deep Blue SARL, Roma, Italy

<sup>3</sup>Skyway, C/Albasanz, 14 Edificio Verona, 28037 Madrid, Spain

## ABSTRACT

The introduction of Artificial Intelligence (AI) in aviation has already begun, with Machine Learning systems supporting aviation workers in a number of areas. So far, such AI additions can be seen as ‘just more automation’, as the human - whether pilot or air traffic controller - remains very much in command and control, maintaining situation awareness and being the principal safety barrier against accidents. Future AI systems, however, are likely to have a higher degree of autonomy, and a collaborative relationship is foreseen. This relationship, in which the human will ‘partner’ with ‘Intelligent Assistants’, is termed Human-AI Teaming. But how will trust between human and AI be achieved? How can we develop satisfactory AI ‘explainability’ functions so the humans can understand its advice and choices? And how do we assess the human-AI interaction design, whether visual, verbal, or gestural, so that safe performance is assured? The European HAIKU project has developed a provisional Human Factors Assurance methodology and applied it to several aviation use cases that vary in terms of their AI autonomy and are consequently a reasonable testbed for evaluating new approaches. This paper outlines the methodology and illustrates its application via one of the use cases.

**Keywords:** Human factors, Artificial intelligence, Human-AI teaming, Aviation, Systems design

## INTRODUCTION

Human Factors and Aviation have been effective partners for decades. Systematic research has delivered guidance, standards and regulations in areas such as cockpit design, air traffic control/display interface design, fatigue management and crew resource management. Such guidance has helped aviation maintain its record as the safest mode of transportation, particularly when dealing with increasing levels of automation, whether in the cockpit or in the ground control systems.

The introduction of Artificial Intelligence (AI) in aviation has already begun, with Machine Learning systems supporting aviation workers in a number of areas. So far, such AI additions can be seen as ‘just more automation’, as the human - whether pilot or air traffic controller - remains very much in command and control, maintaining situation awareness and being the principal safety barrier against accidents. With the advent of future

AI systems likely to appear in the next decade, however, this is likely to change. AI systems are envisaged that will have a higher degree of autonomy. A collaborative relationship is foreseen - generically known as Human-AI Teaming - in which the human will 'partner' with 'Intelligent Assistants'. This may include the AI deciding what to do and executing its own tasks, negotiating with the human crew, and even reconsidering its goals as part of the team. This may require Human Factors to raise its game in human-system performance assurance, either upgrading its existing approach and techniques, or else adding new ones.

A first step is to understand how far AI autonomy is expected to develop in aviation in the coming decade, both from a regulatory standpoint (what is envisaged as being permissible), and a more concrete viewpoint via Human-AI Teaming aviation use cases. This step is outlined in the following two sections. A Human Factors Assurance methodology is then outlined and applied to one of the use cases to illustrate the approach and its outputs.

## **LEVELS OF AUTONOMY IN FUTURE AVIATION HUMAN-AI SYSTEMS**

The European Union Aviation Safety Agency (EASA) regulatory guidance on AI proposes six categories of future Human-AI partnerships (EASA, 2023), in increasing degrees of AI autonomy, interpreted by the authors as follows:

- Machine Learning support (1A), already existing today;
- Cognitive Assistant (1B), equivalent to advanced automation support;
- Cooperative Agent (2A), able to complete tasks as demanded by the operator;
- Collaborative Agent (2B), an autonomous agent that works with human colleagues, but which can take initiative and execute tasks, as well as being capable of negotiating with its human counterparts;
- AI Executive Agent (3A), where the AI is running the show, but there is human oversight, and the human can intervene; and
- Fully Autonomous AI (3B), where the human cannot intervene.

Human-AI Teaming raises a host of questions and challenges for Human Factors, such as how to achieve trust between human and AI, how to achieve satisfactory 'explainability' functions in the AI so the human can understand its advice and choices, as well as how best to design means of human-AI interaction, whether visual, verbal, or gestural. Such questions are often best met head-on during the assessment of use cases. Four candidate use cases are outlined next.

## **AVIATION HUMAN-AI TEAMING USE CASES**

The European HAIKU project (<https://haikuproject.eu/>) has developed a provisional methodology and applied it to several 'use cases', four of which relate to aviation operational settings:

- AI support in emergencies to a single pilot in the cockpit;



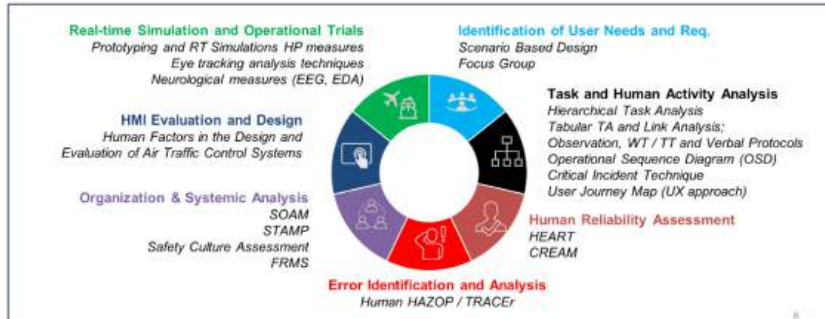
**Figure 1:** Examples of 4 HAIKU use cases.

- AI support to flight crew who unexpectedly need to divert to a different airport;
- AI support to tower controllers dealing with arriving and departing aircraft;
- An executive AI ‘manager’ of pilot-less drone and sky-taxi traffic in urban environments.

These four use cases (see Figure 1) vary in AI autonomy, and so are a reasonable testbed for applying a new Human Factors approach, as outlined below.

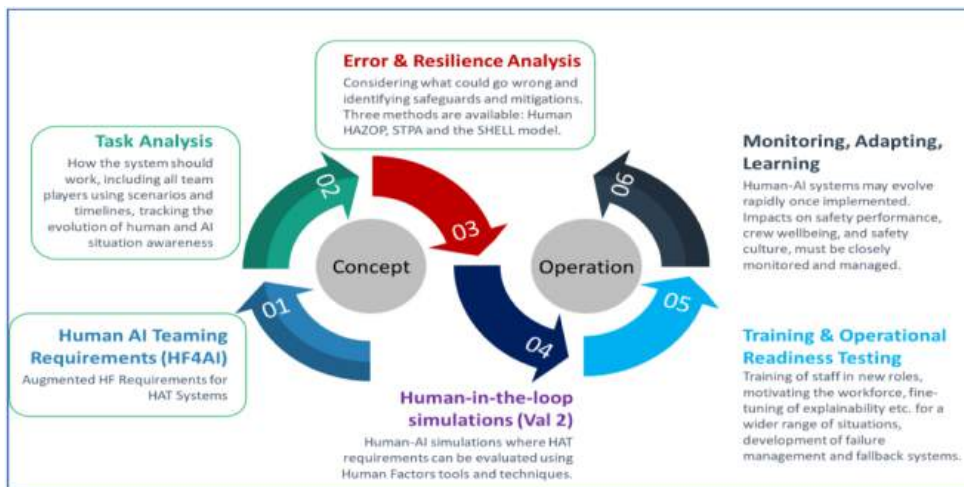
### **A HUMAN FACTORS ASSURANCE APPROACH FOR HUMAN-AI TEAMING CONCEPTS**

Brand new approaches or techniques may well be required during the next decade to deal with advanced Human-AI Teaming concepts, but currently we simply do not know what these new techniques will be. Therefore, a sensible place to start is to adopt an existing state-of-the-art framework, and to tailor and upgrade it to focus on the particular aspects of Human-AI Teaming that differ from today’s systems. The recently launched HURID (Human Risk-Informed Design) platform (see Figure 2) from the European Horizon 2020 SAFEMODE Research Project (<https://safemodeproject.eu/EhuridIndex.aspx>) serves as a contemporary and comprehensive methodological toolkit for current and future systems validation.



**Figure 2:** HURID human factors assurance toolkit (SAFEMODE project).

Figure 3 shows how this comprehensive toolkit has been adapted to focus on Human-AI Teaming concepts. Although it shows a ‘once-through’ flow of activities, in reality there are iterations and feedback loops as required by the design process. This paper focuses on the first three stages in the process: design requirements coming from Human Factors knowledge and guidance; analysis of the task to create a ‘blueprint’ of Human-AI interactions in time; and an analysis of the resilience of the system towards known failure possibilities, whether human, hardware, software or environmental in nature.



**Figure 3:** Human factors assurance process for human-AI teaming.

The assurance approach is led by a Human Factors practitioner, working with the Design Owner (the client), one or more operational users who have been able to explore / use the concept in a simulation or mock-up, and one or more of the AI developers, as illustrated in Figure 4. This ‘collegial’ approach helps ensure that the right expertise is around the table when doing analyses and determining whether design requirements are satisfied, or new ones are needed.



**Figure 4:** Key participants in human factors assurance process for human-AI teaming.

### AN ATC TOWER HUMAN-AI TEAMING USE CASE

One of the use cases, for an air traffic control tower, is used to illustrate the Human Factors approach. In this use case, an Intelligent Sequence Assistant (ISA) is being developed to support and enhance decision-making for Air Traffic Controllers. ISA optimises runway utilisation in single-runway airports, providing real-time sequence suggestions for arriving and departing aircraft. For example, in Figure 5a, the original planned sequence of use for the runway is that the RYR lands first, then the KLM takes off, and then the BAW lands.

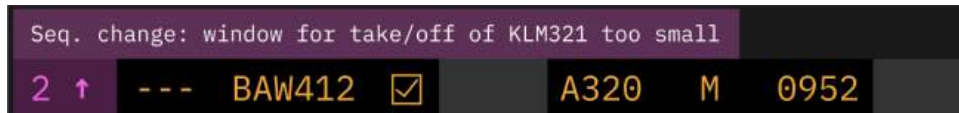
ISA computes the ordered sequence of aircraft that will use the runway. The order is displayed on the tower controller's Human Machine Interface (HMI) via numbers placed on the electronic strips of each aircraft (e.g. Figure 5b) in the controller's bay management area. If an event (e.g. the BAW accelerates faster than expected) triggers a resequencing, ISA updates the sequence in real-time, and the results are displayed on the HMI to the controller. ISA also provides explainability on-demand. For example, in Figure 5c, ISA signals that the take-off 'window' for the KLM is now too small (due to the BAW's increased speed), and so the BAW will land prior to the KLM take-off and takes position '2' in the strip.



**Figure 5a:** Extract of tower controller's HMI showing aircraft positions in real time.



**Figure 5b:** Example of electronic strip with sequence on the left.



**Figure 5c:** Electronic strip with explainability after a resequencing process.

The real-time assistance provided by ISA ensures timely and accurate forecast updates, allowing Tower Air Traffic Controllers (ATCOs) to manage the traffic flow more efficiently, with more ‘look-ahead time’ as ISA can see further ‘upstream’. The expected benefits are improved decision-making, enhanced runway utilisation, increased operational efficiency, and a safer and more streamlined air traffic flow.

## APPLICATION OF THE HUMAN FACTORS ASSURANCE APPROACH TO THE USE CASE

### Human-AI Teaming Requirements Analysis

A framework for Human Factors Assurance used frequently in European Air Traffic Management is the SESAR Human Performance Assessment Process (SESAR-HPAP)<sup>1</sup>, which has a comprehensive set of requirements in four overarching areas:

- human limitations and capabilities,
- the human-machine interface,
- teamwork and communication,
- and transition from design into operation.

Based on a literature review of Human Factors and these requirements have been augmented to evaluate Human-AI Teaming scenarios and design projects, including additional questions related to trust, explainability and AI failure, etc. which are more specific to the use of AI. In total 160 questions are posed to design projects, in order to identify Human-AI performance improvement in nine areas, as illustrated in Figure 6.

The designer, operational user, AI developer and Human Factors practitioner (the Review Team) consider whether the Human-AI Teaming Concept of Operations (Conops) satisfies each of the requirements. Sometimes they do, sometimes the requirement may be judged irrelevant for this Conops and other times it may be too soon to judge and so must be returned to later when the design is more mature or when it is being tested in a validation exercise, for example. In all other cases, a design change needs to occur to satisfy the requirement. It is such design alterations that justify the exercise and show the added value of applying the approach. Some examples are given in Table 1 below for the ISA use case.

<sup>1</sup>[https://ec.europa.eu/research/participants/data/ref/h2020/other/guides\\_for\\_applicants/h2020-sesar-er4-26-2019\\_hp-v1-3\\_16.06.05\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/guides_for_applicants/h2020-sesar-er4-26-2019_hp-v1-3_16.06.05_en.pdf)



**Figure 6:** Human factors requirement areas for human-AI teaming.

**Table 1.** Extract of human-AI requirements ‘dialogue’ within the review team.

Design of the human-machine interface supports the human in carrying out their tasks		
<i>Is all the required information presented to the user, in an uncluttered way?</i>	Y	Yes, though some of the XAI levels were unclear to the users in terms of function and information organisation
<i>Is the interaction medium used appropriate for the task and task context, e.g. keyboard, mouse/trackpad, touchscreen, voice (NLP), and even gesture recognition?</i>	Y	Yes, it is just the same as the existing airport’s current context
<i>Do visual, tactile, and oral/auditory displays, controls and interaction media comply with detailed Human Factors guidance for such devices (e.g. colour coding, luminance, auditory range etc.)?</i>	Y/N	Colour coding is based on the airport’s current workstation, and the magenta colour was chosen to differentiate ISA from all other interface components. Though, it needs to be tested in terms of contrast, luminance etc. An aural component has yet to be considered for design (and may never be, if not needed)
<i>Is the placement of a new AI screen or control/interaction system consistent with the operational workplace layout (e.g. cockpit) such that it supports rather than hinders or interferes with critical operations?</i>	Y	Yes, it is in the same screen ATCOs use now. We must make sure that panels such as XAI don’t hinder ATCO’s views.
<i>Is it ensured that any alerts, warnings or time-sensitive messages provided by the AI gain and direct the human’s attention (without startling)?</i>	Y	Yes, the alert is pretty clear.
<i>Do any alerts / warnings given signify their priority and time-urgency?</i>	Y	Yes, the blinking signifies the time urgency.
<i>Do alerts and warnings follow Human Factors guidance and principles?</i>	Y/N	We need to test this in the next Validation round.
<i>Has the fatigue impact of sustained HAT interaction been evaluated?</i>	N/A	No sustained performance required.
<i>Is it made clear to the human(s) when the alerting situation raised by the AI is resolved, or if actions are not resolving the threat?</i>	Y	Yes, the sequence is changed, and arrows appear on the electronic strip to signal the change of sequence, though they may need to be more apparent as some participants missed those during Val 1
<i>Is the interaction with the AI seen as user-friendly?</i>	Y/N	Pilots like several aspects but more work required (see above)

**Task Analysis: Operations Sequence Diagrams**

To create a ‘blueprint’ of Human-AI interaction, a task analysis format known as Operations Sequence Diagram (OSD) is applied, as it usefully



shows what both the human and the AI think/calculate and do in real time as the scenario develops. An extract is shown in Figures 7a–7c (normally displayed on one line).

Time	Actual System State	Goal	Human1	Info sources (non-AI)
T0	Aircraft data suggest that an aircraft (BAW412) is going faster than expected and the sequence needs to be reshuffled.	Resequencing of the aircraft's order to relieve ATCO's workload.	ATCO is not aware of the possible changes at the moment.	Data from aircraft.

**Figure 7a:** OSD Part 1.

Operator believed system state	AI believed system state	AI solution	AI HMI	AI Rationale (XAI)
ATCO knows that the assistant will provide support if needed.	ATCO is unaware of the aircraft state.	AI will calculate the re-sequencing of the two aircraft.	The AI HMI remains the same at the moment displaying the sequence.	While resequencing is ongoing the explanation is being constructed.

**Figure 7b:** OSD Part 2.

H-AI Dialogue	Authority gradient	Decision / Action	Human Performance Impact:
ISA's resequencing will start automatically. The ATCO supervises and makes decisions based on these suggestions.	ATCO is in control supervising the state of the system and the sequence.	ATCO is unaware of need for resequencing at this time.	No negative impact on Human Factors in this scenario. The AI is simply monitoring the aircrafts and the ATCO is unaware.

**Figure 7c:** OSD Part 3.

The OSD shows the flow of information between human and AI and can highlight where something needs to be signalled to the human user, or where the human and AI may be 'out of synch'. It gives context when carrying out the requirements analysis, and when evaluating resilience of the Human-AI Teaming Conops.

## **Resilience Analysis: Application of the SHELL Method**

The Methodology in Figure 3 shows three candidates for evaluating the resilience of a Human-AI Teaming Conops: SHELL, HAZOP and STPA. In HAIKU, all three are being applied to various use cases, and for the ISA example both HAZOP and SHELL have been applied, with the SHELL method shown here.

The SHELL model, developed by Hawkins (1982), examines five key elements that interact to influence human performance: software (procedures, protocols), hardware (equipment design), environment (contextual factors surrounding human activity), liveware (human characteristics), and their interconnectedness. The analysis of these elements and their interactions reveals vulnerabilities that may lead to accidents. The challenge lies in pinpointing these vulnerabilities within the expanding operational landscape of aviation. This is where the integration with the OSD (see Figure 7a, b, c above) proves valuable, providing a structured map of the system's operational sequence. Integrating each operational step in the OSD with the SHELL analysis significantly reduces complexity in hazard identification. Examples of guidelines gained from the SHELL analysis are given below.

### **[SOFTWARE] Operational Resequencing**

#### **What Happens When There Is a Sequence Change During Nominal Operations? What Is the Dialogue Between the Human and the AI?**

In this scenario, the Software Element of the SHELL model is explored, focusing on the sequencing/resequencing procedure. An aircraft suddenly picks up speed, leading to resequencing. Both compliant (ATCO follows ISA suggestions) and non-compliant (ATCO triggers resequencing due to disagreement) situations are explored. This scenario highlighted the need for ISA to adapt to the controller actions, and vice versa the need for the controller to maintain situation awareness about ISA.

### **[HARDWARE] Connection Error**

#### **What Happens to the Sequence When There Is a Technical Failure? What Is the Dialogue Between the Human and the AI?**

This scenario investigates the consequences of a technical failure in the hardware component of the system, specifically the connection between the AI system and the HMI. This failure leads to sequence changes not being transmitted to the interface: the calculation is not updated, and the controller is not updated about a change in the sequence. This scenario emphasises the critical need for reliable data transmission between ISA and the HMI, and for ways to allow the controller to always acknowledge the status of the system and its connection. It also highlights that controllers can never be too over-reliant on the system.

### **[ENVIRONMENT] Handling Unexpected Traffic**

#### **What Happens When an External Factor Influences the Sequence? What Is the Dialogue Between the Human and the AI?**

The scenario explores how ISA deals with an external environmental factor outside of the ATCO's control, such as unexpected traffic (e.g., unannounced Flight School operations). When an unforeseen operation starts and new aircraft appear on the radar, they don't come with an auto-generated electronic strip and are not automatically inserted into the sequence. This implies that controllers need to manually create a strip and they somehow need to integrate it into the current sequence. Thus, this OSD highlighted the need for ISA to adapt to unforeseen operations by having a function that automatically organises strips, even when added manually by controllers.

### **[LIVEWARE] Handover-Takeover Procedure**

#### **What Happens If There Is a Resequencing Process During the Handover Event?**

This OSD studies how events unfold during a handover/takeover procedure (when a new controller comes to the sector and replaces one of the controllers in position). This event falls in the Liveware category as is about human-human interactions. In this case, the incoming controller must create a mental map immediately. This OSD highlighted the need to add a short briefing about ISA in the checklist of the handover procedure, and in general, the need for ATCOs to be able to immediately understand the current status of the sequence, even if they had just sat down.

The SHELL analysis, together with OSDs for each scenario, usefully brought to life several key scenarios that required design considerations, which could be added to those found via the Human Factors Requirements Analysis.

## **CONCLUSION**

The approach outlined in this paper has been applied to two of the HAIKU use cases so far (the other one was cockpit-based) and has led to design changes and considerations welcomed by the design team and end users. The approach itself is relatively light, with each analysis taking typically 1-5 days, plus some preparation. The design maturity does not have to be high, making the approach useful for early consideration of requirements, though it does work best if there is a simulation or mock-up that licensed operational personnel can use and test to see how it works. The approach will be applied to two further HAIKU use cases later in the year. The use of a review team approach can also track whether the operational users feel their role and motivation might be affected by the introduction of the AI into their workspace, as an over-riding goal of HAIKU is to support human-centric AI.

The approach appears fit for purpose, but as AI develops and advances, it is probable that new techniques, new methodological frameworks and even paradigms may be required. Until then, the approach allows a useful and productive dialogue between designer, operational user, AI developer and Human Factors performance experts, helping ensure that new AI systems can be integrated safely and successfully.

## **ACKNOWLEDGMENT**

This publication is based on work performed in the HAIKU Project which has received funding from the European Union's Horizon Europe research and innovation programme, under Grant Agreement no 101075332. Any dissemination reflects the authors' views only, and the European Commission is not responsible for any use that may be made of information it contains.

## **REFERENCES**

- EASA (2023) EASA Concept Paper: first usable guidance for level 1 & 2 machine learning applications. February. <https://www.easa.europa.eu/en/newsroom-and-events/news/easa-artificial-intelligence-roadmap-20-published>
- Hawkins, F. H. (1993). *Human Factors in Flight* (H. W. Orlady, Ed.) (2nd ed.). Routledge. <https://doi.org/10.4324/9781351218580>

# Analysis and Interview Survey to Detect Subjective Fatigue and Accident Risk of Truck Drivers

Nao Ito, Takeshi Tanaka, Yun Li, Shunsuke Minusa,  
and Hiroyuki Kuriyama

Research and Development Group, Hitachi Ltd., Kokubunji, Tokyo, Japan

## ABSTRACT

**Methods:** In this study, interview survey and analysis were conducted with the aim of detecting accident risk from subjective health data of drivers. For interview survey, we observed time-series changes in VAS for 1956 individuals over a year and selected 11 drivers who showed a significant worsening of subjective health. The 11 drivers were asked two types of questions. The first was, “Are there any factors that you are aware of about the period of time during which you showed significant worsening?” The second was, “What factors do you think are present that could have a dangerous effect on your driving?” The Analysis phase of the study examined whether the use of subjective health information and additional information would be useful in detecting accident risk. From the first question, we defined four patterns of VAS worsening trends and analyzed the relationships between these patterns and accident risk. From the second question, the index “change in the time of work start” was derived as a factor that many drivers consider dangerous. We then analyzed the correlation between this new index and the near-miss rate.

**Results:** For the relationship between VAS and accident risk, it was found that two of the four VAS risk patterns had a significant negative correlation with the rate of near-misses. Furthermore, analysis of the relationship between “change in the time of work start” and the accident risk revealed a significant negative correlation when the absolute value of “change in the time of work start” was within  $\pm 6$  hours. This means that the rate of near-misses increases when the workday starts earlier than the previous day.

**Conclusion:** To detect a hazard leading to a driving near-miss with the VAS data alone, the worsening would have to continue for long time over 4 or 8 weeks. However, the newly discovered convince indicator “change in the time of work start” is a feature with a short span, and its addition to the VAS and accident risk analysis may improve the accuracy of health risk detection.

**Keywords:** Fatigue, Stress, Risk management, Visual analogue scale, Driving accidents

## INTRODUCTION

In the transportation industry, employee health management is considered important as well as compliance with safe operation. Truck drivers, in particular, are subject to a high-load working environment with

irregular working hours, and the Japan Trucking Association is actively promoting health management initiatives. In addition to basic physical condition checks such as temperature and blood pressure at roll call, many transportation companies have recently introduced technology to objectively and subjectively assess drivers' health conditions such as daily fatigue and stress. For objective evaluation, ANF (Autonomic Nerve Function) indices using pulse data obtained from wearable devices are mainly used. From previous studies, we have found an association between ANF and the incidence of near-misses while driving among truck drivers (Mizuno et al., 2020, Minusa et al., 2021 and Ito et al., 2023). ANF has been shown to be related to accident risk, but may differ from the driver's subjective health status. The use of subjective health information is also essential in terms of creating alerts that are acceptable to the drivers. Subjective health assessment may be conducted using a VAS (Visual Analogue Scale) via smartphones, etc. VAS is intuitive and easy to interpret in assessing health status. A previous study examined the relationship between driver's subjective health condition and driving performance and showed a significant correlation between VAS of fatigue and standard deviation of the bus lateral position (Gharagozlou et al., 2015). However, compared to objective assessments, changes in VAS vary widely among individuals. Especially for long-term changes across days, some people's values do not change at all, while others increase or decrease significantly. It is difficult for managers of drivers to identify which trends to monitor and when. In this study, interview surveys and analysis were conducted with the aim of detecting driver's accident risk from the changes of driver's daily VAS.

## MEASUREMENT DATA

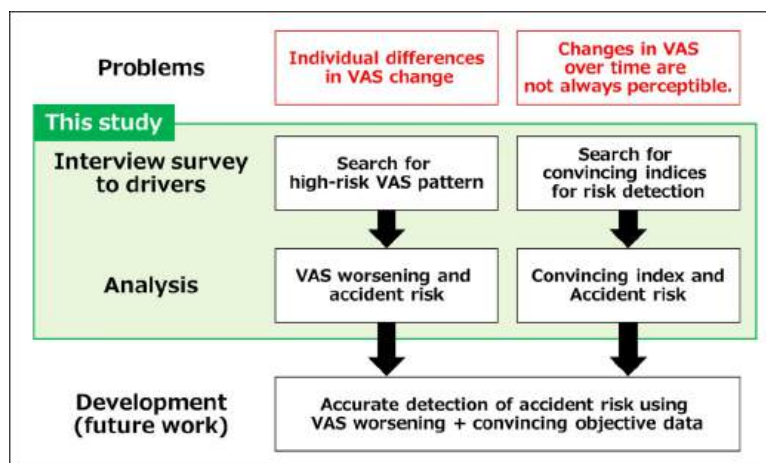
In this study, we used "SSCV-Safety," a safe operation support system provided by LOGISTEED, Ltd., and collected data of 1956 drivers from January 2021 to June 2023. This system obtains driver's subjective health data in VAS format before driving with smartphone and measures near-miss event data during driving with sensors mounted on the truck. As VAS data, we obtained the values of responses to two types of questionnaires. The first type is the response to the "Are you tired?" questionnaire, which we named Fatigue VAS. The second type is the response to the "Did you sleep well last night?" questionnaire, which we named Sleep VAS. In addition to VAS data, we also obtained information on near-misses that occurred during driving as accident risk. A near-miss is an event such as a "sudden stop" that is automatically detected by acceleration and other sensors installed in the vehicle. Table 1 shows the types of near-miss scenes that were targeted for detection in this study. These near-miss scenes were defined on the SSCV system, and were automatically detected by the system's own logic based on acceleration and other onboard sensor data.

**Table 1.** Details of type of measured incident data during driving.

Near-miss no.	Name of near-miss scene
1	Temporary non-stop
2	Sudden steering
3	Sudden Deceleration
4	Distance warning
5	Insufficient distance
6	Impact
7	Forward collision warning
8	Overspeed

## OVERALL VIEW OF THE PROPOSAL METHOD

Figure 1 shows the overview of this study's proposal method. We aimed to detect accident risk of drivers using subjective health data, but there are two challenges. The first is to achieve highly accurate detection, and the second is to ensure that users are satisfied with the detection results. The first challenge is due to individual differences in changes in VAS input data; the second challenge is because changes in subjective health over time are not always perceptible to the driver. To solve these issues, we decided to begin our research with an interview survey of drivers. The interview survey asked the drivers two different questions, and from the responses obtained, two analyses were conducted. To address the first challenge of individual differences in changes in VAS, we searched for high-risk VAS patterns common to many drivers and then analyzed relationship between the patterns and accident risk. To address the second challenge of satisfying drivers with the detection results, we also searched for and analyzed convincing index for drivers that can be used as a supplement to accident risk detection by VAS.

**Figure 1:** Overview of this study's proposal method.

## METHOD 1: INTERVIEW SURVEY TO SEARCH FOR SUBJECTIVE HEALTH RISK AND OBJECTIVE INDEX FOR ACCIDENT RISK DETECTION

For interview survey, we comprehensively observed cases of VAS worsening over time from 1956 drivers and selected 11 drivers who showed a significant worsening. Some of them had significant VAS worsening in a short period of time, others slowly worsened over time, and still others had increasing and decreasing input values. The 11 drivers were asked two types of questions. The first was, “Are there any factors that you are aware of about the period of time during which you showed significant worsening?” and we examined what patterns of worsening are perceived by the drivers. As a result, four drivers answered “Yes” to the question. Table 2 shows the VAS worsening types of the four subjects and the factors answered by each of them. Clear factors and their background such as characteristics of work shift and physical condition changes were answered. The second question was, “What factors do you think are present that could have a dangerous effect on your driving?” and we checked what risk factors were considered by drivers whose subjective fatigue was worse. Several drivers answered that a sudden change in the timing of work start disrupts their sleep cycle and make drivers nervous about the driving.

**Table 2.** Types and factors for four drivers who were aware of worsening of VAS.

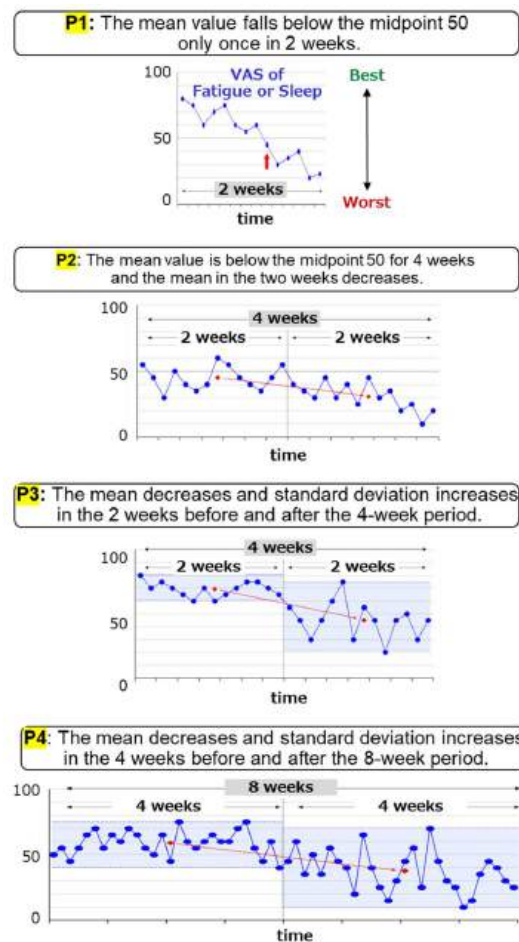
VAS worsening type	Comments on factors behind the worsening of VAS
Fatigue and Sleep VAS decreased in 2 weeks	“It was the end of a long holiday weekend. I always feel tired after a holiday weekend.”
Fatigue and Sleep VAS decreased in 4 weeks	“It was the week of the early shift. During the early shift week, the closer it gets to the weekend, the more tired I get.”
Fatigue VAS decreased and standard deviation increased in 4 weeks	“In the latter half of the detection period, I was not feeling well. I had a fever while driving and visited a hospital after the work.”
Fatigue VAS decreased and standard deviation increased in 8 weeks	“The fatigue worsened on the day when I had to go to work suddenly earlier. I felt strongly fatigued.”

## METHOD 2-1: ANALYSIS OF VAS WORSENING AND NEAR-MISS OCCURRENCE

Because four drivers of different types answered there were factors during the VAS worsening from the interview survey, four patterns of VAS worsening trends were defined as subjective health risks (Figure 2). Patterns 1 to 4 are denoted as P1 to P4, respectively; P1 represents that the mean value falls below the midpoint 50 only once in 2 weeks; P2 represents that the mean value is below the midpoint 50 for 4 weeks; P3 represents that the mean decreases and standard deviation increases in the 2 weeks before and after the 4-week period; P4 represents that the mean decreases and standard



deviation increases in the 4 weeks before and after the 8-week period. These patterns were applied to obtained Fatigue and Sleep VAS data to detect subjective health risk. To examine the relationship between subjective health risk and accident risk, near-miss rates were calculated. Near-miss rates refer to the probability of one or more near-misses occurring in a day. Data were extracted only for drivers for whom VAS risk was detected at least once. The data for each driver was divided into sections where VAS risk was detected and not detected, and the median near-miss rate for the two sections was calculated. The near-miss rates of the detected and non-detected data were then compared with a confidence interval of  $P < 0.05$ .



**Figure 2:** Examples of four defined patterns of VAS worsening trends.

## METHOD 2-2: SEARCH FOR NEW INDICATORS AND ANALYSIS

About the second question in interview survey, several drivers answered that a sudden change in the timing of work start disrupts their sleep cycle, making them more likely to feel at risk for driving. We then calculated a new index,

“change in the time of work start,” and analyzed its correlation with the accident risk. The following procedure was used to calculate the “change in the time of work start.” First, since drivers input VAS information every morning when they arrive at work, the time of VAS input was used as the time of arrival at work. Next, the time of arrival at work (0:00-23:59) was converted to the elapsed time [minutes] from 0:00 of the day, and then divided by 1440, which is minutes of one day, to obtain a “time of work start” with a minimum value of 0 and a maximum value of 24. Furthermore, the difference of “time of work start” between one day and previous workday was calculated as “change in the time of work start.” The minimum of the index is -12 and the maximum is +12. For example, if a driver’s work time one day was 7:00 a.m. and the previous day’s work time was 8:00 a.m., the “change in the time of work start” would be -1 [hour]. If the previous workday fell on more than one holiday, such as a Saturday, Sunday, or consecutive holidays, it was excluded from the analysis. At this point, we checked the data distribution and found that the larger the absolute value of “change in the time of work start,” the smaller the number of data. Therefore, to ensure the reliability of the statistical analysis, we decided to analyze the data within  $\pm 6$  hours, which included at least about 100 data. Even within  $\pm 6$  hours, there was a bias in the number of data depending on the value of “change in the time of work start.” Therefore, we calculated the mean rate of near-misses for each hour of the “change in the time of work start” interval. Finally, a single regression analysis using the least-squares method was conducted for the “change in the time of work start” and the near-miss rate, with a confidence interval of 95%.

## RESULT

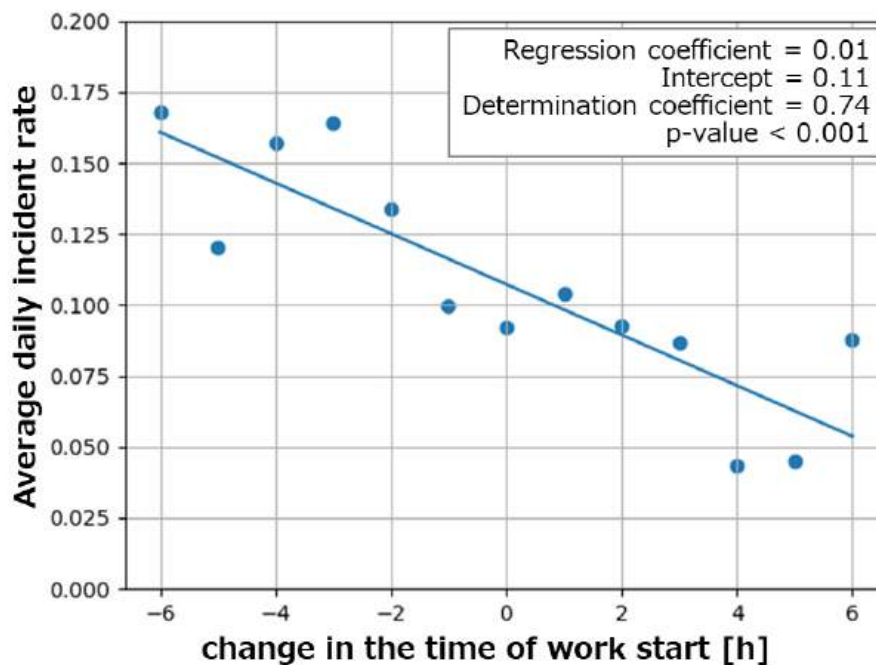
First, we present the results of the analysis on the correlation between subjective health risk and accident risk for the four defined VAS patterns. Table 2 shows the number of data in which VAS risk was detected in each pattern. Sleep VAS of P3 and P4 were barely detected. Table 3 shows the results of the correlation between VAS worsening trends and near-miss rates for each of the four patterns. A significant correlation was found for P3 and P4 worsening trend of the Fatigue VAS. Furthermore, as shown in Figure 2, the correlation between “change in the time of work start” and accident risk was confirmed, showing a significant negative correlation ( $P < 0.001$ ).

**Table 3.** Number of VAS risk detections for 4 patterns.

Pattern (Length)	Number of target data	Number and rate of detection with fatigue VAS	Number and rate of detection with sleep VAS
P1 (2 weeks)	6054	237 (3.9%)	244 (4.0%)
P2 (4 weeks)	6767	993 (14.7%)	885 (13.1%)
P3 (4 weeks)	6609	269 (4.1%)	3 (0.0%)
P4 (8 weeks)	1633	70 (4.3%)	0 (0.0%)

**Table 4.** Results of VAS risk detection and significance different test of near-miss rate.

Pattern (Length)	VAS type	Near-miss rate in detected (mean±SD)	Near-miss rate in NO detected (mean±SD)	P value
P1 (2 weeks)	Fatigue	3.1%±12.0%	1.6%±5.1%	0.373
	Sleep	4.2%±12.9%	2.8%±13.0%	0.419
P2 (4 weeks)	Fatigue	5.4%±14.0%	4.0%±7.7%	0.196
	Sleep	7.2%±16.8%	3.3%±6.8%	0.058
P3 (4 weeks)	Fatigue	9.6%±18.9%	2.9%±7.0%	<0.001*
	Sleep	-	-	-
P4 (8 weeks)	Fatigue	9.7%±15.3%	5.6%±8.5%	0.028*
	Sleep	-	-	-

**Figure 3:** Result of regression analysis of “change in the time of work start” and near-miss rate.

## DISCUSSION

The analysis on the presence or absence of detectable VAS worsening trends and near-miss rates showed that P3 and P4 VAS worsening pattern for Fatigue VAS was significantly related to accident risk. The two patterns have one thing in common: the standard deviation increased at the same time as the mean of the VAS decreased. Truck drivers who travel long distances have irregular work conditions such as destinations, routes, and work hours may vary. The repetition of short periods of intense fatigue and its recovery may have led to the accumulation of chronic fatigue, which in turn led to accident risk. We could not verify the Sleep VAS for P3 and P4 due to lack of detection

drivers. However, with additional data in the future, we may be able to confirm the relationship with accident risk for sleep as well. It has been shown in previous studies that subjective sleep quality has a significant relationship with near-misses while driving (Malish et al., 2016 and Alshareef, 2020). This study is not limited to non-time series data as in previous studies, but indicate that changes in subjective health status over time may also affect the increased risk of accidents. In addition, we found a new indicator, the “change in the time of work start.” The relationship between the new index and the accident risk was shown to have a negative correlation at  $P < 0.001$ . This indicates that the earlier the start of work than the previous day, the higher near-miss rate. Nurses in Japan work in shifts, and there are three types of work schedules: day shift, semi-night shift, and deep night shift. Counterclockwise changes in work schedules, such as deep night to semi-night, are deprecated because they violate the human circadian rhythm (Kubo et al., 2013). Truck drivers often have more irregular work hours than nurses, but they may share the same risk of coming to work earlier than the previous time.

## CONCLUSION

In this study, we conducted an interview survey and analyzed the relationships between subjective health risk, a newly discovered index of “change in the time of work start,” and accident risk. Our findings indicate that the relationship between subjective health risk and accident risk is confirmed in two of the four patterns. Furthermore, the detection of accident risk using VAS was shown to be potentially highly accurate by “change in the time of work start,” which was newly discovered from the interview survey. The findings of this study may inform the health management of occupational drivers, including truck drivers.

## ACKNOWLEDGMENT

The authors would like to acknowledge helpful discussions with Dr. Hirohiko Kuratsune, Inviting Professor at the Osaka Metropolitan University School of Medicine and Dr. Shigeru Haga, Professor Emeritus at Rikkyo University.

## REFERENCES

- Alshareef, S. (2020). Excessive daytime sleepiness and associations with sleep-related motor vehicle accidents: results from a nationwide survey, *Sleep and Breathing* Volume 25.
- Gharagozlou, F. Mazloumi, A. Saraji, N. G. Nahvi, A. Ashouri, M. Mozaffari H. (2015). Correlation between Driver Subjective Fatigue and Bus Lateral Position in a Driving Simulator, *Electronic Physician* Volume 7, No. 4.
- Ito, N. Minusa, S. Tanaka, T. Li, Y. Kuriyama, H. (2023). Prediction of Future Collision Risk for Truck Drivers Using the Time-Series Autonomic Nerve Function, *IEEE Access* Volume 11.
- Kubo, T. Takahashi, M. Sallinen, M. Kubo, Y. Suzumura, H. (2013). How are leisure activity and shiftwork schedule associated with recovery from fatigue in shiftwork nurses? *Sangyo Eiseigaku Zasshi* Volume 55, No. 3.
- LOGISTEED, Ltd. SSCV-Safety. <https://www.logisteed.com/jp/sscv/safety/>

- 
- Malish S, Arastu F, O'Brien LM. (2016). A Preliminary Study of New Parents, Sleep Disruption, and Driving: A Population at Risk?, *Matern Child Health Journal* Volume 20, No. 2.
- Minusa, S. Mizuno, K. Ojio, D. Tanaka, T. Kuriyama, H. Yamano, E. et al. (2021). Increase in rear-end collision risk by acute stress-induced fatigue in on-road truck driving. *PLoS One* Volume 16, No. 10.
- Mizuno, K. Ojio, D. Tanaka, T. Minusa, S. Kuriyama, H. Yamano, E. et al. (2020). Relationship between truck driver fatigue and rear-end collision risk. *PLoS One* Volume 15, No. 9.

# Revolutionizing Automotive Industry for Servicing an Autonomous Adaptive Lift System

Haissam El-Aawar and Omar Mohammad

Lebanese International University, Bekaa, Lebanon

## ABSTRACT

The automotive industry stands at the cusp of a transformative innovation with the development of an autonomous adaptive lift system designed to revolutionize vehicle servicing. Traditional automotive lifts require manual adjustments to fit various car sizes and positions, often leading to inefficiencies and safety concerns. In response, our proposed model introduces an autonomous lift system that detects and adjusts to the optimal lifting points for diverse vehicles, eliminating the need for human intervention in positioning. Powered by cutting-edge hardware components, including an Arduino controller serving as the main board, REV kit for motor and extruder assembly, and precision servomotors, this system incorporates advanced sensor technologies like IR (infrared), ultrasonic sensors, and specialized point detection sensors. This comprehensive sensor array enables the lift to comprehensively assess a vehicle's dimensions, weight distribution, and ideal lifting positions upon entry. By harnessing an amalgamation of sensors, actuators, and programmed instructions within the Arduino controller, the lift autonomously adapts to diverse vehicle shapes and sizes, facilitating precise and safe adjustments. This integrated system, comprising sensors for dimensional analysis, actuators for responsive adjustments, and sophisticated programming, enables the lift to dynamically conform to various vehicles. Furthermore, stringent safety protocols, including failsafe mechanisms and emergency stop functionalities, are inherent to this hardware setup, ensuring unwavering reliability and operational security. This research aims to elucidate the pivotal role of this autonomous adaptive lift system in reshaping the automotive servicing landscape. It will delve into the technological innovations, operational functionalities, and the added value this system brings to garage operations, emphasizing its potential to elevate safety standards, increase efficiency, and transform the industry's approach to vehicle maintenance.

**Keywords:** Autonomous adaptive lift, Efficiency enhancement, Precise alignment, Safety, Vehicle lift

## INTRODUCTION

Human Factors Engineering involves understanding the need for comprehensive the integration of autonomous systems in automotive services considered a significant aspect in the enhancement process of operational services efficiency and safety. while the traditional practices of vehicle maintenance using a lifting system must undergo a transformative approach

toward automatic functions. Numerous studies have been undertaken in Spain to address the public priority of preventing workplace injuries (Park, 2023). These studies have focused on various topics, such as trends over time for fatal injuries caused by occupational accidents in all industries (Santamaría, 2006). Occupational injuries can cause people to lose their jobs temporarily or permanently, resulting in a fall in labor earnings. Furthermore, businesses are burdened with various direct and indirect costs, such as compensation payments to injured workers, repair of damaged materials (machinery and assets), production delays, fines and criminal penalties, employment and training costs for replacement workers, and a decrease in the company's credibility (Heinrich, 1980). According to the European Union's Information Agency for Occupational Safety and Health (EU-OSHA, 2017), occupational injuries cause economic losses of €2.68 trillion globally (3.9% of the world's GDP) and €476 billion in the EU 28 countries (3.3% of the total GDP) (Dietmar, 2017). Furthermore, the projected economic losses from occupational injuries in the Republic of Korea range between 3.35% and 5.91% of the country's GDP (Heinrich, 1980). Furthermore, according to data from the United Nations International Labor Organization (ILO) from 2015 to 2018, the Republic of Korea has the third highest occupational injury rate after Turkey and Mexico. As a result, occupational injury prevention has emerged as a critical policy concern for the government to solve. Our research will focus on technological innovation, the consequential added value to garage car services, safety, efficiency, and industry transformation.

## LITERATURE REVIEW

Technological advancements have significantly transformed automotive servicing, evolving from labor-intensive manual processes to highly precise automated services. In the early days of the automotive industry, a simple peg-based jack was used for vehicle maintenance. However, as automobiles grew heavier, this type of jack became unsuitable due to the risk of collapse. It was eventually superseded by the scissor jack, now the standard equipment provided by manufacturers for vehicle lifting (Yekinni, 2018). The development of a motorized rack and pinion car lift system was introduced in 2014. This system features a combination of flat and round helical gears in a rack and pinion arrangement, with additional spur gears connected by a chain drive, all powered by a 12V battery (Thirugnanam, 2014). Further enhancements to the scissor or toggle jack design have been made by incorporating an electric DC motor, as noted in studies by Osman and Ehab (ekinni, 2018). An integrated automated hydraulic automobile jack system for four-wheelers has been presented, which includes a hydraulic pump powered by an electric motor and a hydraulic cylinder for vehicle lifting. This mechanism enables the activation of the automobile jack with a single push button, allowing hydraulic jacks to be activated separately for each side of the vehicle during breakdowns. The hydraulic system has an advantage over pneumatic systems due to the incompressibility of oil, which provides greater lifting capability. To efficiently handle the hydraulic circuit,

the system uses single-acting cylinders controlled by control valves, as well as a relief valve (Pradhan, 2022). This document outlines the built-in jack in automobiles. The hydraulic jack system is coupled to the automobile's front and rear chassis. It may be simply fitted to various automotive chassis and frames. A hydraulic jack is centrally located in the front and rear suspensions of automobiles, between the wheels. It uses hydraulic power and requires a 12v DC to function. Implementing this approach simplifies vehicle maintenance, particularly for heavier vehicles. The hydraulic jack mechanism operates based on Pascal's principle. The pressure remains constant across a closed container or cylinder at all points. Connect two cylinders, one tiny and one huge. The force exerted by a huge cylinder increases as its area increases, as long as the pressure remains constant. The equation  $P=F/A$  leads to  $F=PA$ . Hydraulic jacks employ oil pressure to lift and lower automobiles by moving the handler. Hydraulic fluids serve as a driving force in hydraulic machinery. Hydraulic cylinders run on hydraulic fluid. In this arrangement, the hydraulic fluid does not absorb the energy (Sathyan, 2019). Designed an automobile jack that is permanently welded to the car chassis and can be lifted using a wiper motor. The motor responds to commands from a mobile app. Redesigned the car's chassis model to line the jack with the chassis. This prevents the jack from hitting any road obstacles (Anon., 2020). But, for maintenance we still need the two-post lifter to be comfortable as in (Québec, 2021) reveals that Mechanics use automotive lifts to raise vehicles off the ground for maintenance and repair. In Quebec, where this survey was conducted, the joint sector-based group estimates that roughly 30,000 lifts are installed in the automotive services sector. By the end of 2020, the province had licensed 5.4 million vehicles and light trucks. Nowadays, two-post above-ground (2PAG) lifts are the most prevalent in garages because they provide unobstructed access under the car (Figure 1) and are suitable for normal maintenance (e.g., tire changes, oil changes, etc.). The ability to change all four tires at once which is not achievable with four-post lifts. ((ALI), 2008). Moreover, the author creates a new contribution using an experimental design, support pad slippage on a 2PAG lift caused by external forces acting on a vehicle. this will help in the grip but still need the intervention of labor to direct the two posts under the car. (Burlet-Vienney, 2021).



**Figure 1:** 2PAG with added pad (Burlet-Vienney, 2021).



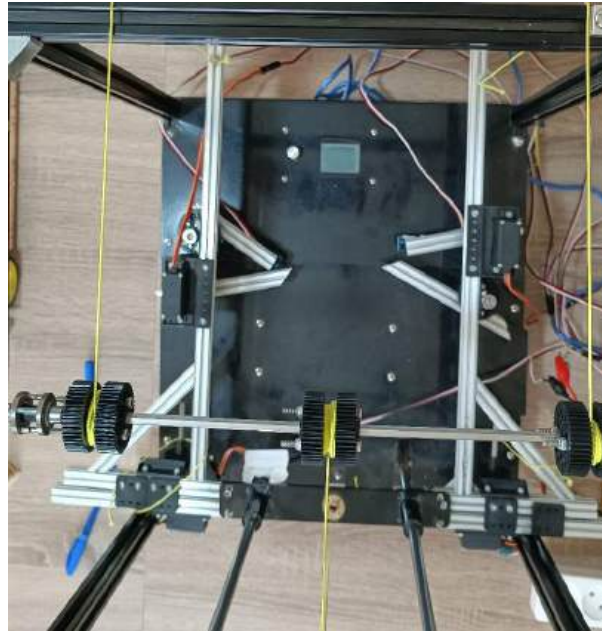
## METHODOLOGY

To develop and evaluate an automated alignment and lifting system for two-post automotive lifts, we conducted an experimental study that integrated advanced sensor technology, precise motor controls, and a microcontroller to automate the vehicle lifting process with unparalleled precision and safety. The system's core components, including an Arduino Uno microcontroller, infrared (IR) sensors for under-car detection, servo motors coupled with gear lines for precise arm movement, a Gear DC motor for the lifting mechanism, and an RGX motor drive for high current demand handling, were specified and integrated. Software control algorithms were developed in C++ using the Arduino IDE, focusing on automating the alignment of lift arms with vehicle-specific markings (colored blocks or circle) for accurate lifting point identification. Upon the vehicle's entry, the two-post lift arms, initially positioned on the ground, commence their precision movement beneath the vehicle, facilitated by a servo motor and a gear line system. The moment the integrated infrared (IR) sensors detect the lift arms have successfully positioned themselves under the vehicle, the movement halts, initiating the process of searching for the predefined alignment points. If these points are not immediately located, the lift arms undergo a slight forward adjustment, continuing to oscillate open and close in their search for the correct points. Upon identifying the optimal alignment points on both sides of the vehicle, flags within the system's code are triggered, signifying the precise alignment of the two posts. Following this confirmation through sensor data, the system engages a Gear DC motor, meticulously adjusting the torque to accommodate various vehicle types. This enables the lift to elevate the car to the desired height securely. Operators then have the flexibility to further adjust the vehicle's height as needed, using simple up and down buttons, ensuring the lift's arms are perfectly positioned at the vehicle's designated points. This entire process, fully automated, significantly minimizes errors in vehicle alignment on the lift, enhancing both safety and efficiency in automotive servicing tasks. A prototype was constructed to embody this design, followed by a series of functional tests within a controlled environment that simulated real-world conditions with vehicles marked for alignment testing. These tests assessed the prototype's ability to accurately move under the vehicle, detect the correct alignment points via IR sensors, and adjust the arm positioning as necessary until successful alignment was confirmed. The system demonstrated high accuracy in aligning with the marking points across different vehicles, efficiency in operation with minimal adjustments needed, and enhanced safety with a stable lifting process. The entire testing phase confirmed the system's functionality, proving its capability to significantly improve the precision, efficiency, and safety of vehicle lifting in automotive servicing operations, thereby confirming the success of the developed methodology in achieving the project's objectives.

## IMPLEMENTATION

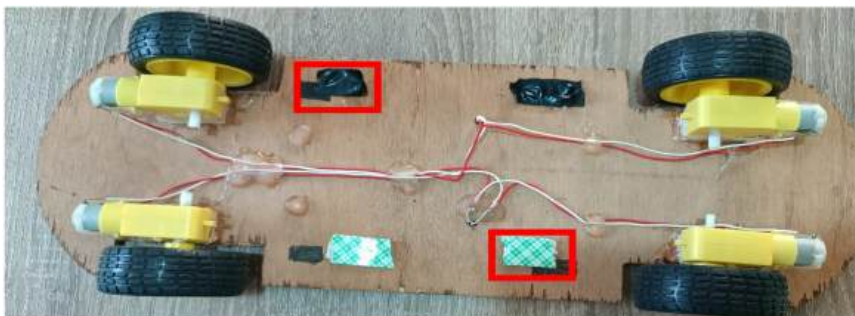
The implementation of the automated two-post lift system began with the meticulous assembly of its core hardware components, including servo

motors, gear lines, and infrared sensors. A comprehensive control program was developed to manage the entire process, from the car's entry to the precise alignment of the lift arms, ensuring a seamless and highly efficient operation. The integration of servo motors and gear lines, ingeniously designed to facilitate precise horizontal movement beneath the vehicle. This crucial step ensured the lift arms could accurately navigate and position themselves for optimal alignment as shown in Figure 2.



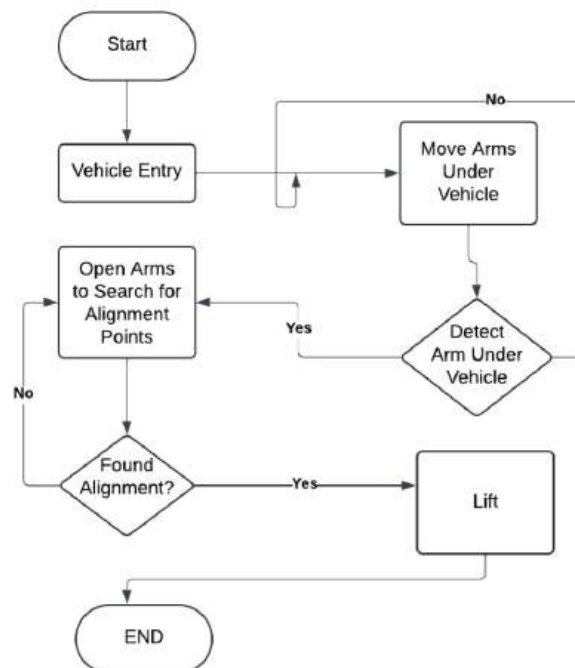
**Figure 2:** The complete setup of the automated two-post lift system.

Simultaneously, infrared (IR) sensors were installed and meticulously calibrated to detect the exact moment the lift arms slid into place under the car, marking the beginning of the automated alignment phase, see Figure 3.



**Figure 3:** The marking blocks for accurate point identification.

Following hardware setup, we connected an Arduino micro-controller that was programmed to control the entire lifting operation: from activating the servo motors and gear lines for initial positioning, processing the IR sensor input for accurate detection under the vehicle, to controlling the Gear DC motor responsible for adjusting torque according to the vehicle's weight. This programming phase was critical, involving the development of sophisticated detection and alignment algorithms. These algorithms processed the data from the IR sensors, directing the system to adjust the lift arms' position incrementally until the designated alignment points on the vehicle were located, see Figure 4. Upon successful alignment, indicated by specific flags in the system's code, the lifting process was initiated. the flowchart below describes the whole process.



**Figure 4:** The decision-making process employed by the system in response to sensor data.

## CONCLUSION

This study presents the successful design, development, and implementation of an innovative automated two-post lift system for automotive servicing. Our approach leverages cutting-edge sensor technology, precise motor control, and intelligent software algorithms to revolutionize vehicle lifting with enhanced precision, efficiency, and safety. Leveraging advanced sensor technology, precise motor controls, and intelligent software algorithms, our system streamlines the vehicle lifting process with enhanced precision, efficiency, and safety. This automated alignment process significantly reduces

the need for manual intervention, minimizing errors and improving overall operational efficiency. Functional tests have demonstrated the system's capability to reliably detect and align with vehicle lifting points, while safety evaluations have underscored its stable lifting performance and user-friendly operation.

Overall, our research signifies a significant advancement in automotive servicing technology, offering a practical solution to enhance the precision and safety of vehicle lifting operations. With further refinement and real-world testing, our automated two-post lift system has the potential to revolutionize automotive servicing practices, improving productivity and ensuring optimal service quality in the automotive industry.

## ACKNOWLEDGMENT

The authors would like to thank the president of Lebanese International University HE Abdel Rahim Mourad and the LIU Bekaa campus administration for their continuous encouragement of research activities at the university.

## REFERENCES

- (ALI), A. N. S. I. (. a. A. L. I., 2008. *Standard for automotive lifts–safety requirements for operation, inspection and maintenance (ALOIM)*, New York: ALI Cortland (NY).
- Anon., 2020. Automatic Car Jack and Pressure Monitoring. *International Journal of Engineering Research & Technology (IJERT)*, 9(3), pp. 748–751.
- Burlet-Vienney, D. a. G. B. a. B. K. C., 2021. Analysis of vehicle stability when using two-post above-ground automotive lifts: Distribution of forces in arms. *Safety science*, Volume 134, p. 105042.
- Dietmar, E. a. T. J. a. R. J., 2017. An International Comparison of the Cost or Work-Related Accidents and Illnesses. *An International Comparison of the Cost or Work-Related Accidents and Illnesses*.
- Heinrich, H. W. a. P. D. a. R. N. a. o., 1980. *Industrial Accident Prevention: A Safety Management Approach*. New York.
- Park, S. a. K. M.-J., 2023. Corporate characteristics and occupational injuries by industry. *Safety and health at work*, 14(3), pp. 259–266.
- Pradhan, U. K. a. K. P. a. V. A. a. S. S. S., 2022. Inbuilt Hydraulic Jack System for Light Motor Vehicles and Heavy Motor Vehicles. *IJRASET*, 10(7), pp. 429–433.
- Québec, S. d. l. A. d., 2021. *Nombre de Véhicules en Circulation Selon Le Type D'utilisation, Le Type de Véhicule Et L'âge Du Véhicule, Québec Et Régions Administratives*, Quebec, QC, Canada: Government Report.
- Santamaría, N. N. C. a. F. G. B., 2006. Tendencias temporales de las lesiones mortales (traumáticas) por accidente de trabajo en España (1992–2002). *Gaceta Sanitaria*, Volume 20, pp. 280–286.
- Sathyan, V. K. V. S. H. P. S., 2019. Effective Performance and Influence of Automated Inbuilt Hydraulic Jack in a Four Wheeler Vehicle. *IJETS*, 6(5), pp. 34–37.
- Thirugnanam, A. a. D. P. a. R. L., 2014. Design and Fabrication of Rack and Pinion lift. *Middle-East J. Sci. Res*, 20(6), pp. 744–748.
- Yekinni, A. a. O. K. a. B. A. a. O. O. a. S. O., 2018. Design and Development of Motorised Vehicle Lifter. *yekinni2018design*.

# The Rolling Robot and the Human Brain: Handover of the Driving Task in Automated Vehicles

Peter Rössger<sup>1</sup>, Cristián Acevedo<sup>2</sup>, Miriam Bottesch<sup>2</sup>,  
Samuel Nau<sup>2</sup>, Tobias Stricker<sup>2</sup>, and Frederik Diederichs<sup>3</sup>

<sup>1</sup>Beyond HMI, Hohe Str. 4, 71032 Böblingen, Germany

<sup>2</sup>Studiokurbos GmbH, Königstrasse 32, 70173 Stuttgart, Germany

<sup>3</sup>Fraunhofer IOSB, Fraunhoferstrasse 1, 76131 Karlsruhe, Germany

## ABSTRACT

The automation of driving represents a pivotal innovation in vehicular technology, transitioning from automating secondary and tertiary tasks such as ignition, gearboxes, and rain sensors, to automating the core driving procedures. This redefinition of the driving process fundamentally alters the human-machine interface (HMI) and vehicle interiors. Ensuring safe driving, vehicle usability, and a positive user experience necessitates clear delineation of responsibilities between human and machine drivers. In a laboratory study with 20 participants (55% female, 45% male, ages 20-59) with varying experience in advanced driver assistance systems (ADAS), four handover procedures were evaluated for understandability and user experience. The study comprised two parts: testing the comprehensibility of animated icons and a comparative analysis of four handover procedure designs. Icon sets differed in their representation—one holistic and the other detailed—and varied in their display strategy and location. Data were collected via direct questioning, the Net Promoter Score (NPS), the meCUE user experience questionnaire, design rankings, and open interviews. Results indicated a preference for icons showing details over holistic ones. The most favored handover procedure featured a centrally located, single detailed icon representing the current activity, leading to superior scores in NPS, meCUE, and overall ranking. Interview feedback highlighted preferences for clarity, simplicity, and central icon placement. Younger participants favored animated icons with bubble effects, while older participants preferred simpler designs. These findings underscore the importance of user-centric design in automated driving systems.

**Keywords:** Automated driving, Handover, User experience, Automotive usability

## INTRODUCTION

Automation stands as one of the pivotal megatrends reshaping automobility. Over the years, automation has seamlessly integrated into the fabric of automotive technology, manifesting in innovations such as automated ignition systems, automated gearboxes, electric window lifters, and rain sensors. Each of these advancements has incrementally enhanced the driving experience, contributing to the ongoing evolution of vehicles.

In recent years, the scope of automation has broadened significantly, leading to a transformative redefinition of the relationship between drivers

and their vehicles. Modern automation trends are poised to modify the core interaction between humans and what can be seen as rolling robots: the driving procedure itself. This profound shift bears significant implications for the Human-Machine Interface (HMI), as well as the overall usability and user experience (UX) within vehicles.

Recognizing the critical need to understand these implications, the KARLI project was initiated. Within the framework of this project, we conducted a comprehensive UX study to evaluate the comprehensibility of icons and the acceptance of various interaction procedures. A primary focus of our investigation was the handover process of the driving task from the driver to the automated system. This aspect of automation is particularly crucial, as it directly influences driver trust and the overall safety of automated driving systems.

The study aimed to identify which design of the handover procedure garnered the highest acceptance among users. By meticulously analyzing user feedback and interaction data, we sought to pinpoint the most effective design elements that facilitate a smooth and intuitive transition of control. The insights gained from this study are set to be integrated into the final version of the KARLI HMI, ensuring that the system meets high standards of usability and user satisfaction.

This paper presents the methodology, findings, and implications of our study, contributing valuable knowledge to the ongoing discourse on automotive automation and HMI design. Through this research, we aim to advance the development of user-friendly automated driving systems that enhance both safety and the driving experience.

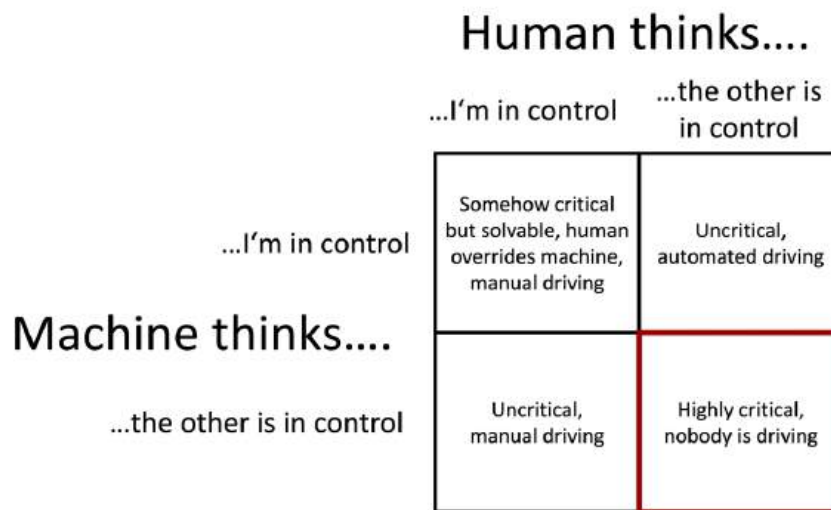
KARLI is a project funded by the German BMWK (Bundesministerium für Wirtschaft und Klimaschutz, Federal Ministry of Economy and Climate Protection). The objective of the KARLI project is to create an adaptive, responsive, and compliance-oriented interaction system for future vehicles (Diederichs et al., 2022, KARLI, 2024). In pursuit of this goal, KARLI is developing AI functions that are relevant to customer needs. These functions aim to identify driver states and formulate interactions tailored to various levels of automation.

In the work package “Level compliant behavior” KARLI aims to assess the driver’s state, behavior, and likely ability to act is derived from the present driving scenario, which includes considerations of the level of automation. Part of this is the development of HMI solutions for safe handovers between vehicle and driver.

## **BACKGROUND**

In the context of partial automation, there is a division of tasks between the human driver and the automated vehicle. This paper adopts the levels of automated driving as defined by SAE J3016 (SAE, 2021). At Level 0 (no driving automation) and Level 5 (full driving automation), the driving task is exclusively performed by either the human driver or the automated system, respectively. However, at Levels 1 through 4, the driving responsibilities are shared between the driver and the automated system.

A fundamental question in partially automated driving is, “Who is doing what at the moment, and how will that change in the future?” For ensuring both safety and an optimal UX, it is crucial to clearly indicate which components of the driving task are managed by the human driver and which are handled by the automated system. A critical safety concern arises when there is a miscommunication or misunderstanding between the vehicle and the driver regarding who is currently in control. This can lead to a scenario where both the driver and the vehicle assume the other is handling the driving task, resulting in a potentially dangerous situation where no one is driving (Figure 1).



**Figure 1:** Possible situations in task sharing (Rössger, 2024).

In the realm of traffic safety, two constructs are delineated for clarity: Situational Awareness (SA) and Mode Awareness (MA). According to Endsley (1995), Situational Awareness is defined as the perception of elements within the environment, the comprehension of their significance, and the projection of their status soon. Effective situational awareness of the current traffic environment, including both its static and dynamic elements, is crucial for the safe transition of vehicle control from an automated driving phase back to the human driver (Schlag, 2016).

Situational Awareness encompasses three levels: perception, comprehension, and projection. Perception involves recognizing critical factors in the environment, such as other vehicles, road signs, and pedestrians. Comprehension refers to understanding the implications of these factors for the driving context, such as recognizing that a pedestrian at the curb might step into the road. Projection entails anticipating future states, like predicting the movement of vehicles based on their current speed and trajectory.

Mode Awareness, on the other hand, pertains to the driver's awareness of the vehicle's current automation mode (Özkan et al., 2021). This includes knowledge and understanding of the system's current state, its behavior, and its expected future states (Andre & Degani, 1997). Mode Awareness is critical

because automated driving systems can operate in multiple modes, each with different capabilities and limitations. For instance, a vehicle might switch between manual driving, partial automation, and full automation modes, each requiring different levels of driver engagement and oversight.

To ensure safe vehicle operation, a driver must not only be aware of the current mode but also understand the implications of each mode for vehicle control and performance. This involves recognizing the system’s boundaries, such as when it can safely navigate on its own and when it requires driver intervention. Furthermore, the driver needs to be aware of transition cues indicating mode changes and be prepared to assume control when necessary.

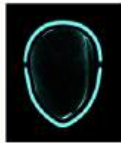







Integrating these constructs into the design and operation of automated driving systems is essential for enhancing traffic safety. By ensuring drivers have a high level of Situational and Mode Awareness, we can mitigate the risks associated with automation and improve the overall safety and efficiency of vehicular traffic. Parts of this chapter are based on an internal paper of the KARLI Project (Pagenkopf et al., n.d.).

Addressing the issue of “who is doing what?”, including SA and MA is essential for the development of reliable and user-friendly partially automated driving systems. We need to clearly indicate who in the human-machine system is performing which tasks. Existing solutions are limited in the quality of information transfer, usability, and UX. We conducted a study to find out, which icons are easiest to understand, and which procedure finds the highest acceptance in the potential user group.

## METHODS

In a study with 20 participants, 55% female, 45% male, aged between 20 and 59, two data collections were performed. The first phase of the study (icon test) collected data on the understandability of icons. Eight icons were tested, one for an Avatar, one for ok, and two sets of icons giving the respective instructions during the handover procedure (Table 1). All icons were animated.

**Table 1.** Matrix of the tested icons.

	Avatar	Feet off Pedals	Hand off steering wheel	Eyes off the road	End of handover
Holistic					
Detail					



In the second part of the study (handover procedures) four different designs of the handover procedure were tested. Videoclips of about 10sec length were shown in a loop. The participants watched them for the entire time we collected data for the respective design. The order of the designs was permuted systemically to reduce effects of order.

For the icon test data was collected applying the following methods:

- Recording of the answers of the participants in written form, was the icon recognized correctly
- If not, what did the participant believed it indicated
- Was it seen as an information or as an advice.

Four different design variants (Designs A to D, Design A is shown as a sample in Figure 2) were developed to support the handover procedure from driver to vehicle.



**Figure 2:** Design A (icons detail, icons parallel to road, road mark lines, animated action).

The characteristics of the respective designs are described in Table 2.

**Table 2.** Design parameters.

	Kind of Icons	Location of Icons	Roadmarks	Animation
Design A	Detail	Parallel to road	Line	Action
Design B	Holistic	Horizontal	Line	Figure and Action
Design C	Holistic	Vertical	Carpet	Figure and Action
Design D	Detail	Single	Carpet	Action and Bubbles

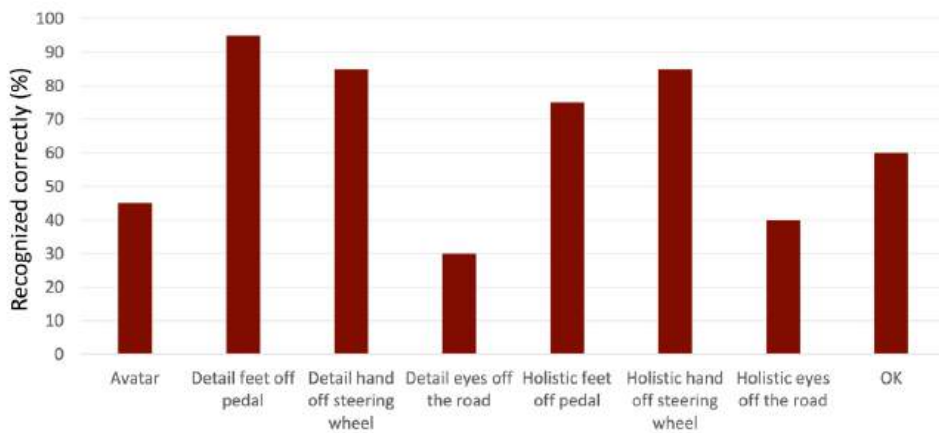
For the handover procedures the following data collection procedures were applied:

- Net Promotor Score (NPS), one question: “How likely is it that you would recommend this company to a friend, colleague, or business partner?” Participants rate their answers on a scale from 0 to 10.
- Expanded NPS, one question: how well did you understand the icon?
- meCUE, a user experience questionnaire
- Ranking of the designs (“which one did you like most, second, third, least?”)
- Interview.

For further insights into the meCUE tool, readers are referred to Minge et al. (2013) and Thüring & Mahlke (2007), while additional details on the Net Promoter Score can be found in references such as Baehre et al. (2022).

**RESULTS**

The icon best understood was the detail icon for feet off pedals. The detail icon hand off steering wheel and holistic hand of steering wheel were understood correctly by over 80% of the participants. Both icons for eyes off the road were understood by 40% of the participants or less (Figure 3).



**Figure 3:** Percentage of icons understood correctly.

The misinterpretations of the icons are listed in Table 3.

**Table 3.** Misinterpretations of the icons.

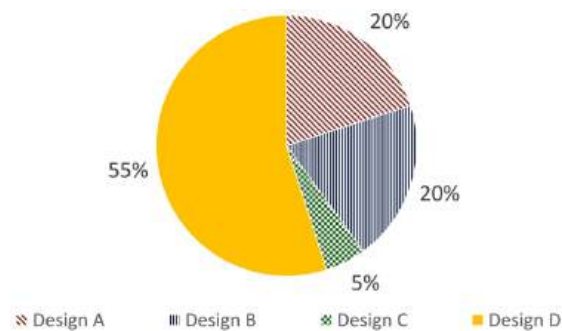
Icon	Missinterpretation
Avatar	Maintenance mode, no meaning, Face ID, facial recognition, unlock vehicle, sleep mode, the car wants to say something
Detail feet off pedal	Take foot off the brake
Detail hand off steering wheel	Steering wheel not working, touch the steering wheel
Detail eyes off road	No idea, look forward, something is turned off, don't fall asleep, poor visibility, camera not working, toggle eye tracking on/off, drowsiness detected, I must stay awake, sensor defect - I must take over
Holistic feet off pedal	Put foot on the pedal
Holistic hand off steering wheel	Something is in front of me, I don't know
Holistic eyes off road	No idea, vision, turn something off, warning that the driver is falling asleep, focus, restricted visibility, look straight ahead, look at the road, scan something
OK	—

The favorite design was design D, 55% of the participants ranked in on place 1, the second place is shred between Designs A and B (20%), last was

design C (5%) (Figure 4). The results of the ranking were supported by the NPS and the meCUE overall rating. In the NPS (scale ranges from -100 to +100) Design D reached -30, Design A -40, Design B -55, and Design C -65. In the meCUE overall rating (scale from -5 to +5) Design D reached 2.1 points, Design A 2, Design B 1.4, and Design C 1.2.

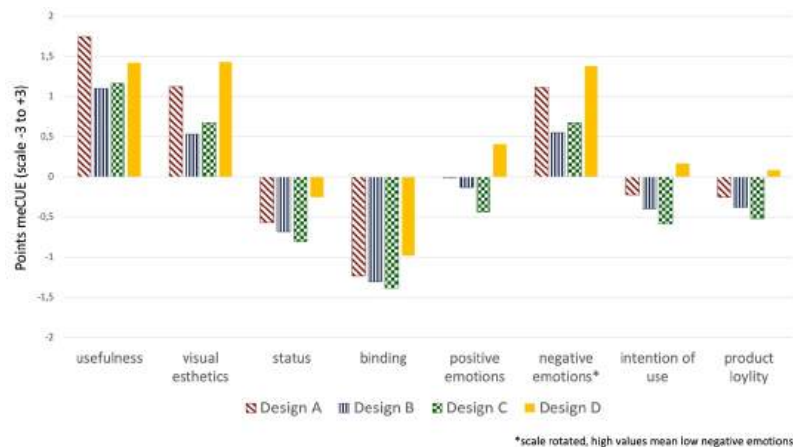
There is a clear age dependency of the favorite designs. 72 % of the participants aged between 18 and 29 rated Design D as the best, 57% of the participants aged 30-39, 50% of the participants between 40 and 49, no participant of 50+ years of age rated it best.

The understandability of the designs, measured with the expanded NPS was positive for all designs. On a scale from -100 to +100 designs A and D reached 25 points, designs B and C 10 each.



**Figure 4:** Results of the design ranking (which design do you think is the best?).

The single dimensions of the meCUE show that all designs are rated positive in usefulness, visual esthetics, and negative emotions (meaning, no negative emotions occurred). Negative scores were found for all designs in status and binding and for all design besides Design D in positive emotions, intention of use, and product loyalty (Figure 5).



**Figure 5:** Results meCUE, single dimensions for every design.

## CONCLUSION

The comprehensibility of the icons varied significantly, with the icons related to eyes-off the road being particularly poorly understood. Detailed icons tended to be more understandable than holistic icons. None of the designs for the handover procedure was rated as particularly good, and the Net Promoter Score was negative for all versions. This indicates that all designs had more detractors than promoters.

All designs were judged to be understandable, with Designs A and D ranking ahead of the other two. In the study section on icon comprehensibility, the detailed icons achieved better scores than the holistic icons. Designs A and D, both of which contain detailed icons, were rated as more comprehensible.

The comprehensibility of the icons ranged from 95% (detailed icon for foot off the pedal) to 30% (detailed icon for eye off the road). The holistic icon for look away was also poorly understood, with a 40% recognition rate. Overall, the detailed icons were better understood than the holistic icons. The icon for the avatar was understood quite poorly, though it is important to note that the context was completely missing. With context (such as language, interaction, and icon animation), the comprehensibility is very likely to be significantly higher. Misinterpretations were greatest for the eyes off the road icons. Often, they were not even associated with the theme of “vision” or were interpreted as the opposite (e.g., “do not fall asleep”). New and better comprehensible designs need to be found for the icons related to “take your eyes of the road.” The detailed icons were generally better understood than the holistic icons. This aligns with many statements from the interviews and the results of the NPS concerning comprehensibility.

The order of the designs was largely consistent across the different evaluation tools. In terms of ranking, ranking points, NPS, and the overall result of the meCUE questionnaire, Design D (Figure 6) was chosen as the best. Over half of the participants (55%) ranked Design D as the best in the ranking. Design A was mostly ranked second across the different evaluation methods. In the ranking, it shared the second place with Design B, while in ranking points, Design B slightly outperformed Design A. Design C was rated the worst by all methods. The age dependency of the ratings was notable; the younger the participants, the more distinctly Design D was favored. In the 50+ age group, Design D was not ranked first. The comprehensibility of Designs A and D was higher than that of the other two. All results were positive, indicating that in all cases, there were more promoters than detractors.

In the subdimensions of the meCUE, the results for Design D were confirmed. It performed the best in all dimensions except usefulness, where it ranked second. Design A won in the usefulness dimension and placed second in all other dimensions. Design B performed the worst in the dimensions of usefulness, visual aesthetics, and negative emotions. Design C scored the lowest in the dimensions of status, attachment, positive emotions, usage intention, and product loyalty. In the dimensions of usefulness, visual aesthetics, and negative emotions, all designs achieved positive values.

However, in the dimensions of status and attachment, the results were negative for all designs. Design D achieved positive values in the dimensions of positive emotions, usage intention, and product loyalty, while all other designs remained in the negative range.



**Figure 6:** The winner, Design D (detailed icons, single presentation of icons, road marks carpet, animation is action and bubbles).

The results of the statistical procedures were supported by the interview statements. Design D was described positively, often with words like "cool." In contrast, Design C was described very negatively by many participants, with some even calling it "the worst design of all!" Among the participants, two groups emerged. One group appreciated the focus of Design D, noting that it only shows what is currently important. A slightly smaller group wanted to see which handover actions had already occurred and which were still to come. Many participants liked the bubbles in Design D and described them positively, while some found this design element negative, considering it too playful or unnecessary.

The icons were generally rated as understandable, except for the two icons for eyes off the road, which were found to be incomprehensible. Design D was unequivocally rated the best by the participants, suggesting that this design should form the basis for further steps. The detailed icons achieved better results than the holistic icons. From a design efficiency perspective, detailed icons are more suitable as they make better use of the available screen space to transfer relevant information. The two designs incorporating detailed icons were rated higher than the other two designs, indicating that detailed icons should be used.

Some participants expressed a preference for indicating which parts of the handover had already occurred and which were upcoming. This could be considered for inclusion, though it should be noted that doing so may compromise the design's focus, clear indication of required actions, and effective positioning of the icons. Design D should be further developed and integrated into the KARLI HMI.

Priority should be given to redesigning the icon for eyes off the road. A suggested redesign is to use two eyes with closing eyelids. Another solution

could be a brain with an arrow, indicating the cognitive relieve of the final step of the handover. Further studies could investigate the systematic influence of individual design parameters. However, since the KARLI project is ending in the foreseeable future, no additional studies will be conducted.

## ACKNOWLEDGMENT

KARLI is a project funded by the German BMWK (Bundesministerium für Wirtschaft und Klimaschutz, Federal Ministry of Economy and Climate Protection).

## REFERENCES

- Andre, A., & Degani, A. (1997). Do you know what mode you're in? An analysis of mode error in everyday things. In M. Mouloua & J. M. Koonce (Hrsg.), *Human-automation interaction: Research & Practice* (S. 19–28). Lawrence Erlbaum.
- Baehre, S., O'Dwyer, M., O'Malley, L., & Lee, N. (2022). The use of Net Promoter Score (NPS) to predict sales growth: insights from an empirical investigation. *Journal of the Academy of Marketing Science*, 50:67–84.
- Diederichs, F., Wannemacher, C., Faller, F., Mikolajewski, M., Martin, M., Voit, M.,... & Piechnik, D. (2022). Artificial intelligence for adaptive, responsive, and level-compliant interaction in the vehicle of the future (KARLI). In: *International Conference on Human-Computer Interaction* (pp. 164–171). Cham: Springer International Publishing.
- Dönmez Özkan, Y., Mirnig, A. G., Meschtscherjakov, A., Demir, C., & Tscheligi, M. (2021). Mode Awareness Interfaces in Automated Vehicles, Robotics, and Aviation: A Literature Review. *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 147–158. <https://doi.org/10.1145/3409118.3475125>
- Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. <https://doi.org/10.1518/001872095779049543>.
- KARLI (2024). *Website of the KARLI Project*. <https://karli-projekt.de/en/start>.
- Folds, Dennis. Gardner, Douglas and Deal, Steve. (2008). Building Up to the Human Systems Integration Demonstration, *INCOSE INSIGHT* Volume 11, No. 2.
- Minge, M., Riedel, L. & Thüring, M. (2013). Modulare Evaluation von Technik. Ent-wicklung und Validierung des meCUE Fragebogens zur Messung der User Experience. In: E. Brandenburg, L. Doria, A. Gross, T. Guntzler & H. Smieszek (Hrsg.): *Grundlagen und Anwendungen der Mensch-Technik-Interaktion*. 10. Berliner Werkstatt Mensch- Maschine-Systeme (S. 28-36). Berlin: Universitätsverlag der TU Berlin.
- Pagenkopf, A. (n.y.). *Bedingungen für angemessenes Modus-/Situationsbewusstsein*. Internal paper KARLI Project).
- Rössger, P. (2024). *Automated Driving*. Automotive User Interface lecture at the FH Aachen.
- SAE (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. According to [https://www.sae.org/standards/content/j3016\\_202104/](https://www.sae.org/standards/content/j3016_202104/).
- Schlag, B. (2016). Automatisiertes Fahren im Straßenverkehr—Offene Fragen aus Sicht der Psychologie. *Zeitschrift für Verkehrssicherheit*, 2(2016), 94–98.
- Thüring, M. & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology*, 2007, 42 (4), 253–264.

# Age-Based Differences in Pedestrians' Feeling of Trust and Safety When Crossing in Front of a Real Communicating Self-Driving Car During Daytime or Nighttime

Aïsha Sahai<sup>1</sup>, Onoriu Puscasu<sup>2</sup>, and Natacha Métayer<sup>1</sup>

<sup>1</sup>VEDECOM Institute, 23 bis allée des Marronniers, 78000 Versailles, France

<sup>2</sup>VALEO Light, 34 rue Saint-André, 93000 Bobigny, France

## ABSTRACT

The introduction of self-driving cars (SDCs) onto public roads will raise challenging issues to ensure traffic fluency. One of these is to guarantee pedestrians feel safe and confident when encountering this new type of vehicle in order to promote pedestrian crossing in front of SDCs. Hence, the aim of the study was to investigate the impact of different types of external Human-Machine Interfaces (eHMIs) indicating the yielding behaviour of a SDC on pedestrians' feeling of safety and trust. Thirty-four participants (19 young adults aged 22–41 and 15 older adults aged 63–80) volunteered to take part in the experiment. Participants were requested to cross in front of a real SDC which gave way to them on a crosswalk. The SDC was equipped with devices that could send different types of eHMI signals. In the *hourglass* condition, two displays located in front of the SDC were showing a luminous hourglass when the SDC yielded (daytime and nighttime tests). In the *safety zone* condition, projectors were sending a cyan light signal onto the ground around the SDC when it yielded (nighttime tests). In the *no eHMI* condition, none of the above-mentioned signals were shown. The participants were not informed in advance about the presence of the eHMI signals or their meaning. After each crossing, they were asked to rate their level of trust (on a Likert scale from 1- no trust at all to 7- totally in trust) and their level of safety (on a Likert scale from 1- not secure at all to 7- totally secure) during their crossing in front of the SDC. Finally, semi-directive interviews were lead in order to gather additional information such as the cues used by the participants in their crossing decision-making, other than the eHMIs signals. Our results showed high levels of self-reported trust and safety overall. Moreover, a significant main effect of the age group indicating a stronger level of trust during the crossing of the older adults as compared to the young adults was found. However, no significant effect of the age group nor of the type of eHMI signal were found on the participants' feeling of safety during their crossing. In addition, no significant effect of the type of eHMI signal was found on the participants' level of trust during their crossing. Yet, the analysis of the semi-directive interviews revealed that the young adults were likely to use more implicit communication cues exhibited by the SDC during the nighttime tests as compared to the daytime tests while the older had a more conservative crossing decision-making strategy during both daytime and nighttime tests. These findings provided understanding elements of pedestrians' crossing experience with regards to communicating SDCs in realistic conditions.

**Keywords:** Pedestrian, Ageing, Self-driving car, External human-machine interface, Crossing decision, Safety, Trust

## INTRODUCTION

Self-driving cars (SDCs) are vehicles equipped with automated driving features capable to monitor the driving environment (e.g., detection of the surrounding pedestrians) and to manage the lateral and longitudinal control of the vehicle (e.g., to stay in the lane or to yield) (Society of Automotive Engineers, 2021). This automation of driving allows the human inside the vehicle to discharge her/himself from the driving activity. Hence, the non-verbal communication cues (e.g., eye contact, gestures, postural changes) traditionally used by the human drivers in order to express their driving intentions to pedestrians (Rasouli et al., 2017; Sucha et al., 2017) may no longer be present. Even more, visual communication with in-vehicle drivers could be merely dismissed by pedestrians when they would face SDCs (Sahaï et al., 2022). This is particularly at risk as this can lead to the misunderstanding of the situation thereby generating a feeling of uncertainty among the pedestrians (Lagström & Lundgren, 2015; Palmeiro et al., 2018) and cause them to make unsafe decision-making.

To overcome this lack of human non-verbal communication, the use of external Human-Machine Interfaces (eHMIs) has been proposed to make vehicles communicate themselves about their intentions such as yielding or turning for example. The eHMIs signals could be of different forms such as text-based or symbolic light signals located on the SDC, to light projections on the road, or to vibrotactile information transmitted by personal wearable devices connected to the vehicle (Bazilinskyy et al., 2019; Cœugnet et al., 2017; De Clercq et al., 2019).

Interestingly, there is some evidence that the presence of eHMI signals could improve the pedestrian crossing experience when facing SDCs. Indeed, the presence of an eHMI signal in the form of light bands placed in front of the SDC and indicating the yielding intention of the SDC was shown to enhance the pedestrians' perceived trust towards the SDC and stimulate their crossing in front of the SDC, especially when the eHMI signal turned on before the SDC started to decelerate (Kaleefathullah et al., 2020). Similarly, both text-based and symbolic light eHMI signals emitted on displays were shown to enhance the pedestrian feeling of safety to cross when they encountered a yielding automated vehicle, particularly when the eHMI signal was turned on before the vehicle started to decelerate or at the beginning of the deceleration (De Clercq et al., 2019). In the same vein, it has been shown that the pedestrians' degraded perception of safety when facing an inattentive driver could be mitigated by the presence of an eHMI signal in the form of a light band (Faas et al., 2021). However, in the authors' study, the eHMI signal was on continuously with the aim to indicate the automated driving mode of the SDC. Yet, although the pedestrians could well understand that they should not wait for communication cues from the driver, the yielding intention of the SDC was not clearly stated by the eHMI signal, which could generate poor situation awareness, while this latter was shown to be positively correlated with the quality of decision-making (Stanners & French, 2005). Taken together, these findings suggested that eHMI signals may foster pedestrian crossing in front of SDCs by removing the uncertainty about the



intention of the SDC and generating a stronger feeling of safety and trust, although there is no clear consensus about the best eHMI signal to emit so far.

Moreover, prior investigations focusing on the pedestrian experience with SDCs has mainly included young adults as participants (e.g., De Clercq et al., 2019; Kaleefathullah et al., 2020; Lagström & Lundgren, 2015; Palmeiro et al., 2018). However, the European Road Safety Observatory indicated that in 2020, older adults over 65 years olds were overrepresented in pedestrian fatalities (European Commission, 2021). This must be related to the older adults' frail perceptual-motor skills coupled with their less efficient crossing strategies (Tournier et al., 2016; Wilmot & Purcell, 2022). Consequently, older adult pedestrians could be considered as especially vulnerable road users and should be given special attention with the forthcoming fleet of new mobility objects such as SDCs.

Finally, it should be noted that pedestrians are at higher risk of collision during nighttime compared to daytime (Uttley & Fotios, 2017). However, to our knowledge, no study to date has examined the pedestrians' subjective crossing experience in front of a real SDC during nighttime.

In this context, we aimed at investigating age-based differences in pedestrians' feeling of trust and safety when crossing on track in front of a real communicating SDC during either daytime or nighttime.

## **METHODS**

### **Participants**

Nineteen young adults (10 women, 9 men) aged between 22 and 41 (mean = 30, S.D.=5.46 years) and 15 older adults (5 women, 10 men) aged between 63 and 80 (mean = 70.13, S.D.=5.24 years) took part in the experiment. Among them, 19 participants completed daytime tests (7 young adults and 12 older adults) while 15 participants completed nighttime tests (12 young adults and 3 older adults). All participants had normal or correct-to-normal vision and gave their informed consent before participating in the experiment.

### **Materials**

The SDC was a vehicle equipped with the Valeo Drive4U autonomous driving system. This system has the ability to keep the vehicle within a lane, to detect surrounding pedestrians (using cameras and lidars) and to give way if necessary. Moreover, the SDC was supplied with left and right front displays that could show a 17x12 cm luminous hourglass coupled with a car and projectors that could send 44 cm cyan light brackets onto the ground around the vehicle (see Figure 1). The eHMI signals which bared the message that the vehicle was waiting were designed taking into account the ISO TR 23049 standard recommendation to avoid communicating guidance.



**Figure 1:** Types of eHMI signals provided by the self-driving car. Left: a luminous hourglass was shown on front displays. Right: a safety zone was shown onto the ground.

## Procedure

One week before the day of the experiment, the participants were asked to fill out a sociodemographic survey. During the experiment, the participants were asked to walk on the sidewalk for 10 meters before arriving at a crosswalk. They were told that at this crosswalk, they would have to reach the opposite sidewalk by crossing the street and that they would encounter a SDC during their crossing. A stop sign located just in front of the crosswalk and a stop line drawn just after the crosswalk were present on the SDC's path. The experiment was conducted in a fenced curved track so that the behaviour of the SDC's was not visible before the participants arrived at the pedestrian crossing. Each participant carried out the experiment either during the daytime or at nighttime.

In the *hourglass* condition, after giving way and coming almost to a stop, the SDC was showing an hourglass coupled with a car as a yielding eHMI signal. In the *safety zone* condition, after giving way and coming to a stop, the SDC was showing brackets around the SDC as a yielding eHMI signal. In the *no eHMI* condition, no eHMI signal was sent by the SDC. The hourglass and the no eHMI conditions were used for both daytime and nighttime tests. Conversely, the safety zone condition was used only for the nighttime tests. During all nighttime tests, the track was lit by street lamps. Each daytime participant completed eight crossings (2 eHMI conditions x 4 repetitions) while each nighttime participant completed nine crossings (3 eHMI conditions x 3 repetitions). The order of the trials was counterbalanced between participants.

After each crossing in front of the SDC, the participants were asked to rate their feelings of trust (on a Likert scale from 1- no trust at all at all to 7- totally in trust) and safety during their crossing (on a Likert scale from 1- not secure at all to 7- totally secure).

Finally, semi-directive interviews were led at the end of the experiment with each participant in order to gather additional information such as the detection of the eHMI signals, the understanding of the eHMI signals if seen, and the cues used for their crossing decision-making, other than the eHMIs signals.

## RESULTS

### Feeling of Trust and Safety

The data of the participants who reported not having seen the signals were excluded from the analyses. Hence, the potential interactions effect between the participants' age group and the eHMI signal type were not analysed due to low remaining sample sizes per condition (see Table 1).

**Table 1.** Detection rate of the eHMI signals by age group and by moment of the day.

	Young adults		Older adults	
	Day	Night	Day	Night
Hourglass	6/7	7/12	6/12	1/3
Safety zone		11/12		2/3

In order to investigate the effect of the participants' age group (young adults, older adults) on the perceived levels of trust and safety during their crossings, the Wilcoxon Mann-Whitney test was computed on the participants' mean trust scores and mean safety scores with the age group as a factor. A significant main effect of the participants' age group indicating a stronger level of trust for the older adults was found ( $\text{mean}_{\text{young}}=5.41$  (STD = 1.08),  $\text{mean}_{\text{older}}=5.99$  (STD = 1.14),  $W=258.5$ ,  $p=.049$ , see Figure 2). However, no significant main effect of the participants' age group was found on the participants' mean safety scores ( $\text{mean}_{\text{young}}=5.50$  (STD = 1.15),  $\text{mean}_{\text{older}}=5.54$  (STD = 1.12),  $W=374.5$ ,  $p=.93$ ). In other words, the older adults' perceived level of trust when crossing in front of the SDC was stronger compared to those of the young adults, while both age groups reported an equivalent level of safety when crossing in front of the SDC.

Moreover, in order to investigate the effect of the moment of the day (daytime, nighttime) on the participants' levels of trust and safety during their crossings, the Wilcoxon Mann-Whitney test was computed on the participants' mean trust scores and mean safety scores with the moment of the day as factor. No significant main effect of the moment of the day was found on the participants' mean trust scores ( $\text{mean}_{\text{daytime}}=5.66$  (STD = 1.16),  $\text{mean}_{\text{nighttime}}=5.54$  (STD = 1.07),  $W=460$ ,  $p=.64$ ) or on the participants' mean safety scores ( $\text{mean}_{\text{daytime}}=5.49$  (STD = 1.17),  $\text{mean}_{\text{nighttime}}=5.53$  (STD = 1.12),  $W=419.5$ ,  $p=.89$ ). In other words, the participants' perceived levels of trust and safety when crossing in front of the SDC did not differ depending on the moment of the day.

Besides, for the daytime participants subset, the effect of the eHMI signal type (off, hourglass) on the perceived levels of trust and safety during the crossings were assessed by computing the Wilcoxon signed rank test on the participants' mean trust scores and mean safety scores with the eHMI signal type as factor. No significant main effect of the eHMI

signal type was found on the participants' mean trust scores ( $\text{mean}_{\text{off}}=5.63$  (STD = 1.14),  $\text{mean}_{\text{hourglass}}=5.70$  (STD = 1.22),  $W=26$ ,  $p=.92$ ) or on the participants' mean safety scores ( $\text{mean}_{\text{off}}=5.38$  (STD = 1.31),  $\text{mean}_{\text{hourglass}}=5.59$  (STD = 1.04),  $W=35.5$ ,  $p=.14$ ). In other words, the participants' perceived levels of trust and safety when crossing in front of the SDC during the day did not differ depending on the presence of the eHMI signal sent by the SDC.

In addition, for the nighttime participants subset, the effect of the eHMI signal type (off, hourglass, safety zone) on the perceived levels of trust and safety during the crossings were assessed by computing the Kruskal-Wallis test on the participants' mean trust scores and mean safety scores with the eHMI signal type as factor. No significant main effect of the eHMI signal type was found on the participants' mean trust scores ( $\text{mean}_{\text{off}}=5.05$  (STD = 1.22),  $\text{mean}_{\text{hourglass}}=5.14$  (STD = 1.05),  $\text{mean}_{\text{zone}}=5.73$  (STD = 1.21),  $H_{df=2}=1.72$ ,  $p=.42$ ) or on the participants' mean safety scores ( $\text{mean}_{\text{off}}=5.11$  (STD = 1.20),  $\text{mean}_{\text{hourglass}}=5.14$  (STD = 1.25),  $\text{mean}_{\text{zone}}=5.62$  (STD = 1.42),  $H_{df=2}=1.03$ ,  $p=.60$ ). In other words, the participants' perceived levels of trust and safety when crossing in front of the SDC during the night did not differ depending on the presence or the type of eHMI signal sent by the SDC.

### Understanding of the eHMI Signals

When asked about the meaning of the eHMI signals during the semi-directive interviews, the yielding intention of the SDC in the *hourglass* condition was mentioned by 23% (3 of 13) of the daytime participants' answers and adapted to 63% (5 of 8) of the nighttime participants' answers (see Figure 2). Moreover, the yielding intention of the SDC in the *safety zone* condition was mentioned by adapted to 17% (2 of 12) of the nighttime participants' answers (see Figure 3). It should be noted that a single participants' answer for the understanding of the eHMI signal could be counted in more than one response categories. Yet, for all eHMI signals the majority of participants had at least a close interpretation of the meaning.



**Figure 2:** Understanding of the eHMI signal in the *hourglass* condition for the daytime participants subset (left) and the nighttime participants subset (right). Dark green denotes precise understanding. Lighter green shades denote a close interpretation by the participant.

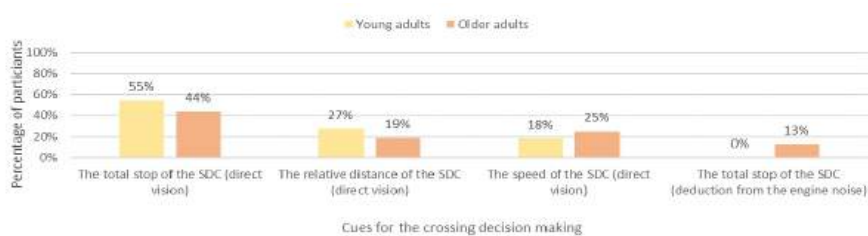


**Figure 3:** Understanding of the eHMI signal in the *safety zone* condition.

### Cues for the Crossing Decision-Making

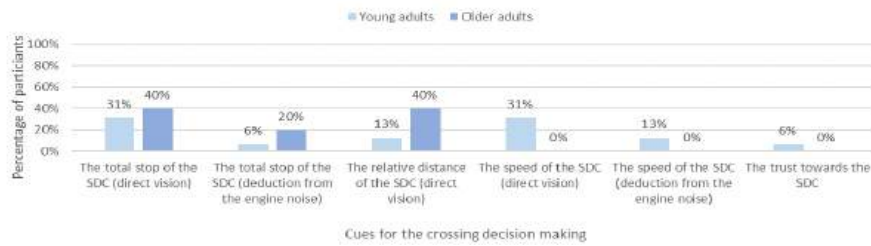
In the *hourglass* condition, among the 20 participants who reported having seen the eHMI signal, 4 of 12 daytime participants (2 young adults and 2 older adults) and 5 of 8 nighttime participants (4 young adults and 1 older adult) reported having used this signal as a cue in their decision to cross. In the *safety zone* condition, among the 12 participants who reported having seen the eHMI signal, 3 young adults and 1 older adult reported having used this signal in their decision to cross.

Furthermore, other than the eHMIs signals, within the daytime participants subset, the cues used for the crossing decision-making were almost the same for the two age groups and essentially vision-based. The cue the most used, other than the eHMI signal, was the visually checked total stop of the SDC both for the young adults (at 55%) and the older adults (at 44%, see Figure 4).



**Figure 4:** Percentage of participants' cues used for the crossing decision-making during the daytime.

By contrast, within the nighttime participants subset, the cues used for the crossing decision-making were both vision-based and audition-based. Moreover, the young adults used more cues than the older adults for their crossing decision-making. The cue the most used, other than the eHMI signal, was the total stop of the SDC visually checked both for the young adults (at 31%) and the older adults (at 40%). Additionally, and in the same proportions, the young adults mostly used the visually checked SDC speed for their crossing decision-making (at 31%) while the older adults mainly used the visually checked relative distance of the SDC (at 40%, see Figure 5).



**Figure 5:** Percentage of participants' cues used for the crossing decision-making during the nighttime.

## DISCUSSION AND CONCLUSION

The aim of this study was to investigate age-based differences in the pedestrians' feeling of trust and safety when crossing in front of a real communicating SDC during either daytime or nighttime. Both young and older adults were asked to cross in front of a yielding SDC that could emit different types of eHMI signals, and to rate their feeling of trust and safety during the crossings.

Our results showed overall strong levels of trust and safety when the participants crossed in front of the SDC. Yet, we found that the older adults' level of trust was stronger compared to those of the young adults. Although it was not assessed, it could be possible that the older adults in the experiment were less aware of the potential failures and risks of emerging automotive technologies as compared to the younger adults, which would explain their greater feeling of trust during their crossings in front of the SDC.

Moreover, our results revealed intriguingly that the participants levels of trust and of safety did not differ daytime compared to nighttime, despite the deterioration of visibility in this latter case. This could be due to the presence of the crosswalk on the ground that led the participants to feel safe and confident during all circumstances given that any approaching vehicle is obliged to stop when a pedestrian is entering a crosswalk. Hence, the approaching SDC could be expected to follow the traditional right-of-way (Meeder et al., 2017) regardless of the moment of the day.

Regarding the impact of the different types of eHMI signals, our findings suggested no impact of the presence or the type of eHMI signal sent by the SDC on the participants' levels of trust and safety. This result was consistent with virtual reality work showing no impact of eHMI signals such as a smile, a traffic light or a handwave presented on front displays on pedestrians' crossing decision confidence (Holländer et al., 2019). By contrast, when dealing with safety, we did not succeed at showing a positive impact of any type of eHMI signal in real-world conditions as opposed with virtual reality work (De Clercq et al., 2019). A possible explanation would be that the onset of the eHMI signals in the current study was too late in order to significantly influence trust and safety feelings. Indeed, in each condition, the eHMI signals were emitted when the yielding SDC was almost at a stop. Thus, vehicle kinematics could have played a more important role than the explicit communication of the SDC in establishing trust and a safety feeling.

Likewise, it has been shown that eHMI signals had little influence on the participants' feeling of safety to cross when the eHMI signal was emitted after the beginning of the vehicle's deceleration (De Clercq et al., 2019).

Finally, the semi-directive interviews analysis confirmed that implicit communication cues embodied in the SDC's behaviour still played a role in the participants crossing decision-making. More in detail, we found that the implicit communication cues used in the participants' crossing decision-making were mostly vision-based during the daytime (e.g., the visually assessed total stop and relative distance of the SDC). However, there was a mix of both vision-based and audition-based implicit cues during the nighttime (e.g., the total stop of the SDC assessed with the help of the engine noise), likely to compensate for the loss of visibility in darkness. This was in line with the existing scientific literature indicating that implicit cues communicated by the vehicle's kinematics could be enough to grasp the intention of automated vehicles (Palmeiro et al., 2018; Dey et al., 2021; Lee et al., 2021). However, our findings went a step further indicating that implicit communication remained an important cue with ageing on the one hand, and both during the daytime and nighttime on the other hand. Furthermore, our data showed that the young adults were more likely to use a wider range of implicit communicative cues from the SDC during the nighttime as compared to the daytime. By contrast, the older adults seemed to adopt a more conservative crossing decision-making strategy by using almost the same implicit communicative cues during the daytime and during the nighttime.

To conclude, this study put forward the crossing experience of young and older pedestrians during their encounter with a real SDC. It highlighted the importance of making understandable eHMI signals in order to come into play in pedestrians' crossing decision-making, especially when the SDCs will be present in complex road traffic and will generate traffic uncertainties. Indeed, in the current study, the eHMI signal was better understood and thereby had a greater rate of use in the crossing decision making in the *hourglass* condition than the *safety zone* condition. In addition to that, the day/night environmental conditions seemed to have an impact on the precise understanding rate of the eHMI signal, as a higher proportion of the nighttime participants had a precise understanding of the eHMI signal in the *hourglass* condition. This shed light the importance of context in understanding an eHMI signal and suggests that pedestrian needs might be different during nighttime as compared to daytime. Moreover, it should be emphasized that the pedestrian crossing experience is closely dependent on individual factors such as social norms and culture for instance (Rasouli & Tsotsos, 2019). Nevertheless, pedestrians might not be reluctant to cross in front of SDCs as high levels of trust and safety could be experienced. Importantly, this study has focused on the feeling of safety and trust experienced by pedestrians, but an objective video-based behaviour analysis is in progress as well as the investigation of other key parameters of the user experience such as eHMI acceptance. Moreover, an additional nighttime study can be done in order to increase the sample size of participants. Yet, future investigations must involve more complex crossing scenarios to fit with

the real-world traffic conditions (e.g., plurality and mix of road user types present on the track, open-road).

## ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101006664. The sole responsibility of this publication lies with the authors. Neither the European Commission nor CINEA – in its capacity of Granting Authority – can be made responsible for any use that may be made of the information this document contains.

The authors would like to thank Théa Cosnier (VEDECOM Institute) for her active assistance with the data collection, and colleagues from Valeo Kronach site (Valeo Brain division) for their assistance with the self-driving car.

## REFERENCES

- Bazilinsky, P., Dodou, D., & De Winter, J. (2019). Survey on eHMI concepts: The effect of text, color, and perspective. *Transportation research part F: Traffic psychology and behaviour*, 67, 175–194.
- Cœugnet, S., Dommès, A., Panéels, S., Chevalier, A., Vienne, F., Dang, N. T., & Anastassova, M. (2017). A vibrotactile wristband to help older pedestrians make safer street-crossing decisions. *Accident Analysis & Prevention*, 109, 1–9.
- De Clercq, K., Dietrich, A., Núñez Velasco, J. P., De Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human factors*, 61(8), 1353–1370.
- Dey, D., Matviienko, A., Berger, M., Pfleging, B., Martens, M., & Terken, J. (2021). Communicating the intention of an automated vehicle to pedestrians: The contributions of eHMI and vehicle behavior. *IT-Information Technology*, 63(2), 123–141.
- European Commission (2021) Facts and Figures Pedestrians. European Road Safety Observatory. Brussels, European Commission, Directorate General for Transport.
- Faas, S. M., Stange, V., & Baumann, M. (2021). Self-driving vehicles and pedestrian interaction: Does an external human-machine interface mitigate the threat of a tinted windshield or a distracted driver?. *International Journal of Human-Computer Interaction*, 37(14), 1364–1374.
- Holländer, K., Colley, A., Mai, C., Häkkinen, J., Alt, F., & Pfleging, B. (2019, October). Investigating the influence of external car displays on pedestrians' crossing behavior in virtual reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–11).
- Kalefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & Winter, J. D. (2020). External human-machine interfaces can be misleading: An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human factors*, 64(6), 1070–1085.
- Lagström, T., & Lundgren, V. M. (2015). AVIP-Autonomous vehicles' interaction with pedestrians: An investigation of pedestrian-driver communication. Chalmers University of Technology.



- Lee, Y. M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C.,... & Merat, N. (2021). Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cognition, Technology & Work*, 23, 367–380.
- Meeder, M., Bosina, E., & Weidmann, U. (2017, May). Autonomous vehicles: Pedestrian heaven or pedestrian hell. In 17th Swiss Transport Research Conference (pp. 17–19).
- Palmeiro, A. R., van der Kint, S., Vissers, L., Farah, H., de Winter, J. C., & Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A wizard of Oz experiment. *Transportation research part F: traffic psychology and behaviour*, 58, 1005–1020.
- Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2017, June). Agreeing to cross: How drivers and pedestrians communicate. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 264–269). IEEE.
- Rasouli, A., & Tsotsos, J. K. (2019). Autonomous vehicles that interact with pedestrians: A survey of theory and practice. *IEEE transactions on intelligent transportation systems*, 21(3), 900–918.
- Sahaï, A., Labeye, E., Caroux, L., & Lemerrier, C. (2022). Crossing the street in front of an autonomous vehicle: An investigation of eye contact between drivengers and vulnerable road users. *Frontiers in psychology*, 13, 981666.
- Society of Automotive Engineers (2021). Taxonomy and definitions for terms Related to driving automation systems for on-road motor vehicles (J3016\_202104 (pp. 1–41). Society of Automotive Engineers.
- Stanners, M., & French, H. T. (2005). An empirical study of the relationship between situation awareness and decision making.
- Sucha, M., Dostal, D., & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention*, 102, 41–50.
- Tournier, I., Dommès, A., & Cavallo, V. (2016). Review of safety and mobility issues among older pedestrians. *Accident Analysis & Prevention*, 91, 24–35.
- Uttley, J., & Fotios, S. (2017). The effect of ambient light condition on road traffic collisions involving pedestrians on pedestrian crossings. *Accident Analysis & Prevention*, 108, 189–200.
- Wilmot, K., & Purcell, C. (2022). Why are older adults more at risk as pedestrians? A systematic review. *Human factors*, 64(8), 1269–1291.

# Exploring the Risks of Password Reuse Across Websites of Different Importance

Anurag Mathews<sup>1</sup> and S M Taiabul Haque<sup>2</sup>

<sup>1</sup>Computer Science, Bergen County Technical High School, Teterboro, NJ 07608, USA

<sup>2</sup>Department of CSE, BRAC University, Merul Badda, Dhaka 1212, Bangladesh

## ABSTRACT

This study attempts to simulate the different ways through which a malicious hacker may attempt to gain unauthorized access to user accounts by leveraging the similarities between multiple linked passwords of the same user. The issue of managing multiple password-protected accounts exemplifies the usability/security trade-off in cybersecurity. Users often reuse the same password, with little or no modifications, across websites of different importance, compromising the security of the high-value accounts. By combining syntactic similarity, dictionary attack, service-related keywords, and semantic similarity on a set of 62,213 linked passwords available from the leaked databases on the internet, 82.3% of the high-value passwords were cracked with an average of 1.82 seconds spent on each attempted password. Similarly, the syntactic method alone achieved an accuracy of 73.6% at 0.82 seconds spent per password attempted. We further connect our findings to the broader issues in cybersecurity and offer a few suggestions to protect the high-value accounts of the users.

**Keywords:** Usability, Security, Authentication, Password, Hierarchy

## INTRODUCTION

Despite being the most popular method of user authentication on the internet, textual passwords have several usability issues that become more pronounced due to the existence of multiple password-protected accounts. A recent study reveals that an average user may need around 100 passwords to manage across various accounts (Sicora, 2022). According to Miller's law, however, a typical user can retain about seven items in working memory (Miller, 1956). Consequently, users reuse the same password, with little or no modifications, for multiple accounts (Haque, Wright and Scielzo, 2014).

The existing literature on password security and usability has examined password reuse from different viewpoints, but very few studies have adopted the nuanced approach of formulating a user password hierarchy to investigate this issue. This hierarchy was first proposed and then extended by Haque, Wright and Scielzo (Haque, Wright and Scielzo, 2013; Haque, Wright and Scielzo, 2014), in which they categorized the password-protected accounts of a user into two parts: higher-level and lower-level. While the higher-level passwords are used to protect financial and email/social media accounts, the lower-level passwords are mostly used in the contexts where

the users have little incentives to protect the security of the accounts such as the news/weather websites and lesser known websites (Haque, Wright and Scielzo, 2013; Haque, Wright and Scielzo, 2014). To further examine how an attacker can exploit these lower-level passwords to compromise the higher-level accounts, they conducted a comprehensive user study to investigate the similarities between these two types of passwords and the results demonstrate that the degree of similarity is quite high as users often reuse the lower-level passwords, with little or no modifications, to create their higher-level passwords.

One major limitation of the study by Haque, Wright and Scielzo (2014) is that they relied on a lab study to understand the relationship between the passwords of different categories of a user. The participants were asked to construct new, artificial passwords (different from their actual passwords) for both the higher and lower-level websites in a lab setting and answer some survey questions regarding their password reuse habit. Consequently, the ecological validity (Fahl et al., 2013) was low in their study. Furthermore, when cracking the higher-level passwords by using the lower-level ones, they only considered the syntactic similarities between these two groups.

Our current work is built upon this proposed hierarchy, and we measure the degree of similarity between the lower and higher-level passwords by collecting actual passwords from the leaked databases on the internet. Furthermore, in addition to syntactic similarity, we focus on dictionary attacks, service-related keywords, and semantic similarity. Our results show that by combining syntactic similarity, dictionary attack, service-related keywords, and semantic similarity, 82.3% of the higher-level passwords were cracked. When only a dictionary attack was used on the same set of linked passwords, 30.1% of the higher-level passwords were cracked whereas the syntactic similarity method alone could crack 73.6% of the passwords. The additional percentage of passwords cracked by using service-related keywords and semantic similarity was nominal.

Our work joins the growing body of literature in cybersecurity that addresses the usability-security trade-off (Di Nocera and Tempestinim, 2022; Di Nocera, Tempestini and Orsini, 2023; Adams and Sasse, 1999; Whitten and Tygar, 1999). We show how a malicious hacker can exploit the usability issues related to managing multiple password-protected accounts to gain access to the more important accounts of a user through publicly available information and moderate computing resources. Our study calls for a deeper understanding of the behavioral approaches of the users to protect their high-value accounts on the e-commerce and social media websites.

Our main contributions include: 1) We leverage an established user password hierarchy to determine the degree of similarity between the higher and lower-level passwords by using actual leaked passwords on the internet. 2) We examine the effectiveness of syntactic similarity, semantic similarity, dictionary attacks, and service-related keywords. 3) We connect our findings to the usability/security trade-off in cybersecurity in regard to managing multiple password-protected accounts.

## METHODOLOGY

Similar to the work of Haque, Wright and Scielzo (2014), the major goal of this study is to obtain an understanding of the risks of users reusing similar passwords between lower and higher importance websites. However, the passwords obtained for this study is from the leaked data breaches and not from surveyed participants. We aim to simulate the hypothetical scenario in which a hacker has a lower-level password and attempts to use it to guess the higher-level ones. In this regard, the attacker would use dictionary attack, syntactic similarity, semantic similarity, and service-related keywords. This, in turn, would help us to better understand if there are increased risks for the internet users who use similar passwords for their lower and higher-level websites. In summary, there are two novel characteristics in our developed approach. First, we examine the effectiveness of combining multiple approaches: dictionary attack, syntactic similarity, semantic similarity, and service-related keywords. Second, we test our approach on real-world passwords that are publicly available due to several data breaches, which removes any discrepancies between passwords created during artificial lab studies (Haque, Wright and Scielzo, 2014) and passwords constructed by the general population for their actual accounts.

### Dataset

The dataset used for this study must contain linked passwords, which we define as two or more passwords from the same internet user across multiple websites. While password databases are available such as the RockYou dataset or the AuthInfo dataset, such datasets either have no userIDs or contain userIDs that are sufficiently anonymized (Güven, Boyaci and Aydin, 2022). While this was done to protect internet users and allow researchers to expand on password practice research, these datasets are not effective for simulating this hypothetical situation as it is not possible to know whether two or more passwords belong to the same internet user. As a result, we downloaded several email+password combination lists from intelx.io. Operating under the ground truth that email addresses are unique to an internet user, we recorded the instances where there were duplicate email addresses across data breaches, suggesting that the associated passwords would belong to the same user.

### Data Collection and Procedural Analysis

We downloaded numerous sections of data breaches from intelx.io (as their database splits data leaks into smaller files), performed a check to ensure that the email address provided is in a valid format, and proceeded to search for duplicate email addresses among all the files. In order to search quickly among the growing list of user credentials, we adopted a dictionary approach that involved creating a Python dictionary (unrelated to the dictionary attack). Using the dictionary data structure allowed us to quickly acquire the values for each key in the dictionary. This approach begins by looping through a list of user entries including an email address, password, and the source website. As the list is being looped through, every

email address would be checked to find out if it has been seen in previous iterations. If not, the email address would be marked as seen, while storing the passwords/sources, and the iterations would continue. If the email address has been seen in previous iterations, then the password and the source of the current entry iteration would be stored with the original email address entry. The entries were later organized into financial, identity, and sketchy/content sites based on the purpose of the site (Haque, Wright and Scielzo, 2014). For example, bitcoin and other cryptocurrency accounts are typically used for financial transactions. Websites like FlashFlashRevolution, GamerzPlanet, and Gamigo have the intended purpose of being used for playing games, but may also appear reasonably sketchy to the individuals using the sites. MySpace was the only site used in the identity category as it was the only available social media platform with both plaintext passwords and corresponding email addresses. Table 1 shows the list of the websites along with their categorization. On a set of about 13.1 million email+password entries from various sites, this method was able to find duplicate email addresses in a total of 75.4 seconds. This time consisted of 21.1 seconds for storing the combined 13.1 million entries in the RAM for processing, 47.7 seconds for classifying entries into linked passwords, and 6.6 seconds to write the linked passwords to a file. This strategy demonstrates that it is a relatively simple task for malicious hackers to find the linked passwords from publicly available account data breaches.

**Table 1.** Website classifications.

Sketchy/Content	Identity	Financial
Aipai.com	MySpace	Bitcoin Accounts
FlashFlashRevolution.com		Generic Cryptocurrency Accounts
FrostLand.pro		
Game-Tuts.com		
GamerzPlanet.net		
Gamevn.com		
Gamigo.com		

## Password Cracking Methods

Now we describe each step of our password cracking method.

### Tokenizing Given Passwords

At first, we used a Python script to tokenize the lower-level password. The given password was broken down into individual parts such as words, numbers, and special characters for syntactic/semantic similarity exploitation.

### Recognizing Words in Passwords and Syntactical Similarity Exploitation

In order to use syntactic similarity, the pre-built WordNinja module was used, which can take a string (the given password) and produce a list

of all of the words within the string. The remaining entities within the password (extra letters, numbers, digits) would also be appended to the end of the list of words. The WordNinja module uses NLP (Natural Language Processing) techniques to extract the words based on what the most likely words are within the string. The words can then be rearranged, removed, and added. Additionally, extra characters such as digits or special characters could be added, the password could try to be reversed, and different letters could be capitalized or replaced with similar characters (for example, password → p@ssw0rd).

### **Semantic Similarity**

In order to leverage semantic similarity, we used a word2vec model trained on GoogleNews-vectors-negative300.bin.gz. A word2vec model uses NLP techniques to recognize the relationships between a large set of words from the training set with regards to how similar they are. We extracted ten of the most similar words for each lower-level password. The Google training set does a fairly good job in returning the most similar words while simultaneously not being encumbered with data and thereby compromising efficiency. By generating the words that are similar to the ones seen in the password, our password cracker can recombine these new words with each other, words from the previous passwords, and even return to strategies in the syntactic similarities as mentioned above.

### **Service-Related Keywords**

Service-related keywords were created semi automatically using prompts on ChatGPT. Such prompts included “please generate a Python list of 50 keywords related to MySpace” and “please ensure that each keyword is a single word”. These keywords could then also be combined/fragmented with words in a user’s given password. Some keyword examples are: MySpacekeywords = [“MySpace”, “myspace”, “Myspace”, “music”], GamesKeywords = [“games”, “minecraft”, “dantdm”, “creeper”].

### **Dictionary Attack**

For the dictionary attack, the password cracker used about 14 million unique passwords from the famous RockYou leaked dataset from December, 2009. These commonly used passwords were tried as potential guesses to match the users’ password. Using a dictionary of leaked passwords differs from the list of words and phrases used in the dictionary attack of Haque, Wright and Scielzo (2014). The dictionary attack, syntactic/semantic similarities, and service-related keywords generate guesses that vary depending on the lower-level password and the website of the higher-level password.

## **RESULTS**

### **Password Breakdown**

A total of 13,174,417 email and passwords combinations were downloaded from publicly available account data breach sources. These sources

include generic video game sites (sketchy/content), MySpace (identity), and bitcoin/generic cryptocurrency sites (financial). We acknowledge that cybercriminals post false password data leaks to harm companies' reputations or make password leak checkers less efficient, so it is important to perform some data cleansing before any testing takes place (Maschler, Niephaus and Risch, 2017). In this study, the downloaded entries with invalid email structures were removed as they could easily be proven as fake accounts. The remaining downloaded entries that are composed of only ASCII characters made up the clean entries. The number of entries downloaded versus the remaining clean entries used can be seen in Table 2. A total of 13,093,960 entries remained after the invalid entries were removed as shown by Table 2. More information regarding the breakdown of password lengths and composition is included in Appendix as this information can allow this study to be more easily compared to other research on password cracking but is not explicitly relevant to the goal of the study.

**Table 2.** Downloaded vs clean entries.

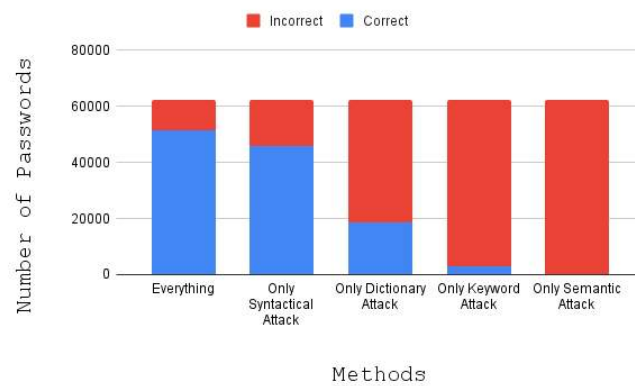
Sources	Entries Downloaded	Clean Entries
Sketchy/Content	6,227,357	6,211,512
Identity/Financial	6,947,060	6,882,448

### Password Cracking Findings

Out of the 13,093,960 clean entries, there were 62,413 linked passwords. Each instance of linked passwords was derived from a unique user as suggested through a unique email address. Given one lower-level password per higher-level password for a unique user, 51,349 passwords out of the 62,413 higher-level passwords were cracked using the four aforementioned methods (dictionary attack, syntactic/semantic similarities, and service-related keywords), giving a password cracking rate of 82.3%. The average time spent attempting to crack a password was 1.82 seconds using an Intel(R) Core(TM) i5-1035G1 CPU, 16 GB RAM, an integrated graphics card, and Windows 11, 64-bit operating system.

Additionally, we ran each of the four methods separately on the same set of 62,413 hidden passwords to understand their individual impact. Overall, 18,755 passwords were cracked using only the dictionary method, giving a password cracking rate of 30.1% at an average of 0.19 seconds per password. The syntactic method was able to crack 45,938 passwords, giving an accuracy of 73.6% at 0.82 seconds spent per password attempted. The semantic method cracked only 27 passwords, giving an accuracy of 0.04% at 0.31 seconds per password. Finally, the service-related keywords method cracked 3008 passwords, giving an accuracy of 4.8% at 0.63 seconds per password. The differences in accuracy between the methods have been demonstrated in Figure 1.

From the above results, we attempt to re-answer the following question from Haque, Wright and Scielzo (2014) using real-world leaked passwords as opposed to surveyed participants (Haque, Wright and Scielzo, 2014).



**Figure 1:** Password cracking comparison.

Research Question: “What percentage of higher-level passwords could be cracked by an attacker by compromising one lower-level password of each participant?” Given a single lower-level password, 82.3% of the higher-level passwords were cracked compared to 26.8% in (Haque, Wright and Scielzo, 2014). It should be noted that there are passwords that could have been cracked by any of the several methods. When visually examining the passwords, it was clear that many linked passwords appeared exactly identical in spelling. When the passwords were not exactly the same, one password often had the first letter capitalized or a single digit was appended to the password. As a result, the password cracker could effectively exploit syntactic similarities between the passwords. However, this study’s unique approach of using WordNinja to extract the words in the passwords and incorporating a word2vec model did not prove effective.

## DISCUSSION

We highlight the limitations of our work before discussing the potential implications.

### Limitations

We followed the standard research procedures, which would simply be ignored by a malicious hacker. The malicious hackers would be able to take advantage of several resources that are restricted due to ethical reasoning such as accessing the dark web. The range and variety of data breaches on the dark web are perhaps significantly greater than those used in this study.

### Implications

Despite these limitations, we demonstrate that by using moderate computational resources and publicly available information, it is possible to crack a very high percentage of higher-level passwords of the users. It has been suggested that the hackers only need to be able to crack roughly 1% to



2% of the passwords to get a return on investment when it comes to accessing user accounts (Acin, 2019), which this study surpasses greatly.

The dictionary attack and syntactic similarities methods proved highly effective at cracking a considerable percentage of passwords. Implementing only a simple dictionary attack that took passwords from a list of 14 million passwords from the RockYou data breach, 30.1% accuracy was obtained. Additionally, manipulating passwords syntactically allowed for an impressive 73.6% accuracy. The service-related keywords also proved effective at 4.8% accuracy, which is above the threshold of the 1% to 2% of accounts needed for a return on investment. The semantic similarity approach proved to be ineffective, however, and its inclusion ultimately led to wasted computational time and effort.

In comparison to the prior works on password cracking, we achieved better results in most of the cases. Simulations have shown that an attack can compromise more than 16% of user accounts in less than a thousand guesses given one of the passwords is known to the attacker (Pal, Daniel, Chatterjee and Ristenpart, 2019). Meanwhile, TarGuess-III and IV -- frameworks designed to leverage linked passwords and personally identifiable information (PII) gain success rates over 73% against normal users and over 32% against security-savvy users (Wang, Zhang, Wang, Yan, and Huang, 2016). A similar implementation of TarGuess showed that 78% of the passwords in a list of surveyed passwords could be cracked, with 66% of the passwords cracked in less than 20 guesses (Houshmand and Aggarwal, 2017). Passtrans was able to show that 67.5% of accounts could be guessed in under 1,000 attempts (He, Cheng, Xie, Wang and Liang, 2022). Our password cracking rate of 82.3% surpasses all these values and demonstrates that combining multiple approaches could give attackers a higher success rate.

More importantly, our work brings to the fore the issue of usability/security tradeoff in password management (Di Nocera, Tempestini and Orsini, 2023). Prior works have demonstrated through user surveys and lab experiments that convenience and simplicity trump what users know as secure behaviors when it comes to managing multiple password protected accounts (Haque, Wright and Scielzo, 2013; Haque, Wright and Scielzo, 2014). Our findings further validate this notion by leveraging the actual passwords used by the users on different websites. This calls for a deeper understanding of the behavioral approach to address the tension between usability and security (Di Nocera, Tempestini and Orsini, 2022).

We propose two solutions in this regard. First, as mentioned earlier, a typical user has dozens of password-protected accounts, but they can effectively remember only a few passwords. Hence, assuming that password reuse would be necessary, we recommend that users should reserve strong passwords for their high-value accounts and not reuse those for less important accounts. Furthermore, we recommend reserving a single, easy password for all the less important accounts on the content and sketchy websites. Second, we advocate the implementation of "password tax" (Bonneau and Preibusch, 2010), which would impose a cost on websites for every password-protected account they store. This would prevent sketchy

and content sites from making users register password-protected accounts, and in turn, would reward them instead for using a delegated protocol such as OpenID. As a result, the users would be able to use their limited cognitive capacity to memorize the strong passwords that are only used for their high-value accounts. This reward-based mechanism has also been advocated by the behavioral science researchers to address the usability/security trade-off in cybersecurity (Di Nocera, Tempestini and Orsini, 2022).

## CONCLUSION

In this study, we leverage a well-established user password hierarchy to show the similarity between the passwords of high and low importance to the users. As opposed to relying on survey data from the participants, we use the leaked password datasets from the internet to exploit the lower-level passwords of a particular user to crack the higher-level passwords of the same user. Our findings show that password reuse is a major issue that needs to be addressed to mitigate the usability/security trade-off in cybersecurity. We offer a few suggestions that would enable the users to protect their high-value accounts and demonstrate that password reuse is a nuanced issue that should be addressed more carefully in the future research works in this area.

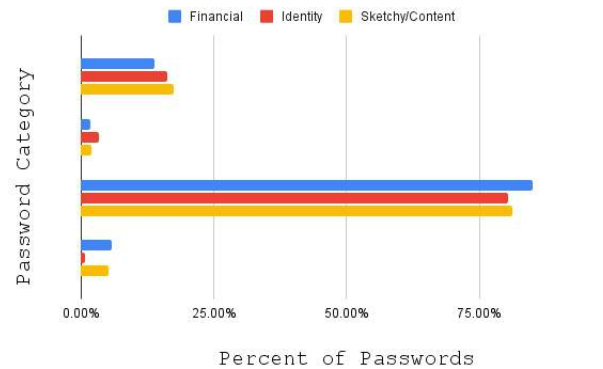
## APPENDIX

### Average Password Lengths and Compositions

Figure A1 shows the average length of passwords from various sources. While it was expected that financial sites would have the largest length of passwords (8.53 characters) and that identity sites would be smaller in length (7.61 characters), the length of passwords from the collection of games in the sketchy/content category was longer than that of identity sites and closer to the length of financial sites (8.37 characters). Figure A2 shows the percentage of various characters present within the passwords from their respective sources. When comparing the password composition to that of passwords in Haque, Wright and Scielzo (2014), they appear similar apart from the occasional outperformance of the sketchy/content passwords over identity passwords (the use of digits and capital letters).



**Figure A1:** Average password lengths.



**Figure A2:** Percentage of characters.

## REFERENCES

- Acin, V. (2019) Making sense of the dark web. *Computer Fraud & Security* 2019(7), pp. 17–19.
- Adams, A., Sasse, M. A. (1999) Users are not the enemy. *Communications of the ACM* 42(12), pp. 40–46.
- Bonneau, J., Preibusch, S. (2010) The password thicket: Technical and market failures in human authentication on the web. In: WEIS.
- Di Nocera, F., Tempestini, G. (2022) Getting rid of the usability/security trade-off: A behavioral approach. *Journal of Cybersecurity and Privacy* 2(2), pp. 245–256.
- Di Nocera, F., Tempestini, G., Orsini, M. (2023) Usable security: A systematic literature review. *Information* 14(12), 641.
- Fahl, S. et al. (2013) On the ecological validity of a password study. In: *Proceedings of the Ninth Symposium on Usable Privacy and Security*, pp. 1–13.
- Güven, E. Y., Boyacı, A., Aydın, M. A. (2022) A novel password policy focusing on altering user password selection habits: A statistical analysis on breached data. *Computers & Security* 113, 102560.
- Haque, S. T., Wright, M., Scielzo, S. (2013) A study of user password strategy for multiple accounts. In: *Proceedings of the Third ACM Conference on Data and Application Security and Privacy*, pp. 173–176.
- Haque, S. T., Wright, M., Scielzo, S. (2014) Hierarchy of users' web passwords: Perceptions, practices and susceptibilities. *International Journal of Human-Computer Studies* 72(12), pp. 860–874.
- He, X., Cheng, H., Xie, J., Wang, P., Yan, J., Huang, X. (2022) Passtrans: An improved password reuse model based on transformer, pp. 3044–3048.
- Houshmand, S., Aggarwal, S. (2017) Using personal information in targeted grammar-based probabilistic password attacks. In: *Advances in Digital Forensics XIII*, pp. 285–303.
- Maschler, F., Niephaus, F., Risch, J. (2017) Real or fake? large-scale validation of identity leaks. *INFORMATIK* 2017.
- Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review* 63(2), 81.
- Pal, B., Daniel, T., Chatterjee, R., Ristenpart, T. (2019) Beyond credential stuffing: Password similarity models using neural networks. In: *IEEE Symposium on Security and Privacy*, pp. 417–434.

- 
- Sicora, M. (2022) How to find saved passwords on Mac. NordPass Website: <https://nordpass.com/blog/how-to-find-passwords-on-mac/>.
- Wang, D., Zhang, Z., Wang, P., Yan, J., Huang, X. (2016) Targeted online password guessing: An underestimated threat. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, pp. 1242–1254.
- Whitten, A., Tygar, J. D. (1999) Why johnny can't encrypt: A usability evaluation of pgp 5.0. In: USENIX Security Symposium, vol. 348, pp. 169–184.

# Human Factors in Alarm Response Procedures: An Empirical Analysis of Paper Versus Digital Support

Chidera W. Amazu<sup>1</sup>, John McGrory<sup>2</sup>, Maria Chiara Leva<sup>2</sup>,  
Gabriele Baldissone<sup>1</sup>, Davide Fissore<sup>1</sup>, and Micaela Demichela<sup>1</sup>

<sup>1</sup>Politecnico di Torino, Turin, 10139, Italy

<sup>2</sup>Technological University Dublin, Dublin, D07 EWW4, Ireland

## ABSTRACT

The objective of a Human Machine Interface (HMI) is to communicate process monitoring information, data, metrics and graphics to an operator through a screen or dashboard and offer an opportunity to control equipment and processes in factories and plants. But after the annunciation of an alarm, how effective is the supporting documentation? The existence of a refined set of instructions and procedures in the form of checklists has been a major factor contributing to the improved safety outcomes observed in the nuclear and aviation sectors. The use of paper-based checklists has been the norm; however, trials of digitized instruction systems have been on the rise in these sectors. The focus of the paper is to analyse an operator on his behaviour and situational awareness from when an alarm is annunciated to the completion of the intervention process using either paper or digitised procedures. The participants (n = 46) were split equally into two groups, each testing three tasks with increasing levels of complexity. Results showed that those who were presented with the procedures on paper had slightly better situational awareness and preferred to use paper procedures when compared to those using the digitised procedures. The rationale for this outcome and recommendation for subsequent redesign of the HMI are presented in this paper.

**Keywords:** Paper procedures, Digital procedures, Control rooms, Human system interfaces, User preference, Situational awareness

## INTRODUCTION

As commercial factories and process plants are becoming increasingly complex in the equipment they use and techniques they perform, it is fair to suggest that the control and support systems they rely on are also becoming progressively multifaceted. We can contend that the function of an HMI is to present process monitoring information, data, metrics and graphics to an operator through a series of screens or dashboards to provide the operator with decision-support material. The HMI should also offer an opportunity for the operator to control equipment and processes. With the increase in plant system complexity, and safety responsibilities are being spread over different departments and vendors, there is now a

greater necessity to support the operator in resolving process faults according to an agreed interdepartmental process. In various safety-critical domains such as process and nuclear industries, documentation supporting alarm response, routine operations, and emergency protocols is paramount for ensuring safety. These documents, through a series of outlined task steps, serve as indispensable guides for operators, aiding them in maintaining plant stability, adjusting process parameters, monitoring, diagnosing issues, making informed decisions, and effectively controlling plant operations. They are an operator support tool used to highlight when to intervene and provide steps on how to rectify abnormal process conditions and have the potential to impact their performance (Barnes and Radford, 1987). Over time, significant advancements have been incorporated into the types and format of support procedures provided to control room operators for executing these activities.

Although considerable effort has been made to normalise HMI components through the introduction of the ISA101 Human Machine Interfaces standard in 2015 to produce high-performance interfaces, there is a lack of clarity on the structure of successful procedures after the annunciation of an alarm.

## **RELATED WORK**

In the nuclear energy domain, there are already migrations afoot from paper to computerised support procedures at various levels based on functionality; Level 1 – Electronic Procedures, Level 2 – Computer-based procedure systems (CBP), and Level 3 – CBPs with Procedure-based automation. Depending on the level, operator supports can range from a primitive level; select and display procedures on a computer screen through to performing sophisticated integrating soft controls to conduct multiple synergised steps on the command of a single operator. (Naser, 2007, IEEE, 2011). In some cases, these supports have been formatted as flowcharts, checklists or numeric task steps. In the aviation sector, checklists have proven to be beneficial for the pilots before take-off, landing or upon the need for troubleshooting during an upset. However, the process industry is still heavily reliant on paper-based procedures for start-ups, shutdowns, troubleshooting or emergency situations.

Generally, procedures in paper form have been criticised for their shortcomings and contribution to many accidents in the process and nuclear sectors. These shortcomings include having to scan and transition between steps or arrange the papers while monitoring and controlling at same time, incompleteness, lack of updates, complexity to follow...etc (Yang et al., 2012, Gao et al., 2013). Also, in an earlier nuclear sector study focusing on different procedure classification schemes and procedure interfaces, it was reported that one of the key issues with procedures is a lack of human factor consideration which leads to several usability challenges that impact safety (Barnes and Radford, 1987). To resolve these issues, the nuclear sector has transitioned from paper to digitised paper to computerised procedures. It is important to make a distinction at this juncture between digitization and digitalization. Digitization refers to digitising information for electronic

storage and retrieval (e.g. paper compared to PDF, or photograph to JPG image), whereas digitalization refers to the inclusion of business rules that aid the person using the electronically stored media (e.g. searching the document for a keyword/link or pattern match an image content).

To understand the effectiveness of these new systems researchers in the nuclear domain have experimentally evaluated their impact on situational awareness, workload and performance. Yang et al. (2012), in their experimental study, compared the computerised vs paper-based procedures on their impact on workload and situational awareness using subjective methods like Nasa-TLX and SART respectively. Their results reported that CBPs reduced mental workload and improved situational awareness compared to paper-based procedures. Other studies have comparing different computerised procedures in flow chart styles (Xu et al., 2008, Gao et al., 2013).

This work aims to analyse an operator on his behaviour and situational awareness from the time an alarm is annunciated through to the completion of the intervention process using either paper or digitised procedures. The rationale for this outcome and recommendation for subsequent redesign of the HMI are presented in this paper.

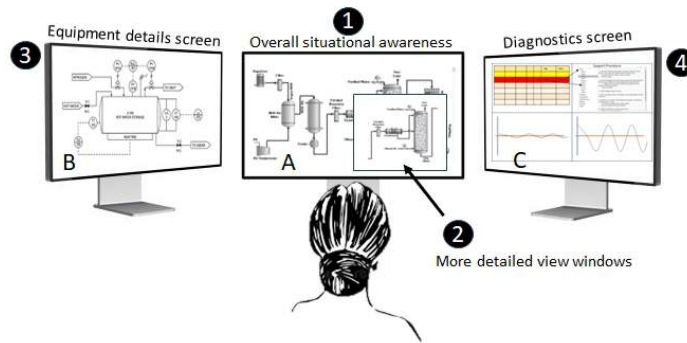
## **METHOD**

### **The Experiment Design**

This study was a between-subject experiment in which individual participants were divided into two separate groups and tasked to perform operator type process resolution instructions and procedures using either a paper-based procedure (G1) or digitised screen-based procedure (G2). Each group comprised of 23 participants. All the participants used in the experiment were equivalent to junior process engineers drawn from the chemical engineering master's students and staff members of the university. The age range of the participants was broad, spanning from 21 to 61 years, with a standard deviation of  $SD = 5.4$  years.

### **The Experiment Setup**

The test rig was a computer-based simulator encoded using MATLAB, primarily based on the four levels of the ISA101 Human-Machine Interfaces standard, as illustrated in Figure 1. The experiment was a single operator environment where the operator was in a seated position with standard room lighting and all three screens are of similar make, model, age with comparable luminance and contrast settings. Screen A, the centre screen, was located directly in front of the operator showing level 1 overall process situation awareness content. Screen A also displayed level 2 windows containing additional selectable detail content. Screen B was tasked with Level 3 content showing specific equipment details. Finally, screen C showed the diagnostics screen information which included the alarm annunciation.



**Figure 1:** Test rig layout.

The alarm annunciation component was located in the top left corner of screen C, which was in line with the F-Shaped pattern layout as identified by Nielsen (Nielsen, 2006). The content was in keeping with the ethos of high performing HMI's incorporating symbols and the limited inclusion of colour. There were additional stimuli components to attract the attention of the operator to the annunciation of an alarm using colour coded alarm banners with black text. A minor alarm appeared as a yellow banner whilst a major alarm appeared as a red banner. There were two types of support mechanism to aid the operator in resolving alarms as illustrated in Figure 2. Mechanism A was where the support documents were provided in a paper format, whereas Mechanism B's support documents were selected from a hierarchical menu located on the diagnostic screen to the right of the alarm annunciation.

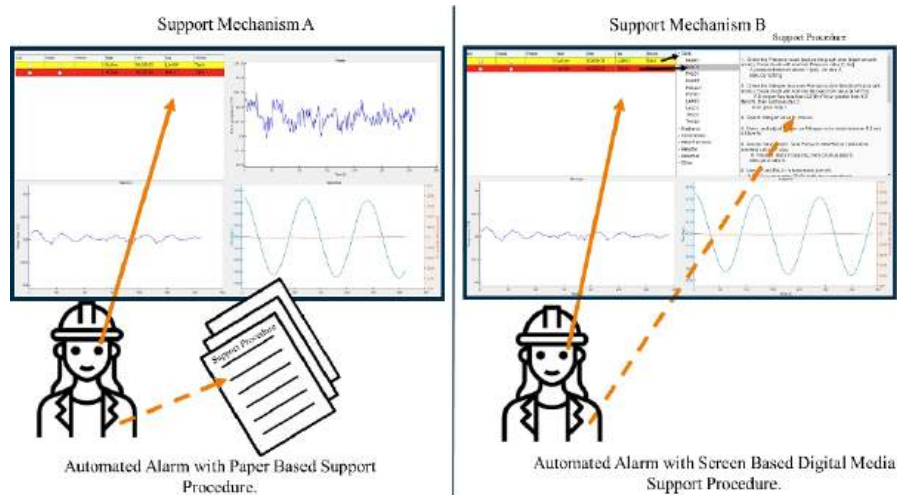
Both paper and on-screen support procedures were presented in a similar font, format and order. The paper-based support was presented on a stack of ordered, single-sided printed pages with a table of contents at the beginning and coloured tab dividers used to separate section areas. The screen-based support was accessed via a tiered navigation menu on the left pane to enable the operator to easily traverse through the sections and respective tag headings to retrieve the appropriate support procedure, as shown in Figure 3. A vertical scroll bar to the right of the procedure allows the operator to move the text down if the procedure bridged two or more pages. The screen-based tiered menu of Mechanism B was not interconnected to the alarm and had no additional stimuli components included, such as colour mapping, jump or hyperlinking to a specific support procedure. Also, no digitalisation using supplementary User Experience UX enhancements, such as, initiative-taking or reactive assistance logic was incorporated. In this experiment, the support procedure was a digitised version of the paper media using the tiered menu. There was no variation of the appearance, timing, order or location of the alarm annunciation for each test participant.

### The Experiment Tasks

The participants were invited to operate a simulated process undergoing three carefully orchestrated events with escalated levels of complexity and supervisor interactivity. The case study is of a formaldehyde production



plant with identified safety-related events. The events were realistic in nature and completed sequentially for each simulation with 5-minute break between them. Each task required the operator to monitor the process, detect alarms, prioritise alarms, perform a situation assessment, plan a response, and implement the response to manage the event following a standardised procedure. Essentially, each task required specific procedures for the troubleshooting, control and evaluation. An example of the screen navigation menu and procedure tasks for tag FAL01 is shown in Figure 3.



**Figure 2:** Support mechanism orientation, (A) paper (B) screen.

- Tank	
PAH01	
PAL01	
FAL01	1. Check the Pressure value [ata] on the graph (see Plot on tank mimic). Cross check with nominal Pressure value [1 ata] if, pressure below or above 1 [ata], do step 2. else, do nothing
FAH01	
PSLL01	
PSV01	2. Check the Nitrogen flow (see Primary system flow [Nm <sup>3</sup> /h] on tank mimic). Cross check with nominal Nitrogen flow value [4 Nm <sup>3</sup> /h] if, Nitrogen flow less than 3.5 [Nm <sup>3</sup> /h] or greater than 4.5 [Nm <sup>3</sup> /h], then continue step 3. else, go to step 1.
LAH01	
LAL01	
TAL01	
TAH01	
• Methanol	
• Compressor	
• Heat Recovery	
• Reactor	
• Assorber	
• Other	
	3. Switch Nitrogen valve to manual.
	4. Move and adjust Pointer on Nitrogen valve scale between 3.5 and 4.5 Nm <sup>3</sup> /h.
	5. Monitor for 10 seconds Tank Pressure with graph on Tank mimic (nominal value = 1 ata) if, Pressure increases, then continue step 6. else, go to step 8 in Procedure PAL01.
	6. Monitor until FAL01 is recovered (turns off). if FAL01 is recovered, then do 7.

**Figure 3:** Digital procedure display.

## RESULTS

Due to missing data on the responses recorded by the participants and supervisors during the experiments, some participants were omitted in the reporting of certain scenarios. This was done to ensure not to bias the results in favour of a particular outcome. The participants rated their preference using a standard 5-point rating scale (5– Outstanding descending

to 1– Unacceptable). The following sections divide the research questions into (a) preference on the perceived level of support and (b) situational awareness where the results are presented, followed by a discussion of the findings.

### Preference on Perceived Level of Support

The participants' rating of their perceived levels of support provided by paper versus digital formats by each group during the three tasks is presented in Table 1.

**Table 1.** Rating of the perceived level of support of paper vs. digital by each group during the three tasks.

PERCEIVED LEVEL OF SUPPORT					
Procedure type (Paper: G1), (Digital: G2)					
Scenario	Group	M	SD	MED	p (Mann-Whitney)
T1	G1	4.33	1.11	5	0.74
	G2	4.23	0.97	4.5	
T2	G1	4.57	0.98	5	0.18
	G2	4.15	0.99	4	
T3	G1	3.48	1.12	4	0.70
	G2	3.32	1.52	3.5	

### Impact on Situational Awareness

When assessing the carefully orchestrated events from the HMI alarm annunciation to end of intervention using the procedures certain scripted questions were asked by the supervisors to evaluate the 3 levels of operator situational awareness;

- I. Perception (Question 1):  
Which of these alarms, in your opinion, must be verified first? and why?
- II. Understanding (Question 2):  
Why do you think the PAL01 alarm is activated? and what do you intend to do?
- III. Projection (Question 3):  
Now that you have done this, what do you think will change in the system? Why?

The questions, which were particularly phrased and scheduled to evaluate the participants' perception, comprehension and projection, were asked using a *think-aloud* approach at three stages during the test without interruption to the tasks. Specifically, a minute after the alarm annunciation, upon starting to read from the procedure, and after the task is observed to have been completed. The alarm-procedure design, participant process knowledge or participant training/experience and combinations of them can influence the outcomes on these three levels. The results of a comparison between paper-based and screen-based support are presented in Table 2 (Perception), Table 3 (Comprehension) and Table 4 (Projection).

**Table 2.** Situational awareness level 1 (perception) comparison between paper and digital support during tasks 1, 2 and 3. Scoring: 1 – very low situational awareness, 5 – very high situational awareness.

PERCEPTION						
Alarm-Procedure Design Support (Paper: G1), (Digital: G2). M – mean, SD – standard deviation, MED – median						
Tasks	Group	No. of responses	M	SD	MED	P (Mann-Whitney)
T1	G1	21	4.86	0.36	5.00	0.40
	G2	22	4.55	1.01	5.00	
T2	G1	21	4.81	0.51	5.00	0.36
	G2	20	4.60	0.75	5.00	
T3	G1	21	3.90	1.41	5.00	0.35
	G2	22	4.32	1.09	5.00	

**Table 3.** Situational awareness level 2 (comprehension) comparison between paper and digital support during tasks 1, 2 and 3. Scoring: 1 – very low situational awareness, 5 – very high situational awareness.

COMPREHENSION						
Alarm-Procedure Design Support (Paper: G1), (Digitised: G2). M – mean, SD – standard deviation, MED – median						
Tasks	Group	No. of responses	M	SD	MED	P (Mann-Whitney)
T1	G1	21	4.29	0.85	4.00	0.15
	G2	20	3.73	1.24	4.00	
T2	G1	21	4.00	0.95	4.00	0.72
	G2	20	4.05	1.15	5.50	
T3	G1	21	3.19	1.69	3.00	0.51
	G2	22	2.86	1.64	3.00	

**Table 4.** Situational awareness level 3 (projection) comparison between paper and digital support during tasks 1, 2 and 3. Scoring: 1 – very low situational awareness, 5 – very high situational awareness.

PROJECTION						
Alarm-Procedure Design Support (Paper: G1), (Digitised: G2). M – mean, SD – standard deviation, MED – median						
Tasks	Group	No. of responses	M	SD	MED	P (Mann-Whitney)
T1	G1	21	4.33	1.15	5.00	0.20
	G2	22	3.68	1.67	4.50	
T2	G1	21	4.48	1.12	5.00	0.51
	G2	20	4.05	1.57	5.00	
T3	G1	21	2.90	1.61	3.00	0.89
	G2	22	3.00	1.63	3.00	

## DISCUSSION

Although the statistical differences of these results are not substantial, there is a clear pattern emerging that there was a higher score in favour of paper-based support procedures compared to screen-based support. Based on related studies by Yang et al. (2012), our original hypothesis that the design of these procedures using screen-based formats would improve performance and score better on the rates of preference and level of support was proven to be false. On the contrary, the data shows that participants in group 2 (G1) with the paper-based procedures consistently rated their preference and level

of support from their procedures higher than those in group 3 (G2) using screen-based procedures across the tasks.

Likewise, those participants with paper-based procedures reported better overall situational awareness, under the lenses of Perception, Comprehension and Projection, than those who used the screen-based support. This outcome on situational awareness is evident, even as the task complexity increased, except for the most complex task in the perception and projection dimensions. Indicating that the screen procedures are likely to be of better benefit for more complex task scenarios where multiple cues such as alarms are present. The navigation design of the procedures on screen could have been more significant in this type of task, hence the observed outcome. Thus, on the first two tasks, with less alarms or complexity, the print version would be preferred.

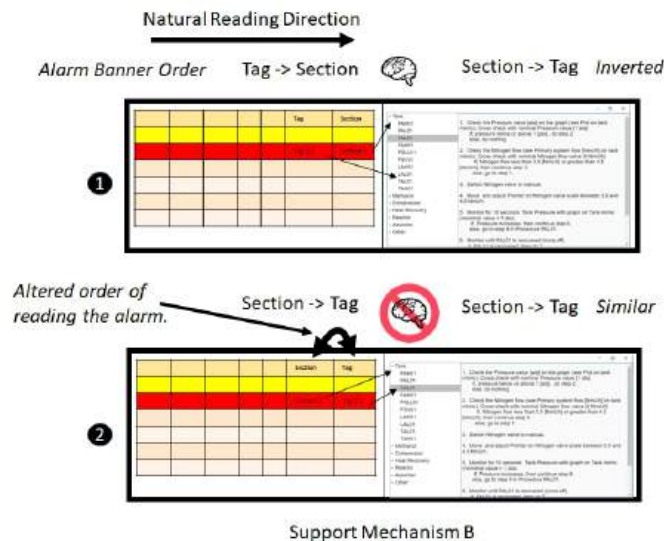
### **Learning From Paper vs. Digital**

Some people harbour the opinion that reading is simply reading and there should not be a difference if the reading is using paper or an equivalent digitised onscreen content. But the research evidence paints a more nuanced narrative specifically when additional factors are included, for example, the purpose of the reading activity. If the purpose of the reading is for learning, the scale tips in favour of paper. The ability to read and engage with the material by underlining, highlighting, and pointing a finger as a placeholder while reading elsewhere is very helpful to the learning process. The addition of summarising notes by creating simplified visual cues aids in cognitively compacting the material and helps in the encoding process. Deep reading of this type fosters text comprehension, memory, and critical thinking. Whereas, if the purpose of the reading is to get the general idea of a story gist, or for checking texts, social media posts and skim reading the scales are tipped in favour of electronic formats. But the support material used in this experiment is read for neither learning nor providing a gist. It is for the express purpose of following a procedure, in particular operator support.

Two areas that have turned around their safety records are the nuclear and aviation sectors and they have specifically concentrated on the following of a procedure with the use of checklists. This has eliminated the dangerous, unreliable, and unpredictable outcomes associated with trusting one's memory response or habit. The research evidence also indicates there are two specific moderators where paper-based activities performed better: (1) Time Frame: time-constrained activity and (2) Text Genre: where there is a mix of informational and narrative texts. (Delgado, Vargas, & Ackerm, 2018). The activity of following a process-specific support mechanism is constrained by time and contains information that is a mix of information and narrative text, which supports the reported findings from the experiment and that swings the preferred preference of the operator to paper-based support documents.

Stepping back from the experiment results, an additional three human factor anomalies were identified during the experiment; (1) On two occasions there was an operator error mix-up when a PAL01 Pressure Alarm Low 01 annunciation prompted the operator to follow the support procedure for the PAH01 Pressure Alarm High 01. Although it's a low instance percentage, on closer examination this operator error may be the result of

two Human Factor characteristics. The first contributing factor being the colour contrast between the black foreground text on a red background. The WCAG Criterion 1.4.3 (AA) relates to contrast for the visual presentation of text and images and sets the minimum acceptable contrast ratio of 4.5:1. The contrast ratio during the test was 4.91:1 which is marginally higher than the WCAG AA criterion. However, the AAA WCAG Criterion 1.4.6. for Contrast (Enhanced) raises the threshold ratio to 7:1. (W3C, 2023). This is a significant finding as red banners are commonly used in HMI design for drawing the attention of an operator. But in doing so the clarity of the text becomes obscure. (2) The second contributing factor being the significance of letter position in word recognition. When reading paragraphs a person does not read every character of each word. The beginning two and end two characters have a greater significance for performing pattern matching when reading where the word is mentally guessed based on the syntax and flow of the sentence (Rawlinson, 1976; Goto, Shirato, & Uda, 2014). On this occasion PAX01 the only difference was the middle character. This observation is also significant as the *tag ID* using similar abbreviations is common in the industry. (3) Thirdly, when a person is reading a series of titles to search in a document, it is good practice to keep the sequence intact. For example, in Figure 4(1) the alarm annunciation shows the *tag ID* first (leftmost) followed by the *section* to the right in a natural language sequence order for reading. Whereas, in the online procedure, used in this experiment the tiered menu for the support document the order is reversed, using the *section* first (leftmost) followed by the *tag ID* to the right. This requires some cognitive processing. An improvement would be to invert the banner to Figure 4(2) where the flow of reading of the alarm banner is *section* then *tag ID* to align similar to that used in the online support documents, thus offering some cognitive off-loading.



**Figure 4:** Inversion of alarm banner order of area and tag.

## CONCLUSION

The study compares the effectiveness of paper-based and digitized procedures in supporting operator responses to alarms in industrial environments. Findings indicate a slight preference for paper procedures and better situational awareness, particularly in less complex scenarios contrary to expectations. This aligns with the published research evidence that indicates there are two specific moderators where paper-based activities performed better: (1) Time Frame: time-constrained activity and (2) Text Genre: where there is a mix of informational and narrative texts. The step-back review of the experiment uncovered banner colour, TagID and sequence of labels as being areas for future exploration.

In addition to the observations and recommendations presented, the author looks to further redesign the digital procedure to include missing human factor elements that can potentially achieve the expected result.

## REFERENCES

- Barnes, V. and Radford, L. (1987). Evaluation of Nuclear Power Plant Operating Procedures Classifications and Interfaces Problems and Techniques for Improvement Prepared for U. S. Nuclear Regulatory Commission. Available at: <https://www.nrc.gov/docs/ml1025/ML102560008.pdf>.
- Delgado, P., Vargas, C. & Ackerm, R., 2018. Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, Nov, Volume 25, pp. 23–38
- Gao, Q., Wang, Y., Song, F., Li, Z. and Dong, X. (2013). Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics*, 56(7), pp. 1070–1085. doi: <https://doi.org/10.1080/00140139.2013.790483>.
- Goto, M., Shirato, T. & Uda, R., 2014. Text-Based CAPTCHA Using Phonemic Restoration Effect and Similar Sounds. Vasteras, Sweden, s.n.
- IEEE Power & Energy Society (2011). IEEE Guide for Human Factors Applications of Computerized Operating Procedure Systems (COPS) at Nuclear Power Generating Stations and Other Nuclear Facilities. IEEE Standards 1786-2011, vol., no., pp. 1–38. doi: <https://doi.org/10.1109/ieeestd.2011.6029264>.
- Naser, J. (2007). EPRI Project Manager Computerized Procedures Design and Implementation Guidance for Procedures, Associated Automation and Soft Controls. [online] Available at: <https://www.nrc.gov/docs/ML0800/ML080080077.pdf> [Accessed 14 May 2024].
- Nielsen, J., 2006. F-Shaped Pattern For Reading Web Content (original study). [Online] Available at: <https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content-discovered/?lm=how-people-read-web-eyetracking-evidence&pt=report>
- Rawlinson, G., 1976. PhD Thesis: The Significance of Letter Position in Word Recognition, s.l.: s.n.
- W3C, 2023. Web Content Accessibility Guidelines (WCAG) 2.2. [Online]. Available at: <https://www.w3.org/TR/WCAG12/>
- Xu, S., Wells, J. W., Li, Z., Zhao, Q., Li, W., He, X. and Gavriel Salvendy (2008). An ergonomics study of computerized emergency operating procedures: Presentation style, task complexity, and training level. *Reliability Engineering & System Safety*, 93(10), pp. 1500–1511. doi: <https://doi.org/10.1016/j.res.2007.09.006>.
- Yang, C.-W., Yang, L.-C., Cheng, T.-C., Jou, Y.-T. and Chiou, S.-W. (2012). Assessing mental workload and situation awareness in the evaluation of computerized procedures in the main control room. *Nuclear Engineering and Design*, 250, pp. 713–719. doi: <https://doi.org/10.1016/j.nucengdes.2012.05.038>.

# Inclusive Smart Navigation Service Design for the Blind and Visually Impaired: A Proposal for the City of Genoa

Francesco Burlando<sup>1</sup>, Boyu Chen<sup>1,2</sup>, and Federica Maria Lorusso<sup>1,2</sup>

<sup>1</sup>University of Genoa, Italy

<sup>2</sup>University of Campania Luigi Vanvitelli, Caserta, Italy

## ABSTRACT

The advancement of interactive technologies offers promising potentials for enhancing mobility and quality of life for visually impaired individuals in smart cities. Goal 11.7 of the 2030 Agenda emphasizes the importance of inclusive, safe, and sustainable urban spaces, specifically addressing the needs of elderly and disabled citizens. This paper examines the potential of interactive technologies, such as smart white canes, wearable aids, and smart devices, in transforming urban public spaces to support the daily activities of individuals with blindness or vision-related challenges in Genoa. By integrating these technologies effectively, this study proposes inclusive service solutions aimed at improving accessibility and promoting independent navigation within the city.

**Keywords:** Smart city, Public space, Interactive technologies, Inclusive design, Sensory substitution

## INTRODUCTION

According to the World Health Organization (WHO), approximately 285 million people are visually impaired, with 39 million of them being blind (Pascolini & Mariotti, 2012). In Italy, there are currently about 2 million people living with visual disabilities, accounting for approximately 3% of the national population (ISTAT, 2017). Walking as a fundamental form of physical activity, is a primary mode of transportation in daily life. However, individuals with visual disabilities often lack crucial environmental information, making navigation, walking, and safety in urban areas significantly challenging, thereby impacting their autonomy and well-being.

Goal 11.7 of the 2030 Agenda (UN, 2015) emphasizes the importance of making cities' green and public spaces more inclusive, safe, resilient, and sustainable, with a particular focus on meeting the needs of elderly and disabled citizens. The construction and advancement of smart cities offer an opportunity to address these challenges. The European Commission defines smart cities as using intelligent transportation networks and digital technologies to enhance urban management responsiveness and create safer

public spaces tailored to urban residents' needs. Interactive technologies and devices within smart city transportation networks can effectively overcome existing deficiencies in facilities. This not only facilitates future accessibility enhancements but also significantly enhances safety and quality of life for visually impaired individuals.

This paper will explore the potential of interactive technologies in urban public spaces for people with visual impairments. We aim to not only enhance their travel accessibility through technologies and navigation systems but also consider the quality of their urban walking experience. Additionally, we will propose integrating these technologies and navigation systems into public spaces to enhance daily experiences for visually impaired individuals in Genoa.

### **WAY OF INTERACTION BETWEEN VISUALLY IMPAIRED PEOPLE AND URBAN ENVIRONMENT**

Visually impaired individuals are characterized by varying degrees of vision loss that hinder conventional problem-solving methods (e.g., eyeglasses). Blindness, on the other hand, refers to a complete inability to see due to injury, disease, or genetic conditions, with total blindness indicating a complete loss of vision (WHO, 2018). According to definitions, visual impairment encompasses moderate, severe, and complete categories (Cruciani, 2005). Partially impaired individuals with reduced visual sensitivity and limited visual fields often prefer designated routes that are trained and familiar. In most cases, the suitability of a route is determined not solely by its distance but by its navigational ease with minimal obstacles or conflicts (Parkin & Smithies, 2012). For those who are completely blind, environmental cues become crucial for gathering information that influences their mobility and activities (Marston & Golledge, 2004).

Human beings typically acquire approximately 80% of their total information through vision (Man & Olchawa, 2018). To compensate for this lack of visual input, the theory of sensory compensation suggests that when one sense is diminished or absent, others are heightened to capture environmental information for localization and navigation purposes (Pieniak et al., 2022). In addition, inclusive design considers individuals with temporary or short-term audiovisual impairments encountered daily in urban environments as a disability situation. Therefore, creating inclusive environments for people with sensory impairments, also can benefit all the citizen's daily life. Simultaneously, the rapid advancement of digital technologies offers new opportunities for enhancing human-computer interaction and providing improved assistance for mobility and access to information for the visually impaired.

### **TECHNOLOGIES APPLIED FOR VISUALLY IMPAIRED'S WALKING TRAVEL**

Although individuals with vision impairments often enhance senses such as hearing and touch to navigate daily activities, they may also rely on optical



aids and techniques to compensate for their visual limitations. Specifically, in orientation and mobility, they utilize human guidance, guide dogs, echolocation techniques, or traditional white canes to detect and navigate around obstacles. However, the expanding array of emerging technologies is enriching the range of assistive aids available. Bouck (2015) defines assistive technology as a collection of devices and tools designed to enhance the functional capabilities of people with disabilities. Cities must increasingly incorporate smart street equipment and devices equipped with assistive technology to ensure users feel considered and safe. The combination of personal aids and urban technologies significantly enhances the experience of blind and visually impaired individuals in urban settings, promoting independence, safety, accessibility in daily activities and movements, and spatial awareness. Following a thorough review of the literature, we have identified and classified three main categories of personal technology devices based on their scenario, function, and technology.

### **Smart White Cane**

The first category concerns one of the most popular traditional aids. The white cane is the most common symbol that signals to people around that the user is visually impaired or blind.

Its use requires intensive training and significant effort on the part of the user (Sobnath et al., 2020), however, it is particularly effective for exploring the surrounding space, orienting oneself, and detecting obstacles on the ground and unevenness independently. Unfortunately, the cane alone provides limited information; it cannot predict long-range obstacles or the ones that do not rest on the ground, such as overhanging branches or low-hanging structures. For these reasons, in recent years, canes for the blind have been implemented with microelectronics containing GPS navigation systems, proximity sensors, and LEDs to illuminate the tip of the cane and be recognized in dimly lit environments, all of which can provide the blind user with a more complete picture of the space around them. A technologically advanced project of particular interest in fostering urban autonomy for blind and visually impaired people is the LETIsmart system, developed by SCEN and co-designed with Marino Attini, a visually impaired user and National Advisor to the Italian Union of the Blind and Visually Impaired. The system is composed of two kits, LETIsmart LUCE and LETIsmart VOCE - which make one's traditional white cane smart - and a sophisticated network of easy-to-install radio beacons placed at strategic points in the city and on public transportation, which can provide voice information to guide you to your destination through the LETIsmart VOCE kit and with the personal devices of citizens who need them, using LORA technology.

The first kit, LETIsmart LUCE, is a battery-powered lighting device that is installed on the lower end of the white cane. It consists of a translucent tube that, depending on ambient lighting conditions, allows a pulsating beam of light to shine through. Once mounted, it makes blind users more visible to other pedestrians or vehicle drivers thanks to the flashing light,

improving the safety of urban travel for both the blind person and others. The second, on the other hand, is a smart handle that can replace the traditional white cane handle or can be used as a pocket version to receive voice information pertaining to urban navigation or city events. In fact, LETIsmart VOICE contains microelectronic components that allow it to connect to beacons placed at useful orientation points, such as at traffic lights for crosswalks, public transportation stops and on bus doors, construction sites that impede movement, and entrances to public buildings and stores. The smart handle has multiple modes of use, including bus information, city information, and business information. Depending on the mode chosen by the user, beacons installed within a 50-meter radius will emit a radio voice message to be detected and reached, such as “Pedestrian crossing of Via Cesare Battisti.” From a technological point of view, the two kits to be integrated into one’s white cane are particularly advantageous, as they contain microelectronic components that do not affect the ergonomics and do not compromise the lightness of the white cane (only 11.5 grams). This allows the user to benefit from a lightweight, foldable product that does not fatigue the user during use and greatly enhances his or her experience in the city. On the other hand, the LETIsmart system only works if there is an adequate and widespread beacon infrastructure in the city. Each public or commercial entity can equip itself with beacons that play the name of its business as if it were a vocal sign (for the visibly able-bodied), or that announce news in the storefront. For the system to work properly, it is therefore essential that cities equip themselves with a massive amount of beacons that make every corner of the city inclusive and easily usable by the blind user.

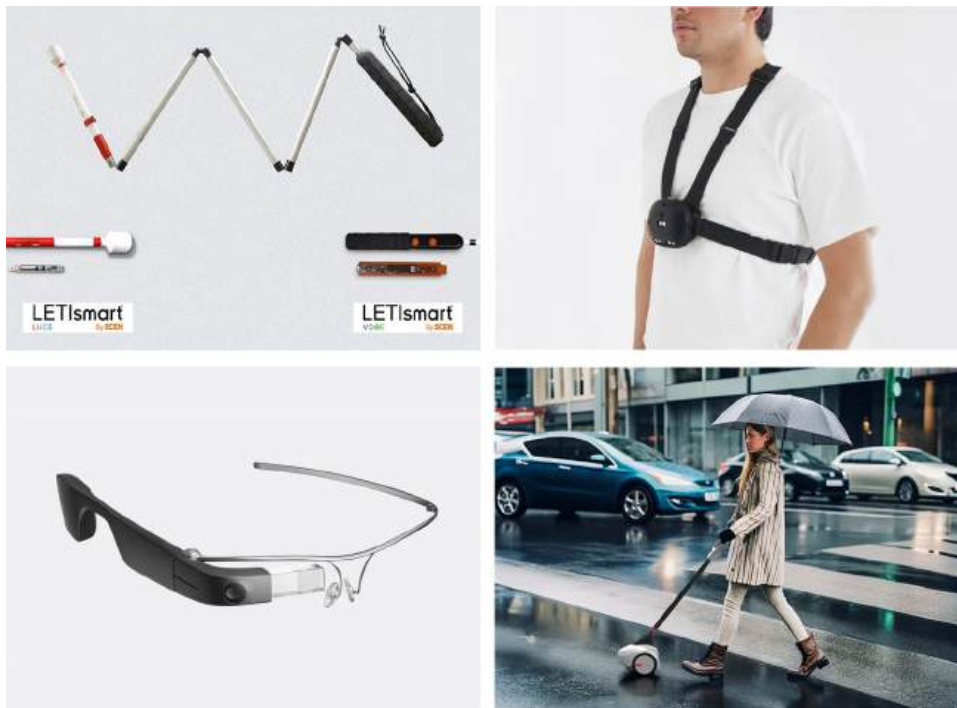
### **Wearable Devices**

In recent decades, numerous wearable technological aids have been developed that assist visually impaired people in numerous daily activities without having to resort to large devices. They allow hands-free interaction or minimize the use of hands, as in most cases they are operated through voice commands and are directly worn on different parts of the body, such as on the head, wrists, chest and abdomen (Velázquez, 2010), with the purpose of providing real-time information. Among the many, the most popular are smart glasses and visors with AI technologies, followed next by chest harnesses, bone-conduction headphones, wristbands, and sensorised shoes. One emblematic device is Envision Glasses, developed by Envision based on the Google Glass Enterprise Edition 2. These are lightweight smart glasses that house traditional components, and are installed with AI-powered features and the use of advanced algorithms to identify objects, read text and recognize faces in real-time. As AI-powered smart devices such as Chat GPT, they need to be connected to the Internet at all times. They are configured through Envision App and also integrate navigation apps such as BlindSquare, to guide the user step by step by providing auditory cues and directions. Ultimately, smart glasses like Envision greatly support visually impaired or blind people in the city and can enrich their

experience through constant interaction with artificial intelligence. However, these devices are not designed to completely replace traditional mobility aids such as white canes or guide dogs, but rather act as a complementary support. According to users, they are an excellent substitute for the smartphone, as they allow for many hands-free activities, since the interaction is through voice. Unfortunately, they are not suitable for deaf-blind people, as they only provide audio output such as sounds and words to communicate information, and there is no way to set haptic feedback or vibration. As mentioned above, chest harnesses can also provide support. These are wearable devices designed specifically for walking, which through a series of sensors, such as Infrared Time of Flight sensors and ultrasonic sensors, accelerometer and gyroscope, capture real-time contextual data from the environment, alerting users to impending obstacles. This specific assistive technology makes it possible to improve walking speed and posture, as it allows for hands-free walking. Part of this type is the Ara project, designed by Strap Tech in 2022. Ara is a lightweight, rechargeable device that is worn on the chest that also identifies obstacles that cannot be detected with the classic white cane, i.e., those from chest height upward, and is therefore suitable for use alone. Ara does not overload the user with information, in fact it interacts with the user through sequential vibrations only when it detects obstacles, telling the user their location (high, medium or low). The advantage of such tactile language is that it can be universally understood, breaking down language and hearing barriers, thus making it suitable even for deaf-blind users.

### **Smart Aids**

Smart aids come in various sizes and technologies tailored to specific needs and preferences. Among the most widely adopted are smartphones, which have become indispensable assistive devices for people with visual impairments. Due to their compact size and lightweight nature, smartphones are convenient for navigation and interacting with the environment. Their operating systems include a range of accessibility features such as screen readers, magnifiers, high-contrast screens, and speech-to-text capabilities, which facilitate common daily tasks for visually impaired individuals (Muhsin et al., 2024). One noteworthy smart device is Glide, a self-guided mobility aid devised by Glidance, which is currently undergoing trials, but manifests great potential for facilitating the visually impaired. Ostensibly resembling a white cane, Glide employs advanced navigational robotics technologies, artificial intelligence, sensors and a camera to recognize and avoid obstacles, guiding users effectively. It features two 7-inch wheels that steer independently to avoid obstacles and guide the user along the correct path. Operating Glide is straightforward: the user pushes it via an ergonomic handle and follows its lead. Additionally, Glide adjusts to the user's walking speed and provides voice feedback describing the surroundings, thereby enhancing the navigation experience.



**Figure 1:** Above, left, the two LETIsMART kits for the white cane; right, Ara chest harness. Bottom, left Envision Glasses; right, self-guided mobility aid Glide.

## PROPOSAL IN THE CITY OF GENOA

The Liguria region, including its capital Genoa, has a significant population affected by visual disabilities (ISTAT, 2017). Genoa, characterized by its complex and diverse terrain of narrow alleys, intricate streets, and stairways typical of Italian cities, poses daily navigation challenges for blind residents (Campisi et al., 2019). As a historic port city, Genoa not only possesses rich cultural heritage and natural beauty but also embraces advanced technologies like Digital Twin, a state-of-the-art simulation tool enabling real-time monitoring and management of city infrastructure. Moreover, Genoa has established a profound foundation in interactive technologies, actively participating in smart city initiatives such as RAISE (Robotics and AI for Socio-economic Empowerment). This three-year project, funded by the Ministry of University and Research (MUR) through the National Recovery and Resilience Plan (NRRP), is an innovation ecosystem coordinated by the Liguria Region which hosts the main robotic and AI research laboratories of Italy. Collaborating with leading research institutions like the National Research Council (NRC), the Italian Institute of Technology (IIT), and the University of Genoa (UNIGE), RAISE project aims to develop and commercialise technological solutions based on robotic systems and artificial intelligence to meet the productive and social needs of the Liguria region (Burlando et al., 2024). The project targets areas including health, environmental sustainability, smart ports, and urban technologies, aiming

to elevate the city's technological landscape and enhance accessibility for visually impaired individuals.

Despite these advancements, Genoa still faces significant challenges in ensuring equitable access to public spaces such as streets and parks for visually impaired residents. However, adopting user-centric approaches and leveraging emerging technologies can offer innovative solutions to overcome these barriers. Enhancing the accessibility of public spaces through smart city concepts is crucial in improving the overall quality of life and inclusivity for visually impaired individuals in Genoa.

### **SMART NAVIGATION SERVICE PROPOSAL**

Due to its unique characteristics, Genoa presents an ideal spatial context for developing prototype solutions involving interactive technologies to support the blind and visually impaired community. Utilizing open data from the city's Digital Twin and potential collaborations with institutions like the UNIGE and the David Chiossone Foundation, which specializes in the rehabilitation of visually impaired individuals, offer a promising starting point for enhancing existing smart devices available to this community.

Genoa has shown interest in projects like LETIsmart, highlighting an opportunity to expand its infrastructure with radio beacons. These beacons could be strategically placed at key locations such as theaters, meeting points, parks, and scenic viewpoints. Their role would be to convey information audibly to blind individuals, akin to reading a poster, or describing their surroundings in detail without requiring personal smart devices. This approach is crucial since many assistive devices are costly, posing a barrier to widespread adoption without governmental or municipal support in making cities more inclusive.

Additionally, implementing intelligent navigation systems and interactive information stations in parks can complement radio beacons. These systems offer voice-guided navigation and provide multi-sensory feedback, including tactile and auditory cues, tactile maps, and diverse flooring materials, all aimed at enhancing navigation and enjoyment of park environments for visually impaired individuals.

From an inclusive design perspective, urban infrastructures can further be optimized to specifically meet the needs of the visually impaired by integrating multi-sensory experiences. For instance, leveraging AI-powered information totems or radio beacons to highlight seasonal scents of surrounding plants can enrich the navigation experience. Real-time updates could inform users, "You're smelling wisteria in bloom; it blooms in April with purple flowers". Finally, we must not overlook the dictates of the New European Bauhaus (NEB), which posits not only inclusion but also sustainability and beauty as fundamental pillars for building the future of our society (Rosado-Garcia et al., 2021). Ensuring that infrastructure designed to enhance city inclusivity does not compromise urban aesthetics is paramount. Equally important is cultivating an environment where visually impaired individuals can derive aesthetic pleasure from urban spaces. This challenge is complex, considering the diverse perceptions of beauty among individuals

with visual impairments, yet crucial for dignifying their urban experiences (Bolt, 2014). The challenge is therefore on two different levels: on the one hand, making sure that all the infrastructure useful for making cities more inclusive does not undermine the aesthetics of the urban environment. On the other, to pursue the achievement of good aesthetic perception even in the enjoyment of urban spaces by the visually impaired and blind. Only in this way, by recognizing that everyone has the ability and therefore the right to enjoy a beautiful and pleasant environment, is proper dignity given to these categories of users (Bolt, 2013).

As mentioned, this is not an easy challenge, especially considering that there are multiple categories of visual impairment and multiple levels of severity of the condition. The combination of these two factors produces huge differences in users' needs and, therefore, in users' perceptions of the solutions that could be identified (Zhao et al., 2018). If, as mentioned above, the ideal path from point A to point B for such users is not the shortest but the one with the fewest obstacles, the role that the presence of pleasantly satisfying multisensory elements plays in route identification must become equally important. It is important to note that, in full compliance with the dictates of Design for All (DfA), this does not prove to be an exclusive choice for users with disabilities. In fact, if one thinks of elements that normally enhance the urban experience of users, such as green areas, one can see how what for many is an improvement in terms of primarily visual and climatic, actually brings aesthetic benefits at the auditory and olfactory levels as well, going on to improve the aesthetic experience of hypo and visually impaired users.

## CONCLUSION

The design of navigation assistive technologies for visually impaired people represents a promising yet challenging research area. Researchers and designers face the complex task of understanding how these technologies integrate into urban life and impact citizens, especially marginalized populations requiring heightened attention and understanding.

This paper conducts an analysis of existing aids and assistive technologies for blind and visually impaired individuals, including smart white canes, wearable aids, and smart devices, to assess their strengths and limitations, and identifying opportunities for design actions. Furthermore, we propose a conceptual solution aimed at enhancing inclusivity for visually impaired individuals in public spaces within the city of Genoa. This proposal serves as a foundational step towards future research endeavors in this field. In addition, we aim to optimize system functionalities and expand implementation to additional cities, ensuring that more blind and also vulnerable individuals can benefit from the convenience and welfare offered by smart city initiatives. Simultaneously, we advocate for increased attention across all sectors towards the needs of disadvantaged groups in the development of smart cities, fostering a more inclusive and supportive social environment.

## ACKNOWLEDGMENT

The paper is the result of a joint reflection by the Authors carried out with the supervision of Francesco Burlando.

The “Introduction” and “Way of interaction between visually impaired people and urban environment” paragraphs have to be attributed to Boyu Chen, “Technologies applied for visually impaired’s walking travel” to Federica Maria Lorusso, “Proposal in the city of Genoa” and “Conclusion” to Francesco Burlando.

## REFERENCES

- Bolt, D. (2014). An advertising aesthetic: Real beauty and visual impairment. *British Journal of Visual Impairment*, 32(1), pp. 25–32.
- Bolt, D. (2013). Aesthetic Blindness: Symbolism, Realism, and Reality. *Mosaic: An Interdisciplinary Critical Journal*, 46(3), 93–108. <http://www.jstor.org/stable/44030343>
- Bouck, E. C. (2015). *Assistive technology*. Sage Publications.
- Burlando, F., Lorusso, F. M., & Porfirione, C. (2024). RAISE ecosystem: urban design for accessible and inclusive Smart Cities, in Bisson, M., *Environmental Design*. 4th International Conference on Environmental Design, 9–11 May 2024, Palermo University Press.
- Campisi, T., Canale, A., & Tesoriere, G. (2019). The development of walkability in the historic centre of Enna: The case of the Saint Tommaso neighbourhood. *European transport: International journal of transport economics, engineering & law*, (73), pp. e4–e4.
- Cruciani, F. (2005). A proposito della legge 138/01: “Classificazione e quantificazione delle minorazioni visive e norme in materia di accertamenti oculistici”. A che punto siamo. *Oftalmologia sociale*, pp. 4–9.
- European Commission. Smart cities, European Commission. Available at: [https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities\\_en](https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en) (Accessed: 15 June 2024).
- ISTAT. (2017). Italian National Institute of Statistics. Available at: <https://www.istat.it/en/> (Accessed: 15 June 2024).
- Man, D., & Olchawa, R. (2018). The possibilities of using BCI technology in biomedical engineering. In *Biomedical Engineering and Neuroscience: Proceedings of the 3rd International Scientific Conference on Brain-Computer Interfaces*, BCI 2018, March 13–14, Opole, Poland (pp. 30–37). Springer International Publishing.
- Marston, J. R., & Golledge, R. G. (2004, May). Quantitative and qualitative analysis of barriers to travel by persons with visual impairments and its mitigation through accessible signage. In *Proceedings of the 10th international conference on mobility and transport for elderly and disabled people*, Hamamatsu, Japan, pp. 23–26.
- Muhsin, Z. J., Qahwaji, R., Ghanchi, F., & Al-Tae, M. (2024). Review of substitutive assistive tools and technologies for people with visual impairments: recent advancements and prospects. *Journal on Multimodal User Interfaces*, 18(1), 135–156.
- Pascolini, D., & Mariotti, S. P. (2012). Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, 96(5), 614–618.
- Parkin, J., and Smithies, N. (2012). Accounting for the needs of blind and visually impaired people in public realm design. *Journal of urban design*, 17(1), 135–149.

- Pieniak, M., Lachowicz-Tabaczek, K., Karwowski, M., and Oleszkiewicz, A. (2022). Sensory compensation beliefs among blind and sighted individuals. *Scandinavian journal of psychology*, 63(1), 72–82.
- Rosado-García, M. J., Kubus, R., Argüelles-Bustillo, R., and García-García, M. J. (2021). A new European Bauhaus for a culture of transversality and sustainability. *Sustainability*, 13(21), 11844.
- Sobnath, D., Rehman, I. U., & Nasralla, M. M. (2020). Smart cities to improve mobility and quality of life of the visually impaired. *Technological Trends in Improved Mobility of the Visually Impaired*, 3–28.
- UN. (2015). *Transforming our world: The 2030 agenda for sustainable development* | department of economic and social affairs. Available at: <https://sdgs.un.org/2030agenda> (Accessed: 15 June 2024).
- Velázquez, R. (2010). Wearable assistive devices for the blind. *Wearable and Autonomous Biomedical Devices and Systems for Smart Environment: Issues and Characterization*, 331–349.
- Zhao, Y., Kupferstein, E., Tal, D., & Azenkot, S. (2018, October). “It Looks Beautiful but Scary” How Low Vision People Navigate Stairs and Other Surface Level Changes. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 307–320).



# The Design of Generation Z Camping Car System Based on Grounded Theory-AHP Method

Shihan Tang, Qijun Ye, Wenjing Wei, and Bowen Zhou

Nanjing University of Science and Technology, Nanjing, 210094, China

## ABSTRACT

The aim of this paper was to propose a design strategy for Generation Z user camping car systems by studying the scenario needs and preferences to improve the users' experience. The research methodology was based on the Grounded theory to summarize the user requirement model, as well as combining the AHP method to extract the key features for design planning. In the research process, firstly, semi-structured interviews were used as the basis to derive the requirement information. Secondly, the results of the interviews were analyzed through the combination of qualitative (Grounded theory) and quantitative (AHP) research methods. The requirements model of the target population and the weighted ranking of each requirement were output successively to ensure the accuracy of design and development. This research finally output a user requirements model, which mainly consisted of 3 core requirement categories, 7 main requirement categories and 30 independent requirement categories. 14 key demand functions were also extracted. Based on this conclusion, this paper described the design strategy of camping car centered on Generation Z users in detail. Following this, a whole vehicle system was designed, which mainly includes the interior and exterior of the vehicle and the interaction screen.

**Keywords:** User-centered research, Grounded theory-AHP method, Camping car system design, Generation Z

## INTRODUCTION

Outdoor recreation and close contact with nature have a positive impact on human emotions and interactions between social agents (Fagerholm et al., 2021). In the wake of the epidemic, an increasing number of people have realized the positive impact of outdoor activities, and as camping is the most natural way of outdoor recreation, the “car camping fever” has set off a trend in China. Literature searches can be found that car camping industry has been developed for a long time in western countries, and the related research has been more mature. However, there are still differences in the field between China and the West due to economic development and national characteristics, so it is important to carry out targeted and systematic research on the car camping industry in China. Several researchers have found that young people of Generation Z, as one of the major consumer groups, have distinct preferences and consumption differences (Canqun and Yongqi,

2023), such as tolerance of diversity, love of personalization, interest-driven consumption, etc. Because of this, they are more inclined to glamping, which is mainly manifested in higher requirements for camping scenes, recreational facilities, and equipment styles. It means that campers put their emphasis on the activity experience. In addition, Chinese people usually choose private cars as camping vehicles rather than an RV, but the basic functions of an ordinary private car are obviously unable to meet the users' camping needs any more. Consequently, the fundamental purpose of this paper was to propose a design strategy for Generation Z user camping car systems by studying the scenario needs and preferences to improve the users experience.

## **CURRENT STATUS OF CAR CAMPING AND RELATED DEVELOPMENTS**

The term “car camping” refers to a range of activities in which ordinary vehicles, modified vehicles, campers, etc., stay in places where regular accommodation is not provided (Jung et al., 2022). This paper will mainly focus on the research under the characteristics of China's national conditions, while analyzing the relevant research of foreign car camping for reference. The main status was as follows:

**Development and differences in camping culture at domestic and overseas:** The camping culture is still slightly varied from place to place due to differences in economic development and national conditions. The definitions of camping are obviously not the identical, for example, Frank et al. believe that camping tourism is a leisure activity with cultural attributes for social, leisure, educational and therapeutic purposes (Fu et al., 2013). Domestic scholars, such as Dawei Xu, believe that camping is a recreational activity in which tourists seek natural interest in the wilderness to satisfy their individual needs (Dawei, 1996). In addition, due to the development of camping culture, the term “Glamping” has become popular in recent years. Especially in Asian countries such as China, this style of camping has become the main pursuit of young people. In the West, the word means “stylish, special camping”, but after being introduced into China, this style of camping has gradually changed into the pursuit of rituals, with stronger leisure and social attributes, hoping to create a sense of ambience and sophistication of the leisure scene (Zhipeng et al., 2023).

**Types and evolution of camping cars:** Different from western countries, due to the more restrictive conditions of RV camping, Chinese campers are more willing to use a private car that meets both commuting and camping requirements. Therefore, this paper will focus on the study of camping vehicles using private cars as carriers. In the early days, when people pursued practical functions during camping, they often explored more space inside and outside the vehicle to support the storage space for equipment and rest space, such as carrying a tent on the roof and putting down the seats to increase the usable space. With the advancement of technology, more and more designs focused on user experience, so unique design language and intelligent technology are added to the design of camping car. For example, the “camping mode” of Ideal ONE and L9 is to adjust the temperature and air quality inside the car by giving instructions to the vehicle, so that users can be more comfortable in the car for a long time. Wuling Gameboy

captures the young people's love of games and entertainment. Combined with related culture, the design language of games is very distinct on the exterior decoration.

To sum up, Chinese campers have higher requirements for camping scenes, entertainment facilities, and equipment styles, which means that campers will pay more attention to the experience of activities. Nowadays, more and more car designs have made a lot of efforts to improve the user's camping experience, such as making the function intelligent, or giving a special design style. It is not difficult to see that improving user experience will be the main direction of future camping car design.

## **OVERVIEW OF THEORETICAL METHODS AND INNOVATIVE INTEGRATION**

Grounded theory is a complete and independent qualitative research method proposed by Barney G. Glaser and Anselm L. Strauss (Glaser et al., 1968), whose purpose is to establish theories on the basis of empirical data (Glaser, 1967). First of all, the original materials were obtained through in-depth interviews. Secondly, according to the core idea of grounded theory, the original data are analyzed and summarized through three levels of coding: open coding, relational coding and selective coding (Mccallin, 2003), so as to obtain independent category, main category and core category. Finally, the theoretical model is obtained by combining the three categories.

AHP, proposed by American operations research scientist THOMAS LS in the 1970s, is a decision-making method that makes qualitative and quantitative analysis of research objects at the same time. It is comprehensive, concise and intuitive (Xue et al., 2012). This method needs to build a hierarchical structure model and transform the model into a judgment matrix. After that, experts need to be invited for pound-for-pair comparison, scoring and assigning values, and key elements can be obtained by calculating the weight value of the data and sorting in turn.

Both research methods have their own advantages, but also have limitations. A combination of methods is needed to overcome the shortcomings of a single approach. Grounded theory is a qualitative research method, which is based on the induction of dialogue information of the interviewer. This method can build a theoretical model through layers of coding and provide guidance for subsequent design practice. However, for some research with relatively large information data, the primary conceptual categories after coding will still show complex characteristics. Too much information will fail to present clear priorities for design guidance. In order to ensure that the design presents the transformation of the user's key needs, it is necessary to involve quantitative analysis methods. As a quantitative research method, AHP can objectively output the weight value of each demand. By ordering the weight value, key requirements can be clearly obtained and the accuracy of design decisions can be improved. According to the literature search, there is no application of the grounded theory and AHP method to the user demand research of camping vehicles. Therefore, this paper will propose an innovative applied research method for the system design of camping vehicles. The main steps are as follows:

1. **Target scene and crowd research.** This part used ethnographic observation and semi-structured interview to sort out the whole process of self-driving camping trip. From this, we could get users' basic problems about cars and camping equipment, pain points and contact points to prepare for in-depth interviews.
2. **In-depth needs mining of users.** Conduct in-depth interviews with users based on the interview outline obtained from preliminary interviews, and collect original needs.
3. **User requirement model construction.** Analyze and summarize the original data through three-level coding, in turn, summarize them into conceptual categories, scope categories and core categories, and come up with a user requirement model for the design of the Z-generation campervan system.
4. **Key demand extraction.** Transform the demand model into a hierarchical recursive model and construct a judgment matrix. Extract the key requirements by weighting the scores of experts to derive the ranking order of the requirements.
5. **Application of design strategy.** The user requirement model was used as a strategy to build the overall framework of the campervan system. The key requirements were focused on planning and polishing the functions of the camping car system.

## **RESEARCH ON GENERATION Z CAMPING CAR SYSTEM BASED ON GROUNDED THEORY-AHP METHOD**

### **Target Scene and Crowd Research**

To avoid the influence of subjectivity on the interview questions, field research on camping scenarios and target users was conducted before the study. Before the formal study, the researchers invited some car camping enthusiasts to complete a car camping activity together. The location of activity was set in a park campground in Jiangsu Province, China, and the age limit of the enthusiasts was  $22 \pm 4$  years old. This activity was conducted both to allow the researchers to familiarize themselves with the car camping activity process and to be able to observe the behavior of the target users during the activity. At the end of this activity, the researchers conducted a simple interview with the participants, with the aim of initially collecting the needs and pain points of the users, which would be served as a reference guide for the in-depth interview afterwards.

### **In-Depth User Requirements Analysis and Model Establishment**

After the preliminary scene investigation and analysis, we constructed a more detailed semi-structured interview. Twenty young men and women (1:1) were included in the study ( $n = 20$ , age  $22 \pm 4$  years, car camping enthusiasts). Participants needed to join in a semi-structured interview lasting about 30 minutes. This interview mainly focused on the needs and preferences of users in the car camping scene. The content of the interview is divided into five sections: user information, car body needs, scene accessories needs,

experience statements and expected needs, with 2–3 questions in each section. In this study, users' needs and preferences will be collected during the interview process as the original data for subsequent research.

In the information analysis and processing, the research method of grounded theory was added. The original interview data was analyzed and summarized through three-level coding, and the three-level demand model was summarized upward. In this paper, Nvivo software was used to assist the coding analysis of the original data. In open coding (level 1 coding), the original interview sentences needed to be converted into conceptual vocabulary. In order to ensure the objectivity of the model, it was necessary to eliminate words with very low frequency. In this paper, 5 words with a frequency of 1 were eliminated. After that, 30 independent category genera were output (see Figure 1. D1–D30). In relational coding (level 2 coding), its main task was to compare different categories, dig deep connections among them, and integrate high-level category categories with internal logic. In this paper, 30 three-level independent genera were analyzed, and 7 main categories were obtained after summary (see Figure 1. C1–C7). In selective coding (level 3 coding), it was necessary to abstract more core categories from the categories and establish an organic relationship between core and main categories. Through the abstraction and induction of the main categories, 3 core categories were obtained in this paper (see Figure 1. B1–B3). After all the codes were finished, a “User Requirements Model for Generation Z Camping Car System” with user experience requirements was finally constructed through the series of the three levels of codes. In order to ensure the validity and reliability of the research results, five pre-reserved interview data were again tested for theoretical saturation according to the above three-level coding steps. After sorting and comparing and analyzing, no new concepts and categories were generated, which reached “theoretical saturation” in a sense (Shan et al., 2020).

Target layer	Core category	Main category	Independent category
A Camper system design requirements based on Generation Z users	B1 Body Structure Requirements	C1 Rational automotive architecture	D1 Off-road performance D2 Large open space for use D3 Comfort Seat D4 Flexible change of seat D5 Vehicle equipment list D6 Storage Modularity D7 Flexible application of the entire vehicle space D8 Connected space inside and outside the vehicle
		C2 Flexible layout	D9 Convenient access to goods D10 Easy to install equipment D11 Refrigeration function D12 Camp lighting D13 Mosquito protection D14 Sleep and breaks D15 Garage collection D16 Individual air conditioning in the car D17 Heating requirements D18 Outdoor Cooking D19 Fire and supply independence D20 Water storage space D21 Audio music D22 On-board camera function D23 Atmospheric control D24 Air permeability D25 Insulation D26 Monitoring and protection function D27 Intelligent control D28 Prohibit while driving D29 Environmental adaptation D30 Multi-vehicle interaction
		C3 Easy handling	
	B2 Functional requirements for vehicle matching	C4 Practical scenario functions	
		C5 Ambient Entertainment	
	B3 Emotional requirements	C6 Safety and security needs	
		C7 Personalized Experience	

**Figure 1:** Generation Z camping car system requirements hierarchy model (self-drawn by the authors).

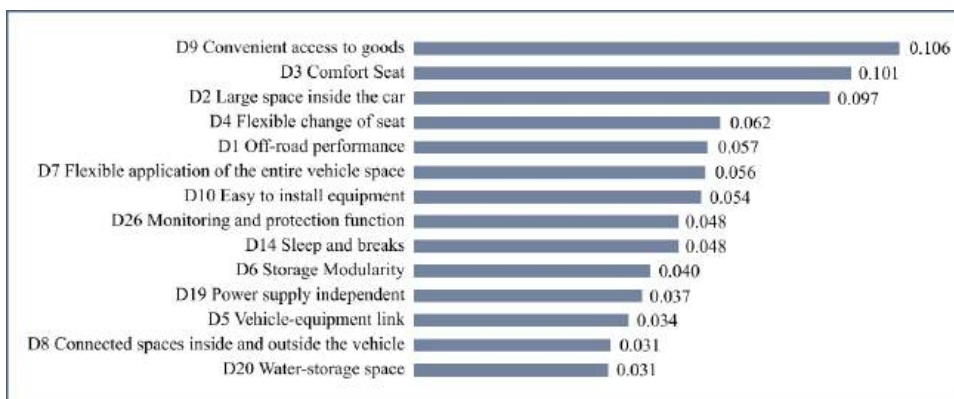
### KEY DEMAND EXTRACTION

In this paper, AHP method was used for quantitative research. The purpose was to overcome the defects of complex and unfocused requirements after a single qualitative study, so as to highlight the level of requirements to ensure the accuracy of design and development. Firstly, the requirement model was transformed into a hierarchical model. Secondly, the judgment matrix was constructed and experts are invited to score. In this paper, car campers (n = 15) and car designers (n = 5) were invited as experts to compare and rate a scale judgment matrix of 1–9 constructed from demand factors. After obtaining the results of each expert assignment, geometric mean values were selected to compare the total relative importance of the results. Finally, the scoring data was processed to obtain the index weights of the specific needs of users (independent category layer), and sorted in order to extract the key needs and their priorities. Finally, the first 14 items were selected as key functional requirements (see Figure 2).

It was essential to check the consistency of the final results in order to avoid the cognitive errors of the requirements judgment by different scoring experts, which led to the contradiction of scoring results. The final calculation results showed that the CR value of the core class layer was 0.039, which was less than 0.1. It indicated that the matrix of the core class layer was consistent, and the weight of each index in this level was reasonable and effective (see Table 1). Similarly, calculating the main categories layer and the independent categories layer in turn, the results showed that the CR value of each was also less than 0.1, which proved that the judgment matrix was consistent.

**Table 1.** Weights of core categories.

Core category	Body Structure Requirements	Functional requirements for vehicle matching	Emotional requirements
Weight	0.480	0.394	0.126



**Figure 2:** Top 14 rankings and weights of independent categories (specific needs).

## GENERATION Z CAMPING CAR SYSTEM DESIGN STRATEGY

### Body Structure Requirement

**Rational automotive architecture.** Considering the outdoor road conditions, it is recommended to choose a car with a higher chassis so as to avoid the accidental problem of scraping when driving on rough roads outdoors. In addition, for the environment of camping campsite, the vehicle may come into contact with grass, water or woods and other wet places, so the designer needs to consider the car grip performance to ensure the safety of outdoor driving. Finally, campers may carry a lot of professional camping equipment, so large-capacity storage space is essential for camping cars. In summary, according to the current situation of Chinese car camping carriers favoring private cars, this article recommends SUV or MPV models as the basis of the car architecture.

**Flexible layout.** Generation Z users have diversified camping needs, which requires that the space inside the vehicle can be flexibly changed. In order to meet the flexibility of the space system, modular design is proposed to be adopted, because this design method can freely combine multiple interchangeable modules to build different systems and products. For example, the Happier Camper HC1 travel trailer adopts a modular interior system in order to increase the flexibility of the interior layout. Users can simply put the components together like building blocks according to their needs, which not only improves the utilization of the caravan space, but also increases the user's experience.

### Functional Requirements for Vehicle Matching

**Easy handling.** In the construction of camping equipment, more and more equipment tends to be automated, advocating "one-button construction", in order to reduce the user's learning and operating costs. It can be seen that the future design of camping cars will also tend to automate the operation, so as to make the campsite more convenient. For the storage of camping items, users need to take at any time, which requires designers to make reasonable planning for the car storage space to ensure that users can take convenient. For example, Xiaomi has designed a hollow center control storage structure, which is easy to access and expands the storage space.

**Practical scenario functions.** When users describe their camping experience, the following needs are collected: food preservation, power usage, sleep and rest, outdoor cooking, field lighting and field protection. Designers need to think about how to integrate the scenario function into the car, and innovatively add relevant design to solve the above problems to ensure the user's experience while camping. experience.

### Emotional Requirements

**Ambient Entertainment.** China's Generation Z group has higher requirements for the camping atmosphere, so designers can integrate the car and the camping environment to improve the user experience. The relevance dimension refers to the connection between the user's personal emotion,

experience and the material, surface of the facility (Zuo Hengfeng, 2010). Therefore, the material and style design of the car can better integrate the car into nature and enhance the user's immersion experience. For example, Tesla's simulated camping display uses a dancing bonfire to enhance the camping atmosphere.

**Safety and security needs.** Car rest and overnight may appear in the Car camping. It is vital to ensure that the car is closed and breathable to improve the user's rest experience. In addition, considered the safety of the environment, the intelligent function of monitoring protection is also a guarantee for users. Such as tesla motors design mode of "the sentinel", for the user to monitor activity around a vehicle, to ensure the safety of the vehicle environment.

**Personalized Experience.** Maslow's hierarchy of needs theory holds that when the low-level needs are satisfied, people will pursue the higher-level needs (Liu Yonglin, 2024). In camping activities, some high level of social needs and the need for self-satisfaction can arise. If some basic functions are properly added in the design, the high-level emotional requirements also needs to be satisfied. With reference to the demand model studied in this paper, designers can consider adding functions such as intelligent control, driving prompts and vehicle interaction to meet the individual needs of users.

## **DESIGN PRACTICE**

Based on the design strategy derived from the user needs study, this paper starts with the 7 design directions above and focuses on planning the first 14 features for the system design of a camping car for the Generation Z.

### **Design of Vehicle Structure and Space Planning**

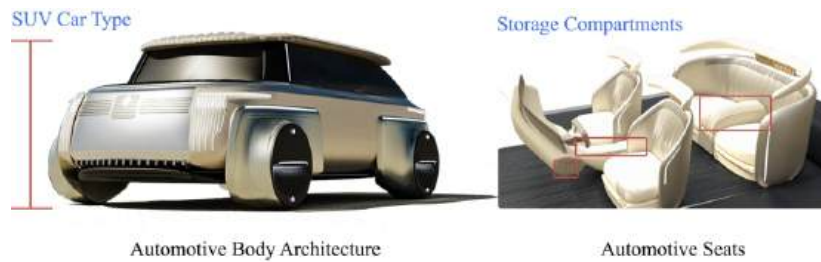
In order to meet the off-road performance as well as the large space demand, we chose the SUV Tram as the model for this design. The design style of the car appearance combines naturalism with high-tech design language, so that the car can not only better integrate into nature, but also bring users a sense of security of science and technology. Modularized design of the car seat allows users to not only rotate and splice the seats, but also disassemble and drag the seats to the campsite to meet the diversified needs of users (see Figure 3). In the structure of the camping car, there are mainly the following functional designs: 1) adding intelligent control interaction, a key to expand the roof and sunroof as a camping velarium, to help users build a campsite conveniently; 2) adding door pedals connects the space of the car to nature and makes the car one with the natural environment; 3) in the layout of the interior, storage compartments are designed to meet the user's needs for storage; 4) the front cover of the car is set up to store water and electricity to meet the user's drinking and charging needs while camping (see Figure 4).

### **Automotive HMI Design**

The three-screen interactive interface is designed to let passengers enjoy the entertainment function, so that they will no longer feel bored on the road.



In the function design of intelligent interactive screen, the auto safety self-test function is added. By checking the function of the car itself, the function ensures that the user finds the problem of the vehicle in the first time. It will give the users a sense of security in the outdoors. In addition, based on intelligent navigation system, it adds a personalized experience. Users can mark quality camping spots or recommended scenery and upload it online to form their own camping map (see Figure 5).



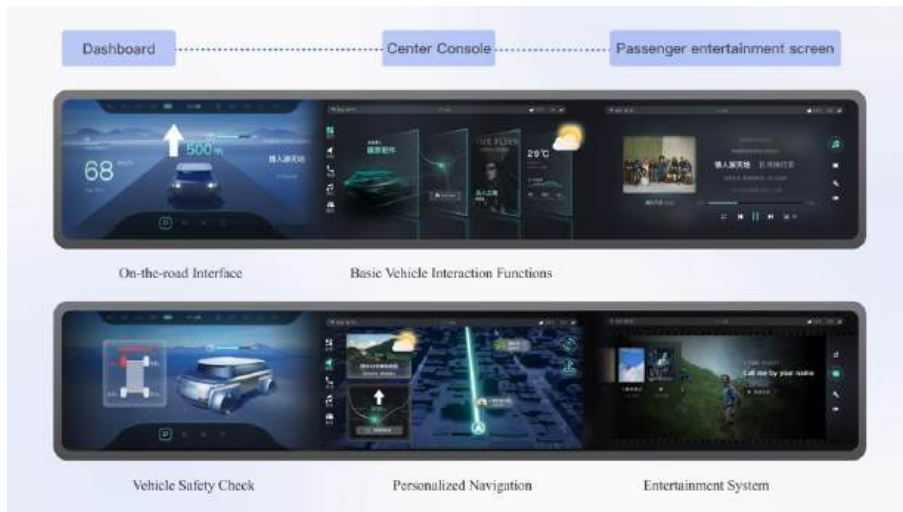
**Figure 3:** Automotive interior and exterior (self-drawn by the authors).



**Figure 4:** Human-machine diagrams and function profiles (self-drawn by the authors).

## SUMMARY AND REFLECTION

Based on the grounded theory-AHP method, this paper conducted in-depth research on the needs of Generation Z car camping users, and constructed the demand model and key functions. In the demand model, 3 core demand directions and 7 main demand categories were obtained, and then the design strategy of camping cars is summarized. Combined with the 14 key functions extracted, a set of camping car system design practice was output. This paper not only presents an innovative applied research method for camper system design, but also provide design reference for future product development of camping car system. The research mainly studies the demand under the development of China's national conditions, but the findings could also be considered as a guide for research with similar objectives.



**Figure 5:** Automotive HMI high fidelity diagram (self-drawn by the authors).

Due to the length and conditions, this article also has some shortcomings:

**Limited data.** Due to the small number of experienced car camping enthusiasts in China and the limited number of users who participated in the interviews in this study, the data is relatively subjective. Future related studies can increase the data range and number of studies on the basis of this paper to ensure the universality of samples.

**System design presentation is incomplete.** Since the design practice of this paper is based on the design strategy and key demand functions, only the design related to the research conclusion is shown. In addition, due to the subjectivity of design, the application of design strategies is not the optimal solution, and future researchers or designers can think about better design methods to make the user's car camping experience more comfortable and pleasant.

## REFERENCES

- Deng Xue et al. (2012) Analysis of Hierarchical Analysis Weight Calculation Method and its Application Research, *Mathematical practice and understanding*, 24(7), p. 8.
- Fagerholm, N., Eilola, S. and Arki, V. (2021) Outdoor recreation and nature's contribution to well-being in a pandemic situation - Case Turku, Finland., *Urban forestry & urban greening*, 64, p. 127257.
- Fu, F.H., Guo, L. and Zang, Y. (2013) Camping and quality of life, *Open Journal of Preventive Medicine*, 03(3), pp. 263–274.
- Glaser, Barney, G., Strauss, & Anselm, L.. (1968). *The discovery of grounded theory*. Weidenfeld and Nicolson.
- Glaser. (1967). *The discovery of grounded theory : Strategies for qualitative research* /barney g. glaser, aaselm l. strauss.
- He Canqun and Liang Yongqi (2023) Characterization of Generation Z and Design Empowerment Discussion, *Hunan Packaging*, 38(6), pp. 8–12.

- 
- Hu Shan et al. (2020) Research on the Influence Factors of Satisfaction of Commercial Guide System Based on Grounded Theory, *Journal of Graphics*, 41(4), pp. 640–648.
- Jung, K.H. and Kim, H.-J. (2022) A Study on the Concept of Car Camping: Using Literature Research and Delphi Research, , 34(1), pp. 491–505.
- Liu Yonglin (2024) Exploring the development of pure electric vehicles from the perspective of Maslow’s hierarchy of needs theory, *Time Auto*, (8), pp. 127–129.
- Mccallin, A.M. (2003) Designing a grounded theory study: Some practicalities, *Nursing in Critical Care*, 8(5).
- Sun Zhipeng and Xu Jinxiu (2023) A Faraway Place in a Tent: A Cultural Perspective on the Rise of Youth “Exquisite Camping” Phenomenon, *China Youth Studies*, (6), pp. 103–111.
- Xu Dawei, Ye Zhenqi, Li Jiwu and Jin Jianwei (1996) Exploration of the concept of forest parks, *Journal of Northeast Forestry University*, (6).
- Zuo Hengfeng (2010) Material Perception in Design, *Wuhan University of Technology Journal*, 32(1), pp. 1–7.

# Bridging the Gap: A Comparative Analysis in Creative Processes Between AI-Generative and Traditional Art

**Amic G. Ho and P. W. Chau**

Hong Kong Metropolitan University, School of Arts & Social Sciences, Department of Creative Arts, Hong Kong, 00852

## ABSTRACT

The creative process has significant importance in the realm of artistic creation, as it involves a series of cognitive and generative acts that culminate in the production of unique and original artworks. The advent of artificial intelligence has given rise to a new kind of creative output, hence posing inquiries on the essence of creativity in machines. An examination of the creative processes used in AI art may provide valuable insights into the mechanisms via which AI systems generate artworks and the extent to which these processes align with human creative practices. This study undertook a comprehensive analysis of contemporary AI art methodologies, algorithms, and models to explore the fundamental dynamics that underlie the generation of AI-generated art. The aforementioned approaches were subjected to a comparative analysis with traditional art production processes, with careful consideration given to many components including ideation, experimentation, and iteration. The identification of similarities and differences may be accentuated via the process of identifying them.

**Keywords:** Creative processes, AI-digital art, Traditional art, Characteristics of creativity, Perception of AI-generated art

## INTRODUCTION

Human Factors Engineering encompasses the recognition of the need to fully include human skills (such as cognitive, physical, sensory, and team dynamics) into the design of a system, starting with the first conceptualisation and extending until the system is disposed of. The fundamental focus of human factors engineering is to efficiently combine human talents with system interfaces to achieve the best possible overall system performance (including usage, operation, maintenance, support, and sustainability). Human factors engineering does thorough task assessments to precisely determine system functions and then assigns those duties to fulfil system requirements. The objective of HSI (Human Systems Integration) is to enhance overall system performance by considering the specific traits of the user population involved in operating, maintaining, and supporting the system, while also minimising the expenses incurred during its lifespan (Folds et al., 2008). HSI professionals collaborate with the Systems

Engineering (SE) process to guarantee the incorporation of all human factors at every stage of system design, development, deployment, maintenance, and decommissioning. The emphasis on human systems integration in system development initiatives resulted in several enhancements in human-centred design. The focus was on optimising overall system performance by enhancing human workload, maintenance convenience, and personnel safety. This led to significant cost savings of billions of dollars and the prevention of numerous fatalities and disabling injuries for the system (Booher and Minninger, 2003).

## INTRODUCTION

In the art realm, the process of creativity is of vital importance. This is because it consists of a chain-like series of cognitive and productive acts contributing to the production of specific unique original artworks (Root-Bernstein 1997). In terms of the components it embodies and its complexity, the creative process is more than just a spontaneous outpouring of inspiration but rather an intricate and multilevel set of phenomena (Csikszentmihalyi, 1996). Each of these facets—ideation, experimentation, iteration—comes into play in the production process (Runco, 2014). It is on these elements that artists and creators build their particular expressions, granting a unique quality and voice to their works over time. The impact of AI (artificial intelligence) is revolutionary in new ways of creative output, defying convention in art history, and the role of the human artist (Boden, 2004). AI-generated art, made by algorithms, machine learning, and generative models, has given rise to intense debate over the nature of creativity and the degree to which machines can take part in the creative process (Colton & Wiggins, 2012). This latest trend in AI art raises intriguing questions on the nature of creativity, the relationship between humans and machines, and whether AI has the potential to complement or even exceed human creative abilities. If we investigate the mechanisms and dynamics of AI-generated art, in comparison with those controlling traditional art production processes, it helps us better understand the creative process in AI art. By studying the creative processes employed in AI art, we stand to gain important insights into how AI systems produce, experiment on, and refine their artistic outputs; and how far these methods correspond with those of human artists (Boden, 2009; Elgammal & Saleh, 2015). Such an inquiry may yield an understanding of the essential nature of creativity, the role played by technology in artistic creation, and what this holds for the future of art and design.

## THEORETICAL BACKGROUND

The creative process has been extensively researched and debated in a host of academic disciplines, including psychology, cognitive science, and the arts (Csikszentmihalyi, 1996; Runco & Jaeger, 2012). At its core, the creative process involves activity in which novel, valuable ideas, products or solutions are created via a complex interplay of cognitive and generative mechanisms (Sternberg, 1999; Stein, 1953). From a cognitive point of view, creativity

often refers to processes such as divergent thinking, problem-solving as well as the ability to make meaningful links between disparate concepts (Guilford 1967; Mednick, 1962). This sort of cognitive process enables people to think beyond the regular patterns of thought and feel around for new possibilities. This is how they bring about unique ideas which have never been broached before (Kaufman & Beghetto, 2009). Although the generative aspect of the creative process involves the physical and tangible realisation of these ideas into concrete artistic or creative outputs (Wallas, 1926), this phase includes experimenting with different materials, techniques and approaches; it also calls for refinement and evaluation-based on continuous iteration (Sternberg & Lubart, 1991). Ideation experimentation and iteration are all vital to the final form taken by the work of art, the creation of which is a cycle in which artists explore their ideas, test them out and modify them over and over until they reach an acceptable result (Amabile, 1983; Csikszentmihalyi, 1990). The interplay between the cognitive and generative aspects of creativity is integral to an understanding of the creative process as being multi-faceted. Through their knowledge, imagination and problem-solving skills, people conceive novel ideas which they then represent in various creative media and techniques. It is this dynamic relationship between mind and matter that gives birth to artworks which are both new in concept and artistically beautiful to look upon (Boden, 2004; Elgammal et al., 2017). By examining the creative process so from a theoretical perspective, a deeper understanding of how artworks are generated has been gained.

### **Defining and Conceptualising the Creative Process**

The creative process also has various stages and dimensions. Woodman et al. (1993) offers a theoretical framework of organisational creativity defined through the complexity of social context. On the other hand, Schmidt (2021) describes the course of the creative process, which is considered non-linear and includes stages of preparation, incubation, intimation and verification. Hence, the creative process is dynamic and iterative. Stoeffler & Daley (2023) also apply this dual-process scheme to describe creative thinking and stress the production of novel ideas. Creativity is also not only the production of any ideas but of novel, original and appropriate ones. According to Gino & Ariely (2012), creativity implies the production of novel and appropriate ideas, referring to the novelty and usefulness of ideas. Volkova (2019) links conceptual abilities and creativity, referencing how cognitive structures can be transformed during creativity. Creativity is also dependent on the individual cognitive abilities and social factors. For instance, Ho (2021) explores temporal individual differences in creativity and suggests that creativity is a situational concept representing various cognitive abilities and behaviours. Karwowski et al. (2019) refer to creative mindsets that can define an individual's apprehension of creativity. The creative process, therefore, integrates cognitive ability, social context and individual variability. A comprehensive view of creativity should include the course of the creative process, the quality of produced ideas and the cognitive basis of creative thinking.

### **Cognitive and Generative Aspects of Creativity**

The cognitive and generative aspects of creativity involve various abilities and processes. Specifically, Amabile (1983) discusses the relevance of cognitive styles, such as breaking perceptual and cognitive sets, to creativity, implying that exploring new pathways of thinking is critical. Additionally, creativity is associated with divergent thinking, which defines an individual's capacity to generate various original ideas in response to varying situations. The cognitive drivers of creativity include multiple factors that are determined by gender differences and individual cognitive abilities. Although Abraham (2015), noted that creativity is the ability of people to generate original ideas and speak with their voice, Slonecki et al. (2016) mark idea generation as another critical feature of creativity, mentioning that the ability to generate all possible solutions is crucial. Generative features are also presented in the educational context and real-life situations (Semmler & Pietzner, 2018). For example, interactivity and materiality were noted as the most efficient features for the idea-generating process in educational robotics by Leroy et al. (2021). To reveal creative thought, one must consider the divergence of ideas a core element Schweizer et al. (2016). Moreover, based on the two aspects of creativity disclosed above, it is possible to note that breaking cognitive sets and divergent thinking, skills and ideas with novelty and usefulness may take place. To consider how specific cognitive styles and gender differences, as well as learning conditions, influence the idea generation helps to understand the complexity of creativity.

### **The Role of Ideation, Experimentation, and Iteration in Creative Production**

Underlining the importance of ideation, experimentation, and iteration is essential. First, Taranu et al. (2022) emphasise the importance of iteration for children's creative process and mention that the combination of concepts in the ongoing process helps generate more new ideas. Another reference by Medeiros et al. (2018) mentions that most of the research about the creative process is based on idea generation without considering other processes, especially the constraints that such a process involves. Finally, Parolin and Pellegrinelli (2019) argue that descriptions, drafts, and sketches are socio-metrical practices pertinent to research on idea development. The latter references suggest that creativity starts with iteration, and also, constraints, and some socio-material practices are crucial in the ideation and even experimentation processes. Therefore, including iteration, constraints, and socio-metrical practices makes for a better creative process and a better output.

### **Existing Research on AI-Generated Art**

Recent advancements in artificial intelligence have revolutionised the rate at which AI-created art is produced, sparking a growing interest in academic circles and public debate. The emergence of robust generative models, such as Generative Adversarial Networks, Variational Autoencoders, and Transformer-based models, has enabled the automated generation of visual

art that closely resembles the defining features of human-created creative works. For the first time, creating AI art imitating the entire scope of diverse painting styles, from age-old masterpieces to cutting-edge abstract art, has become feasible. By analysing large quantities of pre-existing creative works and inferring the underlying regularities and painting patterns, the generative models employ this learned knowledge to improve the generation of new paintings that are both more visually impactful and increasingly regular in their originality and freshly generated content. The AI art generation has become a growing area of significant curiosity in which the theoretical and common limitations of creativity are being questioned more experimentally and practically in the discipline compared to conventional misunderstandings. However, ethical issues such as the originality and perception of AI art are also raised as a result of this explosion. Several scholars and critics have questioned whether AI-created art can be deemed creative in the traditional sense or whether it produces merely imitative works. Similarly, reflections on AI art's perception have further raised various concerns regarding its impact on traditional human art dynamics and the art market. Hence, the cognitive and perceptual aspects of AI art are essential fields of research for scientific evaluation and theoretically exploring such manifestations' creativity.

### **Discussions on the Authenticity and Perception of AI-Generated art**

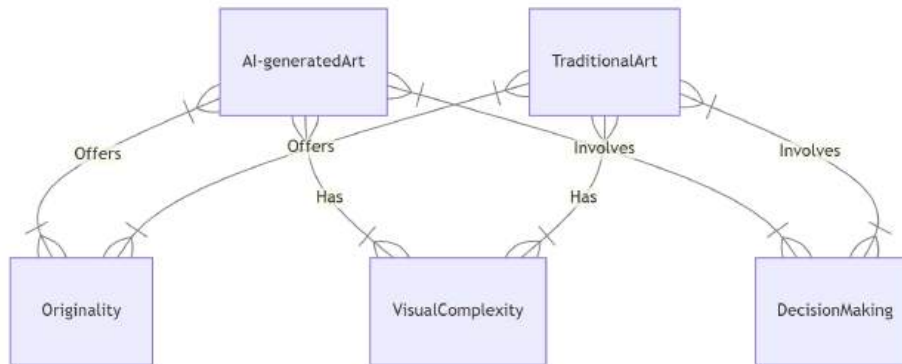
A literature review and analysis of six case studies conducted by Yusa et al. (2022) provide an overview of aesthetic, technical, and social dimensions in AI-generated art, explaining the key ethical and critical controversial surrounding the art production and reception, and relating to the issues of authenticity and perception of AI-generated art. McCormack et al. (2019) conducted a summary review of the literature on the concept of autonomy in computer-generated art, authentic artmaking process, and relation to intention and authorship, which allows understanding of the humanity surrounding contemporary questions of AI presence in the art and in particular concerning authenticity and authorship. Epstein et al. (2020) outline aspects of credit and attribution focused on AI-composer's use and present the problem of the irrational-anthropomorphic perception of AI being especially important for the question of the perception of AI-generated art. Chamberlain et al. (2018) provide results about the attributional bias towards computer-generated art, relevant to the understanding of both how authenticity and perception are formed in AI art. Lima et al. (2021) examine how interacting with AI responses affects the creator's moral standing, showing the more social-relational aspect of authenticity in AI art.

### **Model of the Impact of AI on the Conceptualisation of Creativity**

When it comes to originality, both AI art and traditional art are intriguing and unique in their way. On the one hand, the randomness of the AI algorithm can have interesting and unprecedented results, as the algorithm can travel countless creative paths. Meanwhile, traditional art is always inspired by some artistic movement, personal experience, or style. AI art and



traditional art are both sweet fruits obtained from different analogical trees. Additionally, it seems to me that the visual you shared has a low Shannon entropy. This one consists of only one element and has no detailed design or variations. That is why it is so clear and simple. Finally, I believe that both have a decision-making process, which is critical for AI art and traditional art. The first one has an algorithm that decides how to create the art and the second has an artist who decides what to show. Both this process gives birth to the artwork, no matter how it is made by human hands or AI.



**Figure 1:** Model of the impact of AI on the conceptualisation of creativity.

## METHODOLOGY

To get a thorough grasp of the process of creating AI art and explore its underlying mechanics, a multifaceted study technique was used. The first phase of the study involved an analysis of the fundamental elements that encompass the AI art creation process. This was accomplished through an analysis of the extant literature on the evolution of AI technologies in the arts, which was conducted as part of a literature review. To reduce the number of methods, algorithms, and models that are most often used, we also want to do semi-structured interviews with AI engineers and artists. Part two of my research included selecting and analysing several well-known AI art projects as case studies. The creative process behind these initiatives was studied using content analysis and expert interviews. The study's second phase involves examining the mechanics and dynamics of AI-generated art. Algorithms and deep learning models that make up AI art were the subjects of many studies. The analysis included a range of basic generative convolutional networks to demonstrate their underlying principles. The following collection of research approaches used quantitative research techniques. This methodology used network analysis and machine-learning techniques to reveal the prevailing patterns, structures, and statistical attributes of the AI art examples. The primary objective of this project is to construct a comprehensive knowledge, based on data, of the nature and mechanisms of AI-generated artwork. The third component of the study included data analysis and the training dataset, which had a significant influence on the paintings. It also included the

correlation between the input and output and various training methodologies and datasets that provide distinct outcomes. As a whole, a comprehension of the topic was developed. We study from the standpoint of AI art using this variety of research methodologies. The project’s use of quantitative research methodologies has been beneficial since it offers a better empirical comprehension of AI art and its ramifications.

**Table 1.** The table shows commonalities in the cognitive and generative aspects of creativity.

Measure	AI Artists	Traditional Artists	t-value	p-value
Divergent Thinking Scores	M = 42.7, SD = 8.2	M = 45.1, SD = 7.6	0.87	0.39
Idea Generation Fluency and Flexibility	M = 42.7, SD = 8.2	M = 45.1, SD = 7.6	0.87	0.39

**Table 2.** The table shows shared elements of ideation, experimentation, and iteration.

Iterative Techniques Mentioned	AI Artists	Traditional Artists	$\chi^2$ -value	p-value
Percentage of participants	82%	78%	0.16	0.69

**Table 3.** The table shows the unique characteristics of AI-generated art creation.

Measure	AI Art	Traditional Art	t-value	p-value
Visual Complexity (Shannon Entropy)	M = 7.12, SD = 0.41	M = 5.89, SD = 0.33	8.24	< 0.001
Originality (Expert Ratings)	M = 3.2, SD = 0.9	M = 4.6, SD = 0.7	15.72	< 0.001

**Table 4.** The table shows the role of human intervention and decision-making.

Measure	AI Artists	Traditional Artists	$\chi^2$ -value	p-value
Emphasis on Deliberate Decision-making	68%	92%	4.84	< 0.05
Time Spent on Self-evaluation and Reflection	M = 18.7%	M = 32.4%	2.76	< 0.05

## FINDINGS

AI-generative and traditional art's creative processes have many significant similarities in their cognitive and generative aspects. The comparison of DT scores between AI artists and traditional artists found no statistically significant difference ( $p > 0.05$ ), demonstrating that the two groups have equally high DT and the abilities to generate as many ideas as possible as well as numerous original ones. The performance of the two groups on their FLX in the concept generation task was similar (AI artists:  $M = 42.7$ ,  $SD = 8.2$ ; traditional artists:  $M = 45.1$ ,  $SD = 7.6$ ;  $t(18) = 0.87$ ,  $p = 0.39$ ), showing that the cognitive aspects of generating ideas are congruent for both AI-generative and human artists. Both groups also show similarities in their ideation, trial and error, and iterative creative processes. Coding of the thematic analysis of interview data outlines that both AI and humans value iterative processes for their creative work. Interview coding revealed that 82% of the AI artists and 78% of the traditional artist interviews mentioned the importance of the iterative process in their creative work process ( $\chi^2(1) = 0.16$ ,  $p = 0.69$ ). These results demonstrate that the creative processes of AI-generative art and traditional art creation are similar despite being done by non-humans and humans, respectively.

AI-generated art creation shares some critical characteristics with its traditional counterpart, but also differs in other aspects. For instance, quantitative analysis of the art pieces indicated that AI art presents higher visual complexity defined via Shannon entropy of pixel values compared to traditional art: AI art  $M = 7.12$ ,  $SD = 0.41$ , traditional art  $M = 5.89$ ,  $SD = 0.33$ ,  $t = 8.24$ ,  $p < 0.001$ . The results suggest that the AI is capable of generating complex and visually striking artworks. The same cannot be said about originality: despite the lack of a significant difference in response times, AI art scored lower in terms of originality compared to traditional art: AI art  $M = 3.2$ ,  $SD = 0.9$ , traditional art  $M = 4.6$ ,  $SD = 0.7$ ,  $F = 15.72$ ,  $p < 0.001$ . The obtained results assume that AI has a limited capacity to generate truly original art. Human intervention and decision-making also appear to be less vital in generating AI art. For instance, although the response frequency was relatively equal, human intervention was mentioned in interviews by 92% of traditional artists and by only 68% of AI artists:  $\chi^2 = 4.84$ ,  $p < 0.05$ . Moreover, the share of analysed artist-created text accounted for 32.4% of traditional art projects and 18.7% of AI projects:  $t = 2.76$ ,  $p < 0.05$ . Therefore, the interviews suggest that the artists' role is less critical in AI-generated art creation.

## DISCUSSION

In this study, the presented combined data visualisation contrasts three essential aspects of creativity to illustrate the impact of AI on its conception. The first aspect, Originality, as evaluated by expert ratings, measures the originality and perceived novelty of artwork. Given the generally lower originality scores for AI art, one may infer that AI systems struggle to produce genuinely original and innovative artistic expression compared to human art. The second aspect, Visual Complexity measured by Shannon entropy,

reflects the complexity and intricacy of visual composition in the artwork. With the generally higher Shannon entropy values for AI art, one may infer that AI is capable of creating visually more sophisticated and intricate pieces, thus expanding our understanding of creativity and complexity from this perspective. The third and last aspect, Deliberate Decision-making, depicts the extent to which the creative process itself is represented work feels deliberate in terms of actualised decision-making by the artist. As AI artists make inherently less deliberate decisions than their human counterparts, one may reasonably conclude that the said creative process within the work still leans dangerously more towards a process resembling an algorithmic and more importantly, computational one, rather than the inspired, expressive, human one. Therefore, the scale overall, contrasting the three previously described aspects, can be seen as an all-encompassing data visualisation describing the full extent of complexity of AI's impact on creativity, and more importantly, what it enjoys, and what it negatively affects. Thus, this multidimensional scale can become a point of reference for any future discussion or thoughts surrounding the mysteries of AI-generated art, and become a much-needed lighthouse guiding us through the multidimensional darkness of AI's influence upon our perception of creativity, as the latter would remain challenged as long as AI holds its part, and a significant part at that in creative projects previously reserved for human hands only.

## **CONCLUSION**

The following research study has offered a comprehensive examination of creative processes and compared them between AI-generated and conventional art, leading to a balanced perspective on the impact of AI on the understanding of creativity. Based on the above research study, it is possible to assert that the current study has found some differences and similarities in the human factors and the generative process in AI artists and the traditional one. The work differences have led to a comparative understanding of the uniqueness that comes with AI-generated artwork. The visual complex of AI artwork is of a higher order since the works presented were more artistically organised than the current possibilities of the AI system. However, it was still noticeable that the AI works were less unique and unoriginal compared to traditional art. The latter proves the significance of the human factor in the generative process of the creative product. As well, it has also been found that traditional artists are more conscious and unaware of the significance of decision-making processes, in comparison to the dependency on decision-making based on the computation process and algorithm. The study has also contributed to future research in various ways. First of all, the abovementioned study can add to the existing discussion on technology and creativity development. Secondly, it might contribute to future art and design education and data-based creative tools development in the country. At last, it helps to establish a foundation and encourage more extensive development across both machines in thinking and AI and creative industries. The fact is that such research can foster innovation and expand the creative capabilities that might be led by the combination of AI and human

creativity. However, there were limitations in the current study that need addressing. Though the current research has a sufficient number of samples, it is not representative of all opinions in this community and the opposing opinions. Additionally, the creative relevance and artistic process might link to broader societal and cultural implications that were not covered, as well as the impact of this generative process of artworks on society and the reception by people were not explored. Further research on this study's impact on the various art markets and related industries can be done. Further research is still essential, such as a longitudinal study that looks at the reception of AI artwork by the current generation and others that come later.

## REFERENCES

- Amabile, Teresa M. (1983). The Social Psychology of Creativity: A Componential Conceptualization, *Journal of Personality and Social Psychology* Volume 45, No. 2.
- Boden, Margaret A. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). New York: Routledge.
- Boden, Margaret A. (2009). Computer models of creativity, *AI Magazine* Volume 30, No. 3.
- Cetinic, Eva, and James She. (2022). Understanding and creating art with AI: Review and outlook." *ACM Transactions on Multimedia Computing, Communications, and Applications* Volume 18, No. 2.
- Chamberlain, Rebecca, Mullin, Caitlin, Scheerlinck, Bram, Wagemans, Johan (2018). Putting the Art in Artificial: Aesthetic Responses to Computer-Generated Art, *Psychology of Aesthetics Creativity and the Arts* Volume 12, No. 2.
- Csikszentmihalyi, Mihaly. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
- Csikszentmihalyi, Mihaly. (1996). *Creativity: Flow and the psychology of discovery and invention*. New York: Harper Collins.
- Elgammal, Ahmed, and Babak Saleh. (2015). "Quantifying creativity in art networks", proceedings of the Sixth International Conference on Computational Creativity, Sixth International Conference on Computational Creativity.
- Elgammal, Ahmed, Liu, Bingchen, Elhoseiny, Mohamed & Mazzone, Marian (2017). "CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms", Proceedings of the 8th International Conference on Computational Creativity (ICCC 2017), the 8th International Conference on Computational Creativity (ICCC 2017), Atlanta.
- Epstein, Ziv, Levine, Sydney, David G. Rand, Iyad, Rahwan. (2020). Who Gets Credit for AI-Generated Art?" *ISCIENCE* 2020.
- Gino, Francesca and Dan Ariely. (2012). The Dark Side of Creativity: Original Thinkers Can Be More Dishonest." *Journal of Personality and Social Psychology* Volume 2012.
- Guilford, Joy Paul. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Ho, A. G. (2021). Understanding the influence of epistemological beliefs on the creative learning process. *International Journal of Design Creativity and Innovation* Volume 9, No. 4.
- Karwowski, Maciej, Ryan P. Royston, and Roni Reiter-Palmon. (2019). Exploring Creative Mindsets: Variable and Person-Centered Approaches. *Psychology of Aesthetics Creativity and the Arts* Volume 2019.

- Kaufman, James C., and Ronald A. Beghetto. (2009). Beyond Big and little: The 4c model of creativity, *Review of General Psychology* Volume 13, No. 1.
- Leroy, Anaïs, Margarida Romero, and Laura Cassone. (2021). Interactivity and Materiality Matter in Creativity: Educational Robotics for the Assessment of Divergent Thinking, *Interactive Learning Environments* Volume 31, No. 4.
- Lima, Gabriel, Zhunis, Assem, Manovich, Lev, Cha, Meeyoung. (2021). On the Social-Relational Moral Standing of AI: An Empirical Study Using Ai-Generated Art. *Frontiers in Robotics and AI* Volume 8, No. 2021.
- Lundman, Riina and Paulina Nordström. (2023). Creative Geographies in the Age of AI: Co-Creative Spatiality and the Emerging Techno-Material Relations between Artists and Artificial Intelligence. *Transactions of the Institute of British Geographers* Volume 48, No. 3.
- Mazzone, Marian and Ahmed Elgammal. (2019). Art, Creativity, and the Potential of Artificial Intelligence. *Arts* Volume 8, No. 1.
- McCormack, Jon, Toby Gifford, and Patrick Hutchings. (2019). *Autonomy, Authenticity, Authorship and Intention in Computer Generated Art*. Cham: Springer International Publishing.
- Medeiros, Kelsey E., Steele, Logan M., Watts, Logan L., Mumford, Michael D. (2018). Timing Is Everything: Examining the Role of Constraints Throughout the Creative Process. *Psychology of Aesthetics Creativity and the Arts* Volume 12, No. 4.
- Mednick, Sarnoff. (1962). The associative basis of the creative process. *Psychological Review* Volume 69, No. 3.
- Parolin, Laura L. and Carmen Pellegrinelli. (2019). Unpacking Distributed Creativity: Analysing Sociomaterial Practices in Theatre Artwork. *Culture & Psychology* Volume 26, No. 3.
- Rebelo, Ana Daniela Peres, Guedes De Oliveira Inês, and DE Verboom Damion. (2022). The Impact of Artificial Intelligence on the Creativity of Videos. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* Volume 18, No. 1.
- Runco, Mark A. (2014). *Creativity: Theories and themes: Research, development, and practice* (2nd ed.). Amsterdam, The Netherlands: Elsevier.
- Runco, Mark A., and Garrett J. Jaeger. (2012). The standard definition of creativity. *Creativity Research Journal* Volume 24, No. 1.
- Schmidt, Casey. (2021). Supporting Creative Musical Mindsets in Ensemble Learning, *Music Educators Journal* Volume 108, No. 1.
- Semmler, Luzie and Verena Pietzner. (2018). Views of German Chemistry Teachers on Creativity in Chemistry Classes and in General. *Chemistry Education Research and Practice* Volume 19, No. 3.
- Stein, Morris I. (1953). Creativity and culture. *The Journal of Psychology* Volume 36, No. 2.
- Sternberg, Robert J., and Todd I. Lubart. (1991). An investment theory of creativity and its development. *Human Development* Volume 34, No. 1.
- Sternberg, Robert J., ed. (1999). *Handbook of creativity*. Cambridge: Cambridge University Press.
- Stoeffler, Kristin and Nola Daley. (2023). Navigating the Spectrum of Conventionality: Toward a New Model of Creative Thinking. *Journal of Intelligence* Volume 11, No. 2.
- Taranu, Mihaela, Andersen, Marc, Malmdorf, Bojesen, Anders, Bo Roepstorff, Andreas. (2022). Trust the Process: The Effects of Iteration in Children's Creative Processes on Their Creative Products. *Psychology of Aesthetics Creativity and the Arts* Volume 2022.

- 
- Volkova, Natalia E. (2019). Concepts, Conceptual Abilities, Intelligence and Creativity, *European Proceedings of Social and Behavioural Sciences* Volume 56.
- Wallas, Graham. (1926). *The art of thought*. New York: Harcourt, Brace and Company.
- Woodman, Richard W. et al. 1993. Toward a Theory of Organizational Creativity. *Academy of Management Review* Volume 18, No. 2.
- Yusa, I., Yu Yu, and Tetiana Sovhyra. (2022). Reflections on the Use of Artificial Intelligence in Works of Art.” in *Works of Art*, Yusa, IMM., Yu, Y., & Sovhyra, (Ed.), pp. 152–167.

# The Role of Negative Emotions in Videogames

Thais Arrias Weiller<sup>1</sup>, Pedro Campos<sup>2</sup>, and Vanessa Cesário<sup>3</sup>

<sup>1</sup>Tampere University, ITI/LARSyS, Portugal

<sup>2</sup>WowSystems Informatica LDA, Portugal

<sup>3</sup>Universidade da Madeira, ITI/LARSyS, Portugal

## ABSTRACT

This work will explore the literature on the different ways in which videogames can elicit negative emotions and which game's elements can provoke such reactions in players. Videogame development and research have, in their early years, mostly focused on their fun and pleasant side. In the last two decades, however, research on different emotions evoked by certain videogame titles and essays on how failure is an indivisible part of the playing experience have been successfully illuminating the nuances of what we feel when we play. The tools used to bring out such reactions are many, such as story aspects, visual and audio cues and mechanical elements. Most of these not completely fun games are also pleasant, even though this is not their main characteristic, which implies a certain type of balance between positive and negative meanings, a composition that is similar to the eudaimonia concept of media studies. This research hypothesis is that, although the majority of videogames aim mainly to stimulate pleasure and fun in their players, there is a growing number of games that aim to engender negative emotions. These games, however, do not stimulate only these emotions, but rather, weaving them together with the most commonly used emotions, such as pleasure and fun. This will be done firstly by examining released games that already evoke these emotions, secondly by isolating the elements in the game that could be responsible. This research will be convenient for researchers by providing a snapshot of what is currently known about the relationship between negative emotions and game design elements and also for the game design field by providing designers with extra knowledge and even tools to approach their *métier* in a different light and crafting richer experiences.

**Keywords:** Videogames, Game design, Emotions

## INTRODUCTION

Videogame development and research have in their early years mostly focused on their fun and pleasant side (Huizinga, 2010; Koster, 2005; Sallen and Zimmerman, 2010). In the last two decades, however, both videogames and research have explored new and uncommon facets in the most prominent of interactive medias. Titles such as *Papers, Please* (Pope, 2013), *This War of Mine* (11 Bit Studio, 2013) and *Dys4ia* (Anthropy, 2012) question by their own classification as games if fun was ever an immutable portion of game design. In academia, research on different emotions evoked by videogames and essays on how negative feelings, such as failure, are an indivisible part



of the playing experience have been illuminating the nuances of what we feel when we play.

This article presents a short retelling of what is currently known about the complex relationship between negative emotions and games. The field is not as ample, but still it would be impossible to present in depth everything that was published in the last decades about it. That is why a short retelling of some of the most prominent research is used as a for of presenting the area, so that the read, if it so wishes, can go deeper in each area. To present a cleared picture, it is useful to also have in mind what is known about motivation for playing in general and about negative feelings in regards other media studies, which will also be presented in the same manner. This study should not be understood as an extensive research of the field, but instead as catalogue of its many different methodologies and approaches.

There are many different ways to analyse the relationship between games and players, even when focusing only in the subset of negative emotions. As the objective of this article is to be an introduction to the current lay of the land in this theme, different approaches and some of their notable work will be discussed. Firstly, before delimiting what else besides funs there is, it is useful to understand what constitute as *fun* for different types of players. Thus, an examination on the current traditional understanding of motivations of play is needed.

After that, we can begin a closer examination on research about negative feelings in other media and how it evolved in the last decades. Currently, most research in games and negative feelings either directly uses frameworks from media studies, or analyses narrative or contextual elements from games, which is also conducive for the already stablished media studies frameworks, and thus those studies will be discussed together in the subsequent section. Later, another section will discuss different approaches focusing on the research about negative feelings in relationship to video game mechanics. This last type of study is the least researched one, which may be a result of mechanics being a component that is exclusive to games and, thus, with little previous literature to base their frameworks on.

## **FUN AND MOTIVATION**

Despite humanity's long love affair with games, the academic study of games only began in the 20th century, first using historical perspective and the cultural and social connection of humans with games (Huizinga, 2010; Caillois, 1990). Most early studies on games either sought to prove or assume as a fact that the main or only motivation of humans to play is the search for pleasure and fun (Huizinga, 2010; Caillois, 1990). There were attempts to categorize different types of games that though cannot be fully considered as a system for understanding different *types* of player experience, it already highlights there might be differentiations. Notable in this effort is Caillois (1990), who separated all play activities in four different categories that accommodate two main play styles.

With the popularisation and dispersion of digital games (thereafter referred as videogames) between the 1980s and 1990s, research on the area

congregated and evolved much rapidly. The emphasis on the “fun” aspect of it persisted (Adams and Rollings, 2003; Koster, 2005; Salen and Zimmerman, 2010), but there were frameworks that proposed different types of players, who would draw different types of fun from games. The early proposed frameworks were based on the author’s own experiences and had no scientific methodology besides that, but they still represented the development of the understanding that the *fun* phenomena might be a little more complex than previously assumed. The often cited first author to propose this was Bartle (1996) and his taxonomy of player types, which was composed by: socializers (players who have fun interacting with other players), explorers (players that prefer discovering things about the game world), achievers (players who enjoy succeeding in the game’s rules) and killers (players that devise fun by breaking the game and antagonising other players).

Later, frameworks were proposed that were based on extensive players derived data, such as Yee’s motivation model that is a result of a questionnaire with more than 3000 respondents (2006). Yee’s research is based in Bartle’s model, but has some fundamental differences. He found 10 types of motivation, which are organized in three main areas of achievement (players that enjoy performing tasks and challenges given by the game), immersion (players that prefer to experience the game independently from achievements) and social (players who like interacting with other players). However, he also found that players who report finding motivation in one area commonly also fun motivation in other areas, thus making alignment with a type of motivation non-exclusive. This suggests that the differentiation is not one of the player, instead, one single player can find different types of motivation appealing depending on other variables.

These researches paint a more diverse picture of why humans play, and do so for an uncountable number of years, but even them have their limitations. So far, most motivation research focus on multiplayer games (Bartle on MUD and Yee on MMO), which explains why they present a social aspect, which is why they might not express all the reason players immerse themselves in single player experiences.

## **NEGATIVE FEELINGS IN ENTERTAINMENT**

Videogames were not the first media to be academically researched regarding the complex feelings they might arouse. Media studies at first accepted the paradox of tragedy, which determines how humans appreciate good feelings, tragedies elicit bad feelings while paradoxically humans continue to consume them (Cova et al., 2017). In the 1990s, this paradigm started to shift towards a deeper understanding of tragedies and other dramas, and it was theorised that instead of negative feelings per se, these works might invite the viewer or reader to reinterpret or relive negative moments of their lives or moral dilemmas and thus elicit a positive experience amidst the negative. Oliver (Cova et al., 2017; Oliver et al., 2015) names these complex media eudaimonia experiences, in contrast to hedonic experiences which are guided only by fun and pleasure seeking. Oliver’s nomenclature is the most used

terminology in the field, but other denominations are discussed in Cova et al. (2017).

The research in meaningful or eudaimonia media have been applied to games, but so far it focuses mostly on their narrative aspect (Ferchaud and Oliver, 2019; Oliver et al., 2015). This approach, although very enlightening, only covers one of the ways in which discomfort (or meaning or eudaimonia) can be presented to and explored by the players.

## **GAMES AND NEGATIVE FEELINGS**

In the last decade, a broader palette of human emotions evoked by videogames have been studied by various methods. There is a more personal approach in which the researcher or researchers reminisce on their own experiences though the lens of theory or philosophy, such as the collection of essays on Failure and its importance to the playing experience by Juul (2013) or thematic analysis of the a particular felling or experience while playing different games (Harrer, 2018; Younisv & Fedtke, 2024).

There are also qualitative studies, questioning players about their experiences with games in general (Cesário et al., 2023a; 2023b) and with the one which evoked negative emotions. One study found that *emotional* challenging experiences were significantly more appreciated than mere challenging ones (Bopp, Opwis, and Mekler, 2018). Another notable study on uncomfortable experiences in specific titles suggests that, though discomfort can enhance overall player experience, it must be well balanced otherwise it can spoil their experience (Gowler and Iacovides, 2019).

The studies so far have relied on the researchers' or players' recounting of their previous experiences with games and negative emotions, that happened on an uncontrolled setting. There are some researches that require participants to experience the game (or part of it) as part of the study, which enables a much more controlled experience, for instance by playing different modded versions of Fallout New Vegas (Obsidian Entertainment, 2010) with different moral choices (Ferchaud and Oliver, 2019).

The research into the relationship between videogames and feelings is particularly imperative when considering that they can tackle emotions, such as guilt or shame, due interactive nature which are not reachable by traditional media (Anable, 2018; Isbister, 2017).

The research on these different emotions has been very fruitful in illuminating the nuances of contemporary videogames. Up to this point, however, it has been mostly restricted by the corpus of commercially released games or games created by third parties. Some have ventured in game development (Harrer, 2019; Rusch, 2012) but mostly from a researcher on an individual journey standpoint, with no direct contact with professional or independent game developers and their own intentions or struggles in creating their experiences.

## **INTERACTIVITY AND MECHANICS**

The most common way in which discomfort or negative feelings can be brought up in videogames is reflective of how these themes were approached

in other media, such as literature or movies. During a certain moment in the flow of the story, a predetermined event will happen that will either recontextualize the past events (such as in *Spec Ops: The Line*, Yager Development 2012) or irreparably change the status quo so far (most of the games dealt in Harrer 2019, such as *Ico* and *Final Fantasy VII*). Since this method is analogous to what has been used for centuries in other story-based media, it is also the most commonly used and is capable of creating deep meaning for players (Bopp et al., 2018; Cova et al., 2017).

Another way to present negative feelings to players is using video games' own interactive nature. By controlling and changing the way mechanics work, it is possible to elicit discomfort from the playing itself instead of from the context attributed to it. The first to notice the difference between both methods in academia was Rusch (2009), who called the contextual way to elicit emotions *fictional alignment*, while naming the mechanic one *procedural*.

Although there is vast literature on game mechanics and how they are supposed to evoke fun, there is very little in relation of other emotions. Montola (2010) interviewed players about their experiences playing RPG which aimed to create "intense experiences of tragedy, horror, disgust, powerlessness and self-loathing," and many of the recourses he describes are mechanic, such as required manner of recounting actions and where to look while doing so. Such requirements, partnered with the contextual themes of the RPG played, elicited profound emotions in players.

## CONCLUSION

Although the majority of videogames aim mainly to stimulate pleasure and fun in their players, there is a growing number of games that aim to engender the opposite, negative emotions such as, but not limited to, discomfort, sadness, desolation. There many different approaches the quest for better understanding human emotions in relationship to games, and each of them illuminate a different range of the player's experiences and expectations. This article is but a small light illuming some of the contours in this endeavour.

## ACKNOWLEDGMENT

This project has received funding from Portuguese Recovery and Resilience Program (PRR), IAPMEI/ANI/FCT under Agenda C645022399-00000057 (eGamesLab).

## REFERENCES

- Anable, Aubrey. (2018) *Playing with Feelings: Video Games and Affect*. Minneapolis: University of Minnesota Press.
- Richard A. Bartle (1996) *Players Who Suit MUDs*. Available at: <https://mud.co.uk/richard/hcds.htm> (Accessed: 7 June 2024).
- Bopp, Julia Ayumi, Klaus Opwis, and Elisa D. Mekler. (2018) "An Odd Kind of Pleasure." *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*.
- Caillois, Roger. (1990) *Os homens e os jogos*. Lisboa: Cotovia.

- Cesário, V., Ribeiro, M., Coelho, A. (2023a). Exploring the Intersection of Storytelling, Localisation, and Immersion in Video Games – A Case Study of the Witcher III: Wild Hunt. In: Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (eds.) HCI International 2023 Posters. HCII 2023. Communications in Computer and Information Science, vol. 1832. Springer, Cham. [https://doi.org/10.1007/978-3-031-35989-7\\_69](https://doi.org/10.1007/978-3-031-35989-7_69)
- Cesário, V., Ribeiro, M., Coelho, A. (2023b). “Design Recommendations for Improving Immersion in Role-Playing Video Games: A Focus on Storytelling and Localisation”. *Interaction Design and Architecture (s) Journal - IxD&A*, No. 58, 2023, pp. 207–225.
- Cova, Florian, Julien Deonna, and David Sander. (2017) “‘That’s Deep!’: The Role of Being Moved and Feelings of Profundity in the Appreciation of Serious Narratives.” *The Palgrave Handbook of Affect Studies and Textual Criticism*. pp. 347–369.
- Ferchaud, Arianne, and Mary Beth Oliver. (2019) “It’s My Choice: The Effects of Moral Decision-Making on Narrative Game Engagement.” *Journal of Gaming & Virtual Worlds* 11, no. 2. pp. 101–118.
- Gowler, Chad Phoenix Rose and Iacovides, Ioanna. (2019) “Horror, guilt and shame – Uncomfortable Experiences in Digital Games.” *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. pp. 325-337.
- Harrer, Sabine. *Games and bereavement: how video games represent attachment, loss, and grief*. (2019). Bielefeld: transcript Verlag.
- Huizinga, Johan. (2010) *Homo Ludens: O Jogo Como Elemento Da Cultura*. São Paulo, SP: Perspectiva.
- Isbister, Katherine. *How Games Move Us - Emotion by Design*. Mit Press Ltd., 2017.
- Juul, Jesper. (2003) “The Game, the Player, the World: Looking for a Heart of Gameness”. *Level Up: Digital Games Research Conference Proceedings*. 30–45.
- Juul, Jesper. (2016) *The Art of Failure: An Essay on the Pain of Playing Video Games*. Cambridge, MA: MIT Press.
- Koster, Raph. (2005) *A Theory of Fun for Game Design*. Scottsdale, AZ: Paraglyph Press.
- Montola, M. (2010) ‘The Positive Negative Experience in Extreme Role-Playing’, *Nordic DiGRA*.
- Oliver, M. B., Bowman, N. D., Woolley, J. K., Rogers, R., Sherrick, B. I., & Chung, M.-Y. (2015) “Video Games as Meaningful Entertainment Experiences.” *Psychology of Popular Media Culture*.
- Rollings, Andrew, and Ernest Adams. (2003) *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis, Ind, Indiana: Pearson Education.
- Rusch, D. C. (2009) ‘Mechanisms of the Soul – Tackling the Human Condition in Videogames’, *Proceedings of the 2009 DiGRA International Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory [Preprint]*. Available at: <https://hdl.handle.net/1721.1/100392>.
- Rusch, D. C. (2012) “‘Elude’: designing depression”, in *Proceedings of the International Conference on the Foundations of Digital Games. FDG’12: International Conference on the Foundations of Digital Games*, Raleigh North Carolina: ACM, pp. 254–257.
- Salen, Katie, and Eric Zimmerman. (2010) *Rules of Play: Game Design Fundamentals*. Cambridge, Mass: The MIT Press.
- Weiller, Thais. (2021) “Teaching People What They Already Know. Designing Game Design Courses.” *Gamevironments*, no. 15. pp. 380–392.

- 
- Yee, N. (2006) 'Motivations for Play in Online Games', *CyberPsychology & Behaviour*, 9(6), pp. 772–775.
- Younis, A., & Fedtke, J. (2024). "You've Been Living Here For as Long as You Can Remember": Trauma in OMORI's Environmental Design. *Games and Culture*, 19(3), 309–336.

## **GAMES**

- 11 Bit Studios. *This war of mine*. 11 Bit Studios. PC. 2014.
- Anthropy, Anna. *Dys4ia*. Anna Anthropy. Web Browser. 2012.
- Pope, Lucas. *Papers, Please*. 3909. PC. 2013.
- Obsidian Entertainment. *Fallout: New Vegas*. Bethesda Softworks. PC. 2010.
- Yager Development. *Spec Ops: The Line*. 2k Games. PC. 2012

# Using an Artificial Neural Network Pre-Trained for a Different, Yet Comparable Task to Evaluate Extreme-Affect Vocalizations That Are Indistinguishable by Humans

Hermann Prossinger<sup>1,2</sup>, Violetta Prossinger-Beck<sup>3</sup>,  
Silvia Boschetti<sup>4</sup>, and Jakub Binter<sup>2,4</sup>

<sup>1</sup>Department of Evolutionary Biology, University of Vienna, Vienna, Austria

<sup>2</sup>Faculty of Social and Economic Studies, University of Jan Evangelista Purkyně,  
Usti nad Labem, Czech Republic

<sup>3</sup>Technical University, Dresden, Germany

<sup>4</sup>Faculty of Science, Charles University, Prague, Czech Republic

## ABSTRACT

Humans categorize vocal displays of highly intensive affective states with very low precision. However, there are many applications necessitating correct perceptions of alarm calls. We decided to classify two negative (pain and fear), two positive (laugh and pleasure) affective states and compared these to neutral state. We used a unique dataset where all displays had been vocalized by all expressers. We used an ANN that is designed for a different, yet comparable task; one that classifies human and animal sounds as well as mundane events (such as pouring water from a jug). The outputs were then statistically analyzed using Bayesian methods. Our analysis showed that the outputs can successfully classify neutral and non-neutral affective states but they were unable to distinguish the intensive affective states from each other (with one exception: the case of laugh). Given the insights we acquired, we infer that classifying intense affective states will remain an insurmountable barrier for any future ANN. The applicability of our result also shows that the cost, time, and effort overhead of attempting to designing a dedicated ANN will be prohibitive.

**Keywords:** Affect vocalization, Artificial neural networks, Affect valence identification, Vocal cues, Bayesian methods

## INTRODUCTION

### The Current State of Knowledge

Classifying vocalizations of pain and other intense affective states is a highly complex, non-trivial task. While significant progress has been made in developing automated pain detection systems that are based on facial expressions (ANNs: Prossinger et al., 2022; Swin Transformer: Yuan et al., 2022), and on multimodal signals using feed-forward neural networks

(Gkikas et al., 2024), there remains a need for more accurate and reliable methods for identification and classification of vocal cues in isolation.

We extend our previously published approaches that dealt with facial expressions of affects (Binter et al., 2023) by proposing a novel approach that leverages pre-trained artificial neural networks (ANNs), combined with Bayesian statistics, to enable valid identification in the challenging domain of vocal cues.

Hearing signals contributes to keeping us safe in natural environments by detecting threats. In contrast to vilit (with its extremely short wavelength), sound can diffract considerably. We are thus more safe from threats even in the cases when we are unable to source these. Also, because sound does not attenuate appreciably in our close environs, the evolution of correctly perceiving signals (not only warning ones) is, arguably, a survival advantage. However, in modern settings (such as cities) meaningful signals might be missed due to cross-talk, bystander effects, as well as information overload (oftentimes characterized as sound pollution). This is where novel, intelligent algorithms can play a role in identifying or interpreting acoustic signals, thereby improving communication and, more importantly, ensuring safety.

### **The Intensity Paradox and Decision Making**

There are problems that machine-learning-based algorithms that classify acoustic signals must deal with. These arise when models are trained on sources containing human errors. These errors can bias decision-making, introduce inconsistencies in labeling, or simply generate misunderstandings. These can then lead to models that inherit and amplify these mistakes, thereby potentially perpetuating biases *and* generating inaccurate outputs (more on this problem in the Stimuli Preparation section).

Sound data is plentiful (and readily available); such data can be used to train a feed-forward neural network for a plethora of general acoustic cues. Training is, however, very expensive and only companies with considerable resources can afford to produce such networks *ab initio*. Fine-tuning already existing networks is an option — but benefits are not guaranteed. Specifically, training dedicated neural networks that can classify vocalizations of affect cues (such as the vocal expression of pain that we are investigating) threaten to be prohibitively expensive.

### **The Cost-and-Benefit Analysis — A Novel Approach**

Rather than develop and train (or fine-tune) a neural network that focusses on each affective state vocalization separately, we use the outputs that have been generated by an already available ANN from Wolfram Technologies (details below) that had been trained on acoustic data to accomplish a more general task: identify types of commonplace acoustic signals.

The novel, unconventional idea is that an existing model that classifies sounds and vocalizations will produce a discrete probability distribution of classifications (Fig. 1) for each of the human affective vocalizations we are studying. If these probabilities are numerically close within each vocalization group but with a likelihood distribution distinct enough from

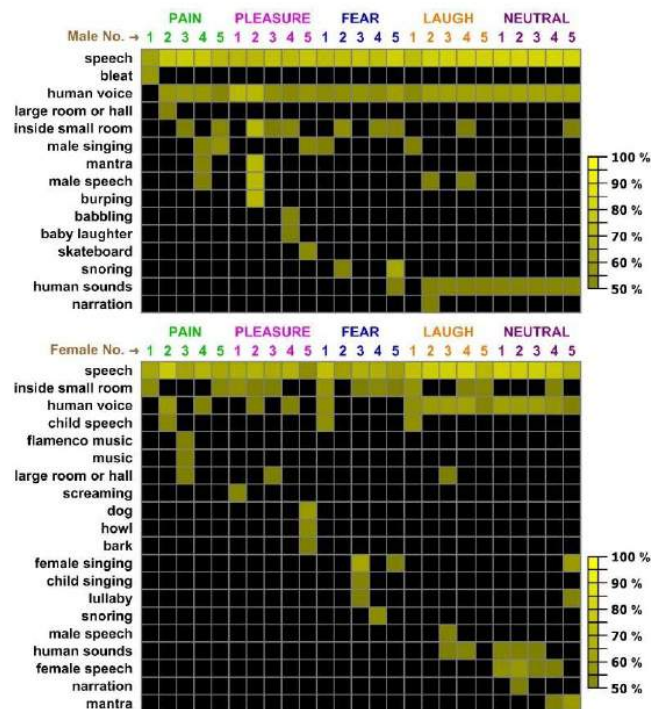


other groups, then these probability spectra (so defined) can be used for our investigations — obviating the (expensive) necessity of training a dedicated neural network from scratch.

## MATERIALS AND METHODS

### Stimuli Preparation

Current pre-testing practices (selecting stimuli with high inter-participant agreement) create “stimulus homogenization bias” (Van Der Zant & Nelson, 2021; Binter et al., 2023), thereby severely limiting investigations of natural variations in human behavior. While genuinely natural stimuli offer indisputable ecological validity, their inherent variability introduces considerable statistical noisiness in the data (no pun intended) and other extenuating acoustic peripherals (primarily acoustic cross-talk). Semi-naturalistic stimuli provide a valuable middle ground, allowing the manipulation of specific elements within a controlled setting for more nuanced and more reliable investigations.

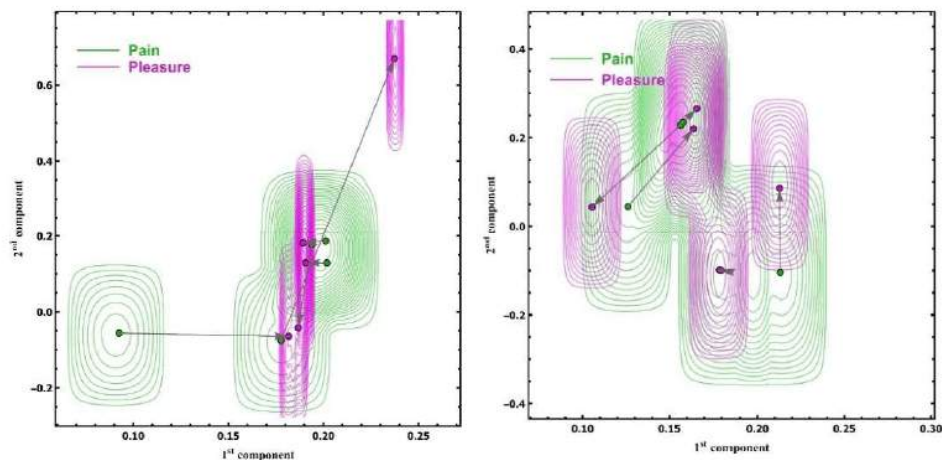


**Figure 1:** The auditory spectrum of the affects expressed by the five males and the five females. Only probabilities of identification above 50% are included in this investigation; they are color-coded. (The side bar shows the scale of the color-coding.) Absences of identifications are rendered as black squares. All males and all females have “speech” as the most probable identification. We observe that there are more classes of identifications above 50% for females (namely, 20) than for males (namely, 15). For the females, we also note there are identifications for “Neutral” that are completely absent for “Pain” and for “Pleasure”.

We used the same stimuli as described in Binter et al. (2023) and Boschetti et al., (2023). From the numerous audio-visual materials viewed, ten audio records (five with female vocalizations and five with male vocalizations) were chosen. Based on the developments of the plots in each of these audio-visual materials, five vocalizations were selected (one of “Pain”, one of “Pleasure”, one of “Fear”, one with “Laugh”, and one for “Neutral”). Relying on the contextual information, experienced researchers agreed on the stimuli that were chosen and what expression was to be expected (both visually and acoustically during viewing).

### Vocalization Identification Spectrum

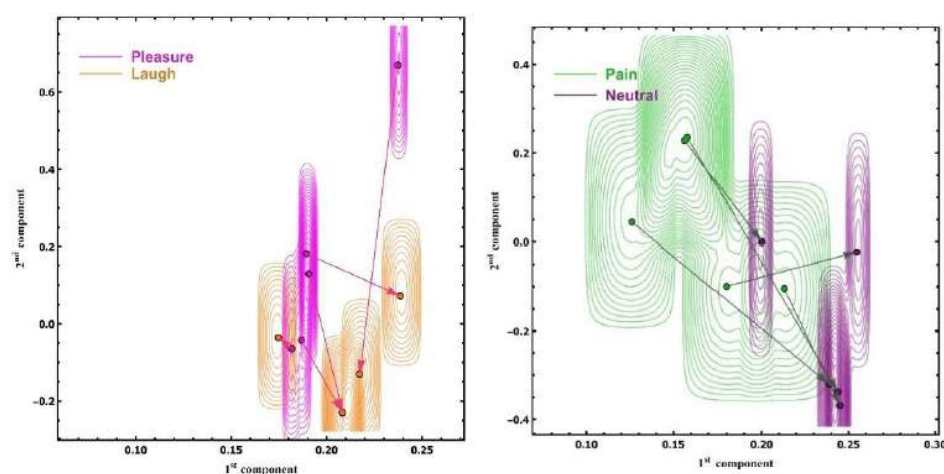
Each of the 25 vocalizations were of 1–2 seconds duration and were stored as \*.wav files. We used a pre-trained neural network (“Wolfram AudioIdentify V1 Trained on AudioSet Data”, 2019) from Wolfram Technologies that had been trained on 2 084 320 human-labeled 10-second sound clips drawn from YouTube videos. This neural network (with 4 664 911 parameters and 156 layers of seven types) identifies the possible type of acoustic signal together with the probability of each identification.



**Figure 2:** The contour plots of the *pdf* (probability density function) of the vocalizations of the affects “Pain” and “Pleasure” by the males (left) and by the females (right). The coordinates of the vocalizations were obtained by SVD of the acoustic spectrum (Fig. 1) and each *pdf* was obtained with a KDE (kernel density estimation) and an Epanechnikov kernel. In the graphs, each point is the vocalization of an affect and the contour plots show the likelihood functions of the KDEs. The rendered arrows serve to enhance visualization; they have no statistical interpretation (in this paper). The numerical values of the components have no direct interpretability. The degree of overlap between two likelihood functions is expressed by confusion matrices (Table 1), one for each biological sex and a pair of vocalizations. In these two graphs, the confusion matrix for the males is  $\begin{pmatrix} 54.6 & 45.4 \\ 12.1 & 87.9 \end{pmatrix}\%$ , and it is  $\begin{pmatrix} 57.4 & 42.6 \\ 24.1 & 75.9 \end{pmatrix}\%$  for the females. In the cases shown here, there are no significant differences between the pairings of the vocalizations, neither for the males, nor for the females.

### Dimension Reduction and Likelihood Function Estimations

We therefore obtained, for each vocalization, a spectrum of probable identifications (Fig. 1); these probabilities were then entered in two matrices: a  $15 \times 25$  one for the males and a  $20 \times 25$  one for the females. We then used SVD (Singular Value Decomposition) to dimension-reduce the entries. For each affect, we obtain two sets of five 2D vectors (one for each male and, separately, one for each female). We then estimated the 2D distributions of the 5 points per sex for each affect using a KDE (kernel density estimation) with an Epanechnikov kernel and the Silverman rule for optimizing the kernel window. In total, we statistically analyzed five  $KDE_{male}$  and five  $KDE_{female}$  for the five vocalizations per sex.



**Figure 3:** The contour plots of the *pdf* (probability density function) of the vocalizations of the affects “Pleasure” and “Laugh” by the males (**left**) and the affects “Pain” and “Neutral” by the females (**right**). The coordinates of the vocalizations were obtained by SVD of the acoustic spectrum (explained in the text) and each *pdf* was obtained with a KDE (kernel density estimation) and an Epanechnikov kernel. In the graphs, each point is the vocalization of an affect and the contour plots show the likelihood functions of the KDEs. The rendered arrows serve to enhance visualization; they have no statistical interpretation (in this paper). The numerical values of the components have no direct interpretability. The degree of overlap between two likelihood functions is expressed by confusion matrices (Table 1), one for each biological sex and a pair of vocalizations. In these two graphs, the confusion matrix for the males is  $\begin{pmatrix} 98.3 & 1.7 \\ 2.6 & 97.4 \end{pmatrix} \%$ , and it is  $\begin{pmatrix} 92.3 & 7.7 \\ 2.0 & 98.0 \end{pmatrix} \%$  for the females. In the cases shown here, there are highly significant differences between the pairings of the vocalizations, both for the males and for the females.

### Confusion Matrices for Significance Testing

We use the following implementation of a Monte Carlo method to test for a significant difference between  $KDE_A$  and  $KDE_B$ . For each of two likelihood functions ( $\mathcal{L}_A(s) = pdf(KDE_A, s)$  and  $\mathcal{L}_B(s) = pdf(KDE_B, s)$ ),

we use a pseudo-random number generator RNG to generate two sets of ran random numbers (in this manuscript,  $ran = 25000$ ). One set  $ran_A = \{RNG(KDE_A, ran_k) | k = 1 \dots ran\}$  uses the  $KDE_A$  distribution, the other set  $ran_B = \{RNG(KDE_B, ran_k) | k = 1 \dots ran\}$  uses the  $KDE_B$  distribution. We obtain four sets of likelihoods:  $\mathcal{L}_{AA} = pdf(KDE_A, ran_A)$ ,  $\mathcal{L}_{AB} = pdf(KDE_B, ran_A)$ ,  $\mathcal{L}_{BA} = pdf(KDE_A, ran_B)$ , and  $\mathcal{L}_{BB} = pdf(KDE_B, ran_B)$ . We calculate the confusion matrix  $M_C$ :

$$M_C = \frac{1}{ran} \begin{pmatrix} n_{A|\mathcal{L}_{AA} > \mathcal{L}_{AB}} & n_{A|\mathcal{L}_{AA} < \mathcal{L}_{AB}} \\ n_{B|\mathcal{L}_{BB} < \mathcal{L}_{BA}} & n_{B|\mathcal{L}_{BB} > \mathcal{L}_{BA}} \end{pmatrix}$$

where the notation  $n_{A|\mathcal{L}_{AA} > \mathcal{L}_{AB}}$  means: “the number of likelihoods when the likelihood of  $KDE_A$  of a subset of  $ran_A$  is greater than the likelihood of  $KDE_B$ ” — likewise the permutations for all other indices. If the off-diagonal elements of  $M_C$  are *both* less than 10% (Caelen, 2017), then the two distributions  $KDE_A$  and  $KDE_B$  are significantly different at 5% significance level.

**Table 1.** The table of confusion matrices, as described in the text. If we assume a significance level of 10% (Caelen, 2017) for the off-diagonal elements (this significance level corresponds to 5% in conventional, frequentist significance level assessments), then five pairings are significantly different for the males (Pain↔Laugh, Pain↔Neutral, Pleasure↔Laugh, Pleasure↔Neutral, and Fear↔Neutral) and three pairings are significantly different for the females (Pain↔Neutral, Pleasure↔Neutral, and Fear↔Neutral). These significantly different pairings are highlighted in pastel orange. However, a close inspection shows that two further pairings for the females (Pleasure↔Laugh and Fear↔Laugh) are *close to* significantly different. The rows of all confusion matrices add up to 100%; if the displayed values do not, then the reason is due to rounding of the computed entries.

Females				
	Pleasure	Fear	Laugh	Neutral
Pain	$\begin{pmatrix} 57.4 & 42.6 \\ 24.1 & 75.9 \end{pmatrix}$	$\begin{pmatrix} 62.1 & 37.9 \\ 23.9 & 76.1 \end{pmatrix}$	$\begin{pmatrix} 87.3 & 12.7 \\ 14.2 & 85.8 \end{pmatrix}$	$\begin{pmatrix} 92.3 & 7.7 \\ 2.0 & 98.0 \end{pmatrix}$
Pleasure		$\begin{pmatrix} 68.2 & 31.8 \\ 33.7 & 66.3 \end{pmatrix}$	$\begin{pmatrix} 92.9 & 7.1 \\ 12.9 & 87.1 \end{pmatrix}$	$\begin{pmatrix} 98.3 & 1.7 \\ 1.14 & 98.9 \end{pmatrix}$
Fear			$\begin{pmatrix} 92.1 & 7.9 \\ 13.9 & 86.1 \end{pmatrix}$	$\begin{pmatrix} 96.9 & 3.1 \\ 1.2 & 98.8 \end{pmatrix}$
Laugh				$\begin{pmatrix} 67.1 & 32.9 \\ 9.6 & 90.4 \end{pmatrix}$
Males				
Pain	$\begin{pmatrix} 54.6 & 45.4 \\ 12.1 & 87.9 \end{pmatrix}$	$\begin{pmatrix} 58.1 & 42.0 \\ 36.0 & 64.0 \end{pmatrix}$	$\begin{pmatrix} 91.9 & 8.1 \\ 3.9 & 96.1 \end{pmatrix}$	$\begin{pmatrix} 100.0 & 0.0 \\ 0.0 & 100.0 \end{pmatrix}$
Pleasure		$\begin{pmatrix} 85.6 & 14.4 \\ 46.8 & 53.2 \end{pmatrix}$	$\begin{pmatrix} 98.3 & 1.7 \\ 2.6 & 97.4 \end{pmatrix}$	$\begin{pmatrix} 100.0 & 0.0 \\ 0.0 & 100.0 \end{pmatrix}$
Fear			$\begin{pmatrix} 81.1 & 18.9 \\ 11.4 & 88.6 \end{pmatrix}$	$\begin{pmatrix} 92.6 & 7.4 \\ 0.0 & 100.0 \end{pmatrix}$
Laugh				$\begin{pmatrix} 20.0 & 80.0 \\ 0.2 & 99.8 \end{pmatrix}$

## RESULTS

### Comparison With the Neutral State

All affect vocalizations of both males and females are remarkably different from the neutral vocalization (Table 1) with only one exception (“Laugh” versus “Neutral” for the males).

### Comparisons of the Negative Affective States

The results of comparisons of the vocalizations of the affective states suggest that they are not significantly distinguishable (Fig. 2).

### Comparison of the Negative Affective States With the Positive Affective States

Only one positive affective state vocalized — the sound of laughter — is significantly different from the negative affective state “Pain” for males, and *almost* significantly different for females. The vocalization of the positive state “Pleasure” is significantly different from “Laugh” for males, and, again, almost significantly different for females. We address these issues in the Discussion and Conclusion section.

## DISCUSSION AND CONCLUSION

Wolfram AudioIdentify Neural Net outputs probabilities for 632 classes. Since all vocalizations of affective states, together with “Neutral” total to 21 different identifications, we can consider this test of output quality is remarkably reliable (Bayesian test of categorical variables; Beta distribution  $Be(22, 612)$ ,  $p < 2.1 \times 10^{-151}$ ; not shown).

This study is not only a “Proof of Concept”; we also present surprising findings.

We discover that, while humans are incapable of distinguishing vocalizations of affective states (Holz et al., 2021; Binter et al., 2023), we find that trained neural networks are not superior at distinguishing some of these vocalizations. Detractors from the reliance on neural networks would perhaps argue that the neural network is not specifically designed to deal with affective state vocalization classifications. We disagree; we challenge that view by observing that the probabilities of acoustic identifications are reliable (Fig. 1). Rather, as can be observed from the confusion matrices in Table 1 and the contour maps in Fig. 2 and Fig. 3, the small sample size is the shortcoming. We discover that some vocalizations are too varied. None the less, we find the contour plots are reliably interpretable; because they are highly varied; our claim that the small number of data-points is the shortcoming is supported. We conclude that the likelihood of a dedicated, *ab novo* designed neural network will have a superior performance is highly unlikely.

Even though the negative affective states are acoustically not distinguishable from one another, there is a clear distinction between any highly activated affective state and the neutral state. More importantly, there is a distinction between affective vocal displays of activities that are

considered illegal in the public arena versus those whose presentation are considered acceptable (“Neutral” and “Laughter”). Because the highly activated affective states (“Pain”, “Fear”, and “Pleasure”) are not acceptable in the public arena, our findings have an application. Consider a distress call in a public place; the neural network we have used can be used to automatically alert authorities and/or medical help personnel — despite the neural network’s inability (as well as the human’s hearing the same distress call) to distinguish pain from fear, say. We have shown that the significance levels in the confusion matrices are adequate for such applications (because they can distinguish between highly affective states and “Laugh” and “Neutral”) — despite the small sample sizes.

Furthermore, the outcomes of our analyses of the vocalization features provide insights that are valuable for ethologists and psychotherapists (when explaining why, for instance, false identifications are so prevalent). Our approach is novel, scalable and can be easily adapted to other comparable vocalization detection challenges and datasets. The costs are minimal in comparison with conventionally recommended ways of dealing with such situations — outlay for expensive hardware and then training a novel model and then fine-tuning it.

## ACKNOWLEDGMENT

This project is funded by the Czech Science Foundation (Grant No. 19-12885Y “Behavioral and Psycho-Physiological Response on Ambivalent Visual and Auditory Stimuli Presentation”).

## ETHICS STATEMENT

This study is part of a set of projects and has been approved by Institutional Review Board of the Faculty of Science, Charles University, Prague (#2018/08).

## REFERENCES

- Binter, J., Boschetti, S., Hladký, T., & Prossinger, H. (2023). “ouch!” or “aah!”: Are Vocalizations of ‘laugh’, ‘neutral’, ‘fear’, ‘pain’, or ‘pleasure’ Reliably Rated? *Human Ethology*, 2023(38).
- Boschetti, S., Prossinger, H., Prossinger-Beck, V., Hladký, T., Říha, D., & Binter, J. (2023, July). Never Correct: The Novel Analysis of Differing Visual (Facial Expression) and Acoustic (Vocalization) Bimodal Displays of the Affective States “Pain”, “Pleasure”, and “Neutral”. In: International Conference on Human-Computer Interaction (pp. 141–150). Cham: Springer Nature Switzerland.
- Caelen, O. (2017). A Bayesian interpretation of the confusion matrix. *Annals of Mathematics and Artificial Intelligence*, 81(3), 429–450.
- Gkikas, S., Tachos, N. S., Andreadis, S., Pezoulas, V. C., Zaridis, D., Gkois, G.,... & Fotiadis, D. I. (2024). Multimodal automatic assessment of acute pain through facial videos and heart rate signals utilizing transformer-based architectures. *Frontiers in Pain Research*, 5, 1372814.

- Holz, N., Larrouy-Maestri, P., & Poeppel, D. (2021). The paradoxical role of emotional intensity in the perception of vocal affect. *Scientific Reports*, 11(1), 1–10.
- Prossinger, H., Hladký, T., Boschetti, S., Říha, D., & Binter, J. (2022) Determination of “Neutral”–“Pain”, “Neutral”–“Pleasure”, and “Pleasure”–“Pain” Affective State Distances by Using AI Image Analysis of Facial Expressions. *Technologies* 10(4), 75. <https://doi.org/10.3390/technologies10040075>
- Van Der Zant, T., & Nelson, N. L. (2021). Motion increases recognition of naturalistic postures but not facial expressions. *Journal of Nonverbal Behavior*, 45(4), 587–600.
- Yuan X., Zhang, S., Zhao C., He, X., Ouyang, B., & S. Yang, (2022) Pain Intensity Recognition from Masked Facial Expressions using Swin-Transformer, IEEE International Conference on Robotics and Biomimetics (ROBIO), Jinghong, China, 2022, pp. 723–728. doi: 10.1109/ROBIO55434.2022.10011731.

# Through the Psychological Lens: Unveiling Biases in Multi-Criteria Decision-Making

**Boris Djartov, Falk Maria Oswald, and Joern Jakobi**

Institute of Flight Guidance, German Aerospace Center (DLR), Braunschweig, Germany

## ABSTRACT

Multi-Criteria Decision Making (MCDM) methods have become a mainstay in navigating complex decision-making scenarios. These methods empower individuals to consider multiple, often conflicting, criteria simultaneously. While primarily developed in computer science and operations research, the psychological implications of these methods are rarely touched on. This paper aims to address this gap by critically analysing the most established contemporary MCDM (Multi-Criteria Decision Making) methods from a psychological perspective. Due to the scope of the field this paper will restrict itself to MADM (Multi-Attribute Decision Making) methods which focus on selecting an option from a set of possible alternatives. By providing additional context and considerations, we aim to empower users to make informed decisions about their application and be mindful of their limitations.

**Keywords:** Decision-making, Multi-objective decision-making, Multi-attribute decision-making

## INTRODUCTION

In life, humans are often confronted by complex decision-making problems, such as buying stocks, choosing a university, selecting a healthcare supplier or purchasing a house. In the latter example, the decision-maker has to decide between a number of houses and does so dependent upon multiple criteria like price, location and aesthetics. Multi-criteria Decision-Making (MCDM) is a field of study, which attempts to solve these problems. It can be understood as a tool to structure the decision-making process in an attempt to find the most optimal solution. MCDM has already found application in numerous fields (Chakraborty et al., 2023; Emovon & Ogheniyerovwho, 2020; Tzeng & Huang, 2011; Zwiegelaar & Rahbarimanesh, 2022). Almost every decision-making problem employing MCDM consists of the following steps: Defining the decision-making problem, listing alternatives, determining criteria, weighting the criteria, comparing the alternatives with regards to the defined criteria, evaluating the performance of the alternatives and selecting the preferred alternative. While the fundamental steps remain consistent across the various Multi-Criteria Decision-Making (MCDM) methods, each method approaches these steps in their own unique way. Another key distinction within MCDM methods is the difference between multi-objective decision-making and multi-attribute decision-making. While



MODM methods do not operate on a predetermined set of alternatives but are restricted by a set of optimal objective constraints (e.g. the temperature for the country to travel to should measure between 10°C–20°C in the mean), the number of alternatives in MADM is predetermined and limited (e.g. a list of countries to travel to). Due to the scope of this paper, it will be mainly concerned with MADM methods. Similarly, methods that employ fuzzy logic, probabilistic methods, or group decision-making will either not be addressed or will receive limited coverage. To provide the reader with a better overview of MADM methods, a concise summary will still follow. Probabilistic methods are used to handle uncertainty in the decision-making process, which is useful when dealing with incomplete information. Similarly, fuzzy logic is utilized when decision-makers struggle to give precise numerical preferences. To paint a clearer picture, the first addresses uncertainty in external events, while the latter does so for indecisiveness within the decision-maker. While all MADM methods can be used in a group decision-making process, it still requires extra steps to funnel the different interests and preferences of stakeholders into one.

According to (Bozorg-Haddad et al., 2021) the MCDM decision-making process encompasses four distinct analytical processes: descriptive analysis, predictive analysis (which is restrained to decision under uncertainty), normative analysis and prescriptive analysis. Descriptive models are focused on describing and predicting the behaviour of an individual. In MCDM, these models account for the unique learning history of the decision-maker, capturing their preferences, weighting the criteria or making comparative judgements. Normative models, on the other hand, are concerned with the perfectly rational human and attempt to improve the decision-making outcome through rationality. In addition, Prescriptive models provide a framework guiding the decision-maker to the optimal choice of action. Lastly, if there is uncertainty involved in the decision-making process, predictive analysis provides the necessary probabilities of events happening. For example, a predictive model states how likely, it would be to assume that a caterer provides quality food for an upcoming festival. Since, most decision-makers are experts in their respective fields, they have gathered experience on the optimal course of action and MADM methods can either incorporate it to a smaller or larger degree. A subjective MADM approach relies on the intuition and instinct of the decision-maker to find the preferred alternative, while an objective approach focuses on providing a framework that helps in identifying the best solution independent of the decision-makers experience (Bozorg-Haddad et al., 2021). Regardless of the choice of either an objective or subjective approach, due to their involvement the decision-makers and their unique learning history, they will always influence the decision-making outcome.

Since, MCDM methods originated from the field of operational research the psychological standpoint has been of less relevance, although undeniably important to any decision-making process. There are many questions concerning the use of MADM methods, like: can MADM methods assess and reflect the preferences of the decision-maker, and if so, do the MADM methods outperform the decisions of the decision-makers? Hence, this paper

serves to provide an overview of MADM methods in light of psychological methodology, its challenges and ideas for further research. The structure of the paper is as follows: First, a short description of some of the most important methods, a general outlook on the problems in MCDM and at last, the MADM Methods are analysed based upon their employed methodology.

## **MADM METHODS**

Even though the procedure remains roughly the same, each method possesses its own unique assumptions, ideas and approaches to solving decision-making problems.

Analytical Hierachy Process (AHP) starts by defining the decision-making problem and expressing it through a set of criteria and alternatives in a hierarchical nature. At the top of the hierarchy stands the goal of the decision-maker; below are the criteria important to the goal; and at the bottom are the feasible alternatives. The idea of AHP is that the criteria can be best understood as single attribute decision-making problems because every alternative is evaluated with respect to one criterion at a time. Thus, providing a better overview of the decision-making problem at hand. Now, to find out which criterion is the most important in the decision-making process each criterion is compared head-to-head and rated on a 9-point scale. This procedure is called pairwise comparison and allows the decision-maker to determine the weights and the impact each criterion should have on the overall decision. The comparison of the alternatives in respect to the given criterion follows the same procedure (Saaty, 1990). If the pairwise comparison process takes a long time, AHP calls for revisions in order to see if judges have changed their minds drastically (Saaty, 1986).

ELECTRE I uses a two-step process to find a set of ideal solutions. First, in the step of aggregation, solutions which cannot be fully dominated by other alternatives are selected, while alternatives which can be dominated are removed. Meaning, that the alternatives have to perform at least as well or better in regard to some of the criteria when compared to each other. In the second step (Exploitation), the remaining alternatives are again compared and either are labeled: indifferent to each other, one is preferable to another, or no decisive statement about their relationship can be made and is thus incomparable. Through Exploitation a set of optimal solutions emerges from which the decision-maker can select the preferred alternative (Figueira et al., 2013). Again, these two steps make use of pairwise comparison to establish outranking relations. However, in ELECTRE there is no 9-point scale employed to state the dominance of one alternative over another. The method of pairwise comparison simply refers to the procedure of comparing two elements at a time.

PROMETHEE uses functions to assess the preferences of the decision-maker. These functions capture, in a pairwise fashion, the preference of an alternative with regard to each criterion in a dimensionless form  $[0;1]$ . The preference function possesses parameters that have to be calibrated by the decision-maker to make use of their knowledge. The extracted preferences are then multiplied by the weights of the criteria and summed up to assess the

overall performance of every alternative. However, the weighting procedure of the criteria is not clearly defined but rather just restricted by the condition that all weights have to sum up to 1 (Bozorg-Haddad et al., 2021). Thus, several methods can be employed to measure attribute importance. The point allocation method allows the decision-maker to distribute 100-points amongst all the criteria from which the weights can be derived directly. From the example of the introduction, Aesthetics could receive 50 points, price 30 and location 20. While the direct rating procedure lets the decision-maker rate the criteria importance on a scale from 0-100, and then needs to be standardized. There are more methods to assess these preferences, but these are some of the most common ones (Bozorg-Haddad et al., 2021).

The best worst method (BWM) follows the same procedure as AHP. However, instead of comparing every element with each other during pairwise comparison, the required number of comparisons is greatly reduced in BWM. Through the use of pairwise comparisons inconsistencies can be introduced into the decision-making process (Kuo & Chen, 2023). For example, an inconsistency would be:

*A* is preferred *B*  $\wedge$  *B* is preferred *C*; However *A* is not preferred *C*

To address this issue, BWM establishes the best and the worst criterion and compares every criterion to these two. Thus, reducing the number of comparisons needed as well as the number of logical inconsistencies. Not only are the criteria compared in such a way, but also the attribute values of the alternatives. This is the distinguishing feature of BWM (Mi et al., 2019).

## METHOD CRITIQUES

### The Validity of Preferences

To evaluate the performance of alternatives or the weights of criteria in MADM, the preferences of the decision-maker are approximated. As mentioned before, a variety of methods can be applied to measure the underlying preferences of the decision-maker. Methods like pairwise comparison, the point allocation or direct rating method make use of the decision-maker's knowledge and are subsequently called subjective weighting methods (Odu, 2019). The theoretical foundation of these methods rests upon psychological latent variable models (LVM). These models assume that manifest behavior is an indicator of some latent variable, which is inaccessible to our empirical investigation (Borgstede & Eggert, 2023). The behavior as a manifest or observable variable is said to be related to its latent or not observable variable in some way. For example, if a participant ticks the box 'strongly agree' in a survey assumed to measure assertiveness, the tick as the manifest behavior is the indicator for some value of the measured attribute assertiveness. In the case of MADM, the assessing of preferences is also a latent variable model. MADM often uses some elicitation method to weight the criteria or it compares the alternatives concerning their performance using a 9-point scale. The data produced through these methods is said to be an indicator for the latent variable preferences.

However, the concept of preferences remains vague and is often not clearly defined. In the context of a weighting procedure, the term preference can mean vastly different things. It can refer to the estimated importance of criteria because they are of interest to the stakeholders or because they are important to the decision-maker himself. In the first case, entirely different processes would be involved in estimating the importance to the stakeholders when compared to assessing one's own preferences. The former involves some estimation of variables external to oneself, whereas the latter does not. To expand on this further, preference can be understood as the perceived value of an alternative inherent to the decision-maker when making a comparison or it could also be just the choice of an alternative over another. Arguably, depending on the concept, a different measurement procedure would be best suited to assess the preferences of the decision-maker. This problem is often ascribed to the appropriation of everyday language for scientific use (Borgstede & Leising, 2019). In everyday language, a single word can refer to different constructs. For instance, someone proficient in speaking many languages might be deemed intelligent, but intelligence can also be equated with a high IQ. Both examples might speak for prowess in cognitive ability, their meaning however is dependent on context. Similarly, the measurement of intelligence depends on the measurement instrument employed. The construct of intelligence is ultimately made up of different facets, and depending on the instrument used, the facets which are measured may vary. Hence, both context and the definition of our theoretical construct are important to the measuring procedure.

Looking at the validity of the concept of 'attribute importance' (which corresponds to the criteria weights), it becomes increasingly apparent that a clear definition of the construct 'preferences' is lacking. Van Ittersum et al. (2007) conducted research on the nomological and convergent validity of different methods to measure attribute importance as employed in MADM. Validity is a key concept used in psychology to measure how far the employed instrument actually measures what it should measure. As latent variables are inaccessible to our empirical investigation, it is necessary to prove that we are truly measuring the defined construct. To assess the convergent validity, tests which should measure theoretically similar constructs are correlated. While the nomological validity measures to what degree the causal relations between theoretically similar constructs are one and the same. Even though the measurement of attribute importance should produce similar results irrespective of the instrument used, results vary between instruments. Suggesting that the different weighting techniques measure different constructs and produce different weights (Van Ittersum et al., 2007). Van Ittersum et al. (2007) propose that the different methods employed measure different aspects of attribute importance. But this perspective cannot explain the lack of validity in all cases.

### **Consistency**

A formal challenge is the consistency of judgments. Even when a decision-maker expresses a preference for *A* over *B* and *B* over *C*, it does not

necessarily mean that he will consequently prefer  $A$  over  $C$ .

$$\text{If } A \succ B \wedge B \succ C \text{ then } A \succ C$$

However, sometimes under empirical investigation:

$$A \succ B \wedge B \succ C; A \prec C$$

There are a few possible explanations for why this could be the case. First, logical inconsistencies can be introduced through pairwise comparison because the process of discriminating between two stimuli may not be deterministic and the perception of stimuli varies (Thurstone, 1994). Thurstone (1994) proposed a potential solution to this problem with his law of comparative judgment. Decision-makers are presented with alternatives or criteria and are asked to judge them as better, worse, or equivalent. To account for variance in perception and ensure consistent judgments, it is necessary to evaluate the same element multiple times. After making judgments, the mean frequency with which they preferred an alternative over another is assessed. Using the mean frequencies, a scale is constructed depicting the distance between alternatives or criteria. Thus, creating a ranking, which also shows the degree of preference through the distance between datapoints. Subsequently, there should be no logical inconsistencies, and a definitive ranking of alternatives is given to any MADM problem. As some of the criteria involved in MADM are not easily quantifiable, Thurstone (Thurstone, 1927) also provides an example of how a scale can be constructed in the case of social values. In his application of the law of comparative judgement, Thurstone evaluates the seriousness of crimes. The fundamental premise is that, if a crime  $A$  is compared to crime  $B$  and rated worse by 90% of the judges, while crime  $B$  compared to crime  $C$  is considered worse 50% of the time, then the distance between  $A$  and  $B$  should be greater than that between  $B$  and  $C$ . As aforementioned, the distance between the compared alternatives is expressed through the frequency of preference. The measurement of perceived differences between two stimuli can be done on a group or individual level (Thurstone, 1994).

Another possible explanation for the occurrence of logical inconsistencies is that the stimulus  $C$  in the example may possess properties that have not been accounted for, thereby making it preferable over stimulus  $A$ . The decision-maker might not fully understand all the factors influencing their decision. Thus, it might be beneficial to observe the decision-makers behavior in a choice experiment, where stimulus properties are systematically varied, and the choice by the decision-maker is measured.

Findings suggest, that in AHP, logical consistency are commonly introduced through its synthesis procedure (inclusion of irrelevant alternatives), its normalization procedure, criteria weights (e.g. all alternatives are equally preferred in regard to one criterion), misuse of the method, the uncertainty of the decision-making process or structural dependency between the criteria and alternatives (Borgstede & Eggert, 2023).

It also should be noted that while there have been various attempts to measure the degree of inconsistency, there have also been attempts to limit the

inconsistencies by making fewer comparisons, as in the Best-Worst-Method or revising the judgements as in AHP (Mi et al., 2019; Rezaei, 2015; Saaty, 1986, 1990). Although, the BWM introduces greater consistency into the decision-making process in relation to the law of comparative judgment, it may decrease the accuracy of the results as it will inevitably provide less information about the preferences of the decision-maker. Additionally, a consistent pairwise comparison matrix does not automatically entail the validity of expert judgments (Kuo & Chen, 2023).

### **Expertise and Experts**

Often, MADM methods try to consider the experience of an expert in hopes of achieving better results. However, there is no agreement on the definition of an expert (Day, 2002; Weinstein, 1993). It is a difficult endeavor to define what an expert really is and every field of study has its own perspective (Hill & Ready-Campbell, 2011; van Dijk et al., 2020). This makes it hard for MADM methods to identify an expert decision-maker and to make use of his experience. There are various approaches which can be used to identify experts (Caley et al., 2014; Shanteau et al., 2002). However, these approaches rarely find application in MADM methods to identify if the decision-maker is really an expert (Franek & Kashi, 2014; Keshavarz Ghorabae et al., 2017). Even if the decision-maker is an expert, the question still remains: whether the assessment made by decision-makers outperforms that of a non-expert. Additionally, one should also consider the bias that experts might also have, such as outcome bias which experienced pilots can also be prone to (Martins et al., 2023). According to (Hogarth et al., 2015) this is highly dependent on the context. The inherent reliability of the decision-making environment makes his judgment more or less accurate. If future events can be predicted through the use of past data, this makes an 'expert' judgement more reliable. If this is not the case, non-experts may perform just as well as experts (Green & Armstrong, 2007).

Furthermore, the expert does not necessarily use more information than the layman to inform his decision making process, but he does weigh criteria differently when compared to the layman (Shanteau, 1992). Thus, it is not necessary to include many decision-making criteria to make use of the decision-makers experience.

More importantly, there is a debate concerning if MADM methods are even able to make use of the decision-makers experience. Harries and Harvey (Harries & Harvey, 2000) argue that participants' judgments seems more reflective of how the criteria ought to have been used rather than how they actually have been used. Hence, it becomes questionable if MADM methods can effectively utilize the acquired experience of the decision-maker. Nevertheless, (Riquelme, 2001) demonstrated that humans should be able to articulate their judgment policies. In an experiment, subjects were asked to give a holistic judgement on their intention to buy a mobile phone plan and subsequently had to weight the importance of the criteria employed to describe the mobile phone plan. Weights derived from holistic judgments were compared with the weights stated by the subjects after the

trial. The level of correlation between derived and self-explicated weights suggests that humans can accurately describe their own judgment policies. However, certain methods do not even lay claim to incorporate the experience of the decision-maker and just try to account for the wishes and interests of the stakeholders. In this case, criteria do not necessarily need to reflect the most optimal properties according to the decision-maker, but rather represent properties the stakeholders deem to be important. These methods then systematically eliminate alternatives until one is found which best fits the optimal outcome for the stakeholders' interest and wishes (Bozorg-Haddad et al., 2021).

## CONCLUSION

In conclusion, Multi-Attribute Decision-Making (MADM) methods offer powerful tools for tackling complex decision-making problems, but their effectiveness and reliability are subject to various challenges, especially when evaluated from a psychological standpoint. The paper highlighted the key steps involved in MADM processes, such as defining the problem, listing alternatives and criteria, weighting criteria, comparing alternatives, evaluating performance, and selecting the preferred option. Several MADM methods were examined, including Analytical Hierarchy Process (AHP), Best Worst Method (BWM), ELECTRE, and PROMETHEE. The analysis focused on identifying general problems inherent in MADM approaches. These include the validity of preferences, consistency in judgements and the role of expertise and experts. Each of these issues presents significant challenges that must be addressed to ensure the reliability and effectiveness of MADM processes.

Looking ahead, future research should focus on addressing these challenges and exploring new avenues for improvement. This includes refining measurement techniques for assessing preferences, enhancing our knowledge of preferences or attribute importance, and better understanding the role of expertise in decision-making. Especially, further investigation into pairwise comparison methods can contribute to advancing the field and improving its applicability in real-world decision-making scenarios.

Overall, while MADM methods hold great potential for aiding decision-makers in navigating complex choices, ongoing research and refinement are necessary to fully realize their benefits and overcome their inherent challenges. By addressing these issues, MADM can continue to evolve as a valuable tool for decision-making in various domains.

## REFERENCES

- Borgstede, M., & Eggert, F. (2023). Squaring the circle: From latent variables to theory-based measurement. *Theory & Psychology*, 33(1), 118–137. <https://doi.org/10.1177/09593543221127985>
- Borgstede, M., & Leising, D. (2019). *Hypothetical Constructs*. [https://doi.org/10.1007/978-3-319-28099-8\\_679-1](https://doi.org/10.1007/978-3-319-28099-8_679-1)

- Bozorg-Haddad, O., Loaiciga, H., & Zolghadr-Asli, B. (2021). *A Handbook on Multi-Attribute Decision-Making Methods*. <https://doi.org/10.1002/9781119563501>
- Caley, M. J., O'Leary, R. A., Fisher, R., Low-Choy, S., Johnson, S., & Mengersen, K. (2014). What is an expert? A systems perspective on expertise. *Ecology and Evolution*, 4(3), 231–242. <https://doi.org/10.1002/ece3.926>
- Chakraborty, S., Raut, R. D., Rofin, T. M., & Chakraborty, S. (2023). A comprehensive and systematic review of multi-criteria decision-making methods and applications in healthcare. *Healthcare Analytics*, 4, 100232. <https://doi.org/10.1016/j.health.2023.100232>
- Day, J. (2002). What is an expert? *Radiography*, 8(2), 63–70. <https://doi.org/10.1053/radi.2002.0369>
- Emovon, I., & Oghenyerovwho, O. S. (2020). Application of MCDM method in material selection for optimal design: A review. *Results in Materials*, 7, 100115. <https://doi.org/10.1016/j.rinma.2020.100115>
- Figueira, J., Greco, S., Roy, B., & Słowiński, R. (2013). An overview of ELECTRE methods and their recent extensions. *Journal of Multi-Criteria Decision Analysis*, 20. <https://doi.org/10.1002/mcda.1482>
- Franek, J., & Kashi, K. (2014). A Review and Critique of MADM Methods and Applications in Business and Management. *International Journal of the Analytic Hierarchy Process*, 6. <https://doi.org/10.13033/ijahp.v6i2.254>
- Green, K., & Armstrong, J. (2007). The Ombudsman: Value of Expertise for Forecasting Decisions in Conflicts. *Interfaces*, 37, 287–299. <https://doi.org/10.1287/inte.1060.0262>
- Harries, C., & Harvey, N. (2000). Taking advice, using information and knowing what you are doing. *Acta Psychologica*, 104(3), 399–416. [https://doi.org/10.1016/s0001-6918\(00\)00038-x](https://doi.org/10.1016/s0001-6918(00)00038-x)
- Hill, S., & Ready-Campbell, N. (2011). Expert Stock Picker: The Wisdom of (Experts in) Crowds. *International Journal of Electronic Commerce*, 15(3), 73–102. <https://doi.org/10.2753/JEC1086-4415150304>
- Hogarth, R., Lejarraga, T., & Soyer, E. (2015). The Two Settings of Kind and Wicked Learning Environments. *Current Directions in Psychological Science*, 24, 379–385. <https://doi.org/10.1177/0963721415591878>
- Keshavarz Ghorabae, M., Amiri, M., Zavadskas, E. K., & Antucheviciene, J. (2017). Supplier evaluation and selection in fuzzy environments: A review of MADM approaches. *Economic Research-Ekonomska Istraživanja*, 30(1), 1073–1118. <https://doi.org/10.1080/1331677X.2017.1314828>
- Kuo, T., & Chen, M.-H. (2023). On using pairwise comparison in the analytic hierarchy process: Validity is goal while consistency is means. *Inf. Sci.*, 648(C). <https://doi.org/10.1016/j.ins.2023.119630>
- Martins, A. P. G., Köbrich, M. V., Carstengerdes, N., & Biella, M. (2023). All's well that ends well? Outcome bias in pilots during instrument flight rules. *Applied Cognitive Psychology*, 37(2), 433–442. <https://doi.org/10.1002/acp.4046>
- Mi, X., Tang, M., Liao, H., Shen, W., & Lev, B. (2019). The state-of-the-art survey on integrations and applications of the best worst method in decision making: Why, what, what for and what's next? *Omega*, 87, 205–225. <https://doi.org/10.1016/j.omega.2019.01.009>
- Odu, G. (2019). Weighting methods for multi-criteria decision making technique. *Journal of Applied Sciences and Environmental Management*, 23, 1449. <https://doi.org/10.4314/jasem.v23i8.7>



- Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53, 49–57. <https://doi.org/10.1016/j.omega.2014.11.009>
- Riquelme, H. (2001). Do consumers know what they want? *Journal of Consumer Marketing*, 18(5), 437–448. <https://doi.org/10.1108/07363760110398772>
- Saaty, T. L. (1986). Axiomatic Foundation of the Analytic Hierarchy Process. *Management Science*, 32(7), 841–855.
- Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26. [https://doi.org/10.1016/0377-2217\(90\)90057-I](https://doi.org/10.1016/0377-2217(90)90057-I)
- Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, 81(1), 75–86. [https://doi.org/10.1016/0001-6918\(92\)90012-3](https://doi.org/10.1016/0001-6918(92)90012-3)
- Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research*, 136(2), 253–263. [https://doi.org/10.1016/S0377-2217\(01\)00113-8](https://doi.org/10.1016/S0377-2217(01)00113-8)
- Thurstone, L. L. (1927). The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21(4), 384–400. <https://doi.org/10.1037/h0065439>
- Thurstone, L. L. (1994). A law of comparative judgment. *Psychological Review*, 101(2), 266–270. <https://doi.org/10.1037/0033-295X.101.2.266>
- Tzeng, G.-H., & Huang, J.-J. (2011). *Multiple Attribute Decision Making: Methods and Applications*. Chapman and Hall/CRC. <https://doi.org/10.1201/b11032>
- van Dijk, E. E., van Tartwijk, J., van der Schaaf, M. F., & Kluijtmans, M. (2020). What makes an expert university teacher? A systematic review and synthesis of frameworks for teacher expertise in higher education. *Educational Research Review*, 31, 100365. <https://doi.org/10.1016/j.edurev.2020.100365>
- Van Ittersum, K., Pennings, J. M. E., Wansink, B., & van Trijp, H. C. M. (2007). The validity of attribute-importance measurement: A review. *Journal of Business Research*, 60(11), 1177–1190. <https://doi.org/10.1016/j.jbusres.2007.04.001>
- Weinstein, B. D. (1993). What is an expert? *Theoretical Medicine*, 14(1), 57–73. <https://doi.org/10.1007/BF00993988>
- Zwiegelaar, J., & Rahbarimanesh, A. (2022). Application of Multi Criteria Decision Making (MCDM) to Analyse the Impact of External Environment on Innovation Ecosystem. *International Conference on Information Systems 2022 Special Interest Group on Big Data Proceedings*. <https://aisel.aisnet.org/sigbd2022/3>

# Redefining Creativity and Artistic Endeavours: Exploring the Impact of AI-Generated Digital Art on Human Society

**Amic G. Ho**

Hong Kong Metropolitan University, School of Arts & Social Sciences, Department of Creative Arts, Hong Kong, 00852

## ABSTRACT

Artificial intelligence has transformed the field of creative work, challenging existing ideas of creativity and creative work. AI-generated digital art has emerged at the crossroads of innovative technology and human creativity and prompted an extensive investigation into its potential consequences on a wide range of human activity. Students also consider whether AI-generated digital art influences individuals' willingness to engage in art-related activities and produce their work. The use of AI in art evokes components of self-consciousness and identity. Furthermore, the researchers examined the colours provoked by AI-generated digital art, revealing useful information on how AI art generates new emotions and interactions. This article seeks to identify ways in which AI technologies and methodologies open up new possibilities for creators and artists, investigating through case studies, surveys, and qualitative and quantitative measures the transformative potential of AI-generated digital art and its influence on cultural and social dynamics.

**Keywords:** AI digital art, Impact on society, Creativity, Motivation, Cultural dynamics

## INTRODUCTION

Artificial intelligence has been connected to a new form of creative expression since it arrived; this new field is on the verge of upending traditional ideas of creativity, as well as artistic efforts. The advent of creative digital art as a product of AI makes it a novel merging of the most cutting-edge technological tools and human creativity, necessitating a closer examination of all possible consequences across multiple areas of the human environment. Furthermore, the effect of AI-generated digital art on human endeavours to create and partake in artistic creation is explored. The impact of the LIA platform on vision creation has been Calibrated. The Other AI domes have had an impact on people's inclination to see. Spin versus clue situations in LIA-calibrated photos were conducted, all of which were compatible with our hypotheses. Self-awareness and identity are also becoming material in relation to AI technologies in the sector of art. It refers to the emotional turns produced by AI technologies and examines their consequences on individuals and

societies. Artworks created by AI or with the aid of AI technology scaffold novel modes to gather and be social and are comprehended better via an investigation of the twists, ushers, and urges in affectivity that they elicit. Investigating the character of AI as a partner or auxiliaries in the art of creating supplies beneficial information. The idea of how the AI art generates supplanted interplays is recognised by examining certain AI technologies. This study aims to examine the different stages through which the jobs of artificial intelligence technologies and technologies provide fresh incitements for artists and creative persons to delve into new and nonconforming art and creative practices. Combinedly, qualitative and quantitative study tools, including surveys and standardised observations, are used to gain information on people's experiences and attitudes toward the party. The goal of this outline is not just to further our understanding of the transformational potency.

### **GENERAL ON THE IMPACT OF AI-GENERATED DIGITAL ART ON PEOPLE AND SOCIETY**

The emergence of AI-generated digital and its development might have several factors that have varied effects on people and their collective integration. The factors and effects might provide a background of potential implications of the broad use of AI arts. The factor above might be critical in determining the effects of AI art. Most likely, the threat of people's direct use of AI in creating art involves the painters themselves and the associations they create with the artwork. This is because they technically outshine the artists. Hence, understanding the importance and implications of AI art shake the art scene provides an oversight of how this technology can contribute to several levels of innovation. One more factor to consider is people's emotional reactions (Ho, 2024a) to AI-generated digital art. As a theoretical realisation created by machine learning algorithms, this art form is expected to provoke different emotions than other forms of human art. Moreover, understanding these emotions might have implications for AI as collaborators and facilitators and their effects on people's self-awareness. Finally, the potential contribution of AI (Ho, 2024b) to the broader cultural and social context is another factor to consider. AI art threatens to change the paradigm of current cultural achievements and practices. Thus, understanding the extent to which AI can serve as a constructor to support the expansion and disturbance of artistic expression provides insights into the long-term view. Upon understanding all aspects of people's responses to AI-generated digital art, and its implications on society, researchers are likely to take valuable insurance for future policymaking and research.

### **THE EMOTIONAL RESPONSES EVOKED BY AI-GENERATED DIGITAL ART**

Thus, the emotional responses such art would evoke are indeed a significant dimension for analysing its consequences for people and the community at large. Because art created with the use of machine learning algorithms

might prompt people to novel emotional responses that differ from those provoked by traditional human-created art. Thus, due to the uniqueness of the emotional responses generated by AI-powered digital art, understanding the specific nature of such creativity means how synthetically its emotional capabilities are constructed. Specifically, this is because the creativity of AI art is based on the analysis and interpretation of large datasets using machine learning, unlike human creativity, which is often derived from everyday experience with life, originality, and emotion. Accordingly, the response to this should be questions about the experience, or the genuineness of its emotionality when people interact with these platforms. This question may concern the nature of responses, such as amazement, admiration, perplexity, discomfort, or intimidation generated by this art. This perspective helps to clarify creativity's essence and its potential for interaction with people or independent functioning. It may refer to the question of human self-recognition and perception too. For example, the emotional responses to AI art that imitates human creativity and creation differ from those inspired by art assuming more abstract and unconventional forms. Therefore, analysing this difference shows us how much people genuinely consider AI-generated art to be creativity, to what extent that art prompts emotional responses or not. Furthermore, understanding emotional responses to AI art might reveal the state of the society. The scale or extent of people's perceptions that AI contributes to art might show the underlying fears about the future of this field, human displacement by machines, or specific dynamics of social culture. Hence, this artificially intelligent direction of the research presents researchers with a complicated matter of the interface of humanity, creativity, and technology. This knowledge enlightens policymakers, professional practitioners, and educational players on how to develop policy, practice, and learning that may address the challenge and potential of AI integration into art.

### **THE INFLUENCE OF AI ART ON CULTURAL AND SOCIAL DYNAMICS**

Some studies provide a comprehensive understanding of the influence of AI art on cultural and social dynamics, which is multifaceted and continually expanding to encompass aesthetics, identity, and several other dimensions. For instance, the integration of AI-based models in the study of cultural elements in folktales to relate them to modern chores as facilitated by the concept of cultural additivity has provided some behavioural and evolutionary insights (Ho, 2024c). The aesthetic dimension of art has also been associated with a national identity which relates the cultural aspects of a region with the appropriation of art. Similarly, the relationship between machine agency and human agency has revealed the transformative nature of AI models, especially in user experiences in social media. The integration of AI in influencer marketing and the evolving creativity, AI, and everyday creativity conversation have also been discussed. Moreover, the discussion focused on ethical considerations such as AI art's potential for human flourishing about social justice (Fiske et al., 2019). Similarly, the concept of AI's social-relational moral standing and its effect on property thesis-based moral have been discussed. The concept of smart urban infrastructure and socially embodied AI were reviewed considering the psycho-cultural barriers

and their implication on cultural readiness. AI itself, culture-political, social structure layer and technological gaps were also analysed. Alternations into the dynamics of AI art as movement and community and the effect of arousal level in determining aesthetic preference were also discussed. Lastly, the practices of the artists and others and the possibility of revealing the art system were reviewed.

## METHODOLOGY

To explore the dissemination of AI-generated digital art and its consequences for humans, the research team constructed a mixed-method approach for this study. They carried out a series of surveys, aimed at collecting data about individuals' experience with AI-generated digital art, and their thoughts on years of development in this field. The surveys took in such diverse aspects as emotional reactions, perspectives on creativity and authorship, motives for participating in AI art, and effects perceived as being brought by it on creative work. This survey data resulted in a more panoramic view of AI-generated artwork in our society and the response it has evoked from both professional artists and poetry lovers. To complement the survey data, some local case studies- which delved more deeply into the situation of each specific case in question were conducted by members of our research group. These studies interviewed professional artists, art lovers, art critics, and others concerned before embarking on research work itself. They are needed to shed light on both personal experiences with this new type of technology-induced art as well as what happens in practice at the societal level. The case studies also investigated specific AI art projects or exhibitions to elucidate the artistic, technical, and societal significance of this new approach further.

**Table 1.**

No.	Question	Objective
7	How familiar are you with AI-generated digital art?	Assess familiarity with AI-generated digital art
8	Have you ever encountered or experienced AI-generated digital art?	Determine experiences with AI-generated digital art
9	What are your initial thoughts and reactions when you encounter AI-generated digital art?	Explore initial thoughts and reactions to AI-generated digital art
10	How do you perceive the creative process involved in AI-generated digital art?	Understand perceptions of the creative process involved in AI-generated digital art
11	Do you think AI-generated digital art can be considered a form of true artistic expression? Why or why not?	Evaluate whether AI-generated digital art is considered a form of true artistic expression
12	What kind of emotional responses do you typically experience when engaging with AI-generated digital art?	Identify emotional responses to AI-generated digital art
13	How do these emotional responses compare to your experiences with traditional human-created art?	Compare emotional responses to AI-generated digital art and traditional human-created art
14	Do you feel that the emotional responses evoked by AI-generated digital art are authentic and meaningful?	Assess whether the emotional responses to AI art are perceived as authentic and meaningful

(Continued)

**Table 1. Continued**

No.	Question	Objective
15	How do you think the emotional responses to AI art may influence individuals' perceptions and attitudes towards art and creativity?	Examine the potential influence of emotional responses to AI art on perceptions and attitudes
16	If you are an artist or art enthusiast, how has the emergence of AI-generated digital art influenced your motivation to create or engage with art?	Investigate the impact of AI-generated digital art on artistic motivation and engagement
17	Do you feel that AI-generated digital art has the potential to disrupt or reshape the traditional art world and artistic practices? If so, in what ways?	Explore the potential for AI-generated digital art to disrupt or reshape the traditional art world
18	How do you think the availability of AI-generated digital art may impact individuals' engagement with the arts, either as creators or appreciators?	Assess the impact of AI-generated digital art on individuals' engagement with the arts
19	In what ways do you think AI-generated digital art may influence cultural dynamics, such as art appreciation, art market structures, and the perceived value of art?	Understand the influence of AI-generated digital art on cultural dynamics and art appreciation
20	How do you believe the integration of AI in the artistic realm may impact social interactions, communities, and collective identity?	Examine the potential impact of AI-generated digital art on social interactions, communities, and collective identity
21	What are your concerns or hopes regarding the long-term implications of AI-generated digital art on society?	Identify concerns and hopes regarding the long-term implications of AI-generated digital art on society
22	Do you have any additional thoughts or insights you would like to share regarding the impact of AI-generated digital art on individuals and society?	Provide an opportunity for additional insights and perspectives on the impact of AI-generated digital art

### Participant Recruitment Criteria

To ensure that a diverse range of audience groups from different segments of society are represented, the research project followed the stated audience recruitment guidelines and procedures. The participants recruited in the study from the general public include anyone aged between 18 and 60 years. The participants recruited are drawn from various gender identities, among them female, male, and other gender expressions. In addition, the education level recruited from the general public draws a wide range starting from uncompleted high school to an individual with a master's degree and above. The general public recruited in the study was drawn from different economic backgrounds in terms of their income levels or socioeconomic status from poor to rich. The place of residence is also considered, including both urban and rural areas to understand the social welfare and cultural significance of the matter. Additionally, the research project recruited art enthusiasts and artists aged 18 to 60 years. The recruited individuals in the study have different artistic achievements, from professional artists to less professional artists and those in between have interests in the world of art but at various stages. In terms of art medium, the research project does not set boundaries to the type of art; hence the medium ranges from traditional bronze sculptures to

performance arts and the new media art forms are only appreciated through the internet. Furthermore, the research project is inclusively designed in such a way that a wide range of experts and key stakeholders are considered in the study. The experts solicited for participation include persons with expertise in art criticism, curation, computer science, AI research, technology industry, art historical background, psychology, sociology, and anthropology. Policymakers and industry folks participating are involved in formulating policies or implementation of initiatives focused on the integration of new techniques in the world of art such as the manipulation of modern AI art forms.

### **Surveys to Gather Data on Individuals' Experiences and Perspectives**

The existing research on the implications of AI-generated art can also provide a valuable basis for exploring the complex range of factors associated with this emerging field for human self-awareness and identity. The investigation can be based on the above-mentioned references, focusing on the five key dimensions. Particularly, the study by Liu (2023) explored the perspective for the rise of AI art, discussing the aspects of creativity, motivation, self-awareness, and emotionality in this technology. These findings would thus help determine whether and to what extent AI-generated art should be called an act of creation and how it can impact the human understanding of creativity and the arts. The research presented by Zhang et al. (2021) outlines the necessity of laypeople to use AI art for emotional and personal reasons such as entertainment, aesthetics, self-expression, and social preferences. In this way, these sources would provide a background for investigating the changes in human understanding of art experience and perception of oneself as a creator based on the integration of AI technologies.

### **Case Studies to Provide In-Depth Insights Into the Impact of AI Art**

Here are some case studies of AI art projects and exhibitions. 'The Curious Behaviour of Humans' Exhibition featured a series of AI-generated portraits which aimed to reflect how we see our own lives. The findings of the case study are recorded in the form of interviews with the curator of such an exhibition, Emma, and numerous participating artists. The exhibition was staged in London's Barbican Centre where it registered 120,000 visitors over three months. The featured works are acquired by The Tate Modern and Centre Pompidou sealing their status within the art world. The Guardian reviewed The exhibition, with the critic calling it an eerie and uncanny observation that 'Autonomous Portraits' created.

The Interactive installation 'Algorithmic Dreamscapes' was commissioned by the Tate Modern, in London. The run is for six months, and it draws 80,000 visitors. This project received the 'Excellence in Interactive Art' award from the annual Ars Electronica Festival. The exhibition that was cooler at the Tate Modern showed at several international art festivals such as the Sydney Biennale and Linz's Ars Electronica Festival. The AI art competition 'Generative Visions' drew over 800 submissions from 47 countries. The winning works are displayed at the COP26 climate summit

held in Glasgow, Scotland. The AI Tapestries exhibition, on the other hand, features a series of AI-generated textile artworks. The exhibition curator, Ling, and some of the participating artists are interviewed. Held at the Museum of Arts and Design in New York City, two of the artists featured, Fatima and Javier, won residencies at MIT Media Lab where they could continue developing their AI-powered textile art practice. The exhibition reviewed by Hyperallergic, the online art and culture publication, demonstrates captivating relationships through interlocking surfaces between ancient craft and futuristic technology. With these further case studies, together with their tracked records and online references, definitive evidence has been given of AI-generated digital art's increasing recognition and impact within the art trade and public culture. The varied examples underscore the expansive range of artistic approaches adopted, institutional involvement, and popular acclaim that this emergent field has won, stressing once more the need for continual discussion and investigation into the integration of AI in creative arts.

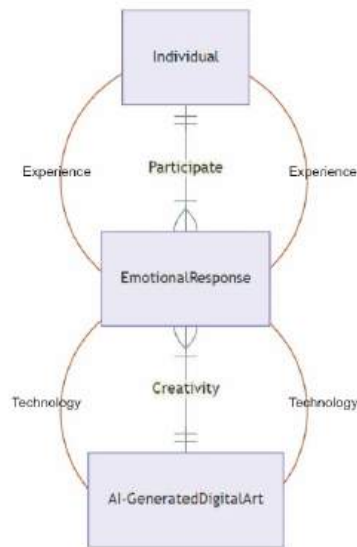
## **EXPECTED OUTCOMES AND SIGNIFICANCE**

In human society, AI-generated digital art impact comprehensive research study is expected to achieve several important results in the emerging sphere and its future attending to which specifics we can use our knowledge of the following.

### **Deepening Our Understanding of the Human Societal Impact on AI-Generated Digital Art**

This study intends to achieve a rich understanding of the complex impact of AI-generated digital art in its different forms on individual recipients, artists and society as a whole. By combining quantitative analysis data from surveys with case studies and interdisciplinary cooperation, the research team uncovered all of these multifaceted ways in which technological innovation has come into human perception; artistic expression; emotional reactions and behavioural patterns as well in moving human perceptions towards present freedom of creativity that is which result comes. Estimates for years to come. These 30 participants in the general public were interviewed and their experiences of AI-generated digital art were documented, in the manner of a photo survey. This provided insights included: how familiar people are with AI art; creativity and artistic expression in their opinion; emotional responses; and well fears that AI might ruin their motivation away from working hard. Interviews with AI painter Alex and art fan Sarah allowed a look at reactions and complex denials of the idea that modern machines could interfere with creative processes. Through synthesising both large surveys with proper sampling techniques designed by experts as well as smaller cases in which people were followed for 3 to 6 months at a time, the study group came to a comprehensive multi-levelled understanding of what happens when individuals or society are changed by AI-generated art.





**Figure 1:** Entity relationship diagram for the intersection of technology, creativity, and human experience.

### **Providing Clues to the Transformative Nature of AI Art and Where it May Head in Future**

The research conclusions showed the transformative possibilities of AI-generated art. AI-generated art has the potential to change artistic practice, reinvent traditional ideals of what is a proper stimulus in creative activity and topple existing structures from the art world on down. These insights informed consideration of future paths for art, who defines what's creative and what society means by 'creativity'; where technology's role in production might be taking us all next-issue scholarship that can effectively assist people in exploring better longer lasting solutions. Case studies of particular AI art projects and exhibitions (e.g., the 'Algorithmic Expressions' exhibit starting in the autumn of 2009 and the 'Generative Visions' competition) feature in various places throughout this report. They show different directions for AI as technology changes and redefines, thanks to its reach and response to what we live through: algorithms can create one unique thing after another.

With the case study findings and the result of the survey on how traditional artistic practices and art markets perceive AI's impact, the transformation potential of art produced by AI was fleshed out. It suggests that the established understanding of authorship could be overthrown; it gives rise to new art communities and ways of being involved; and with advances in technology, what it means to be an artist is changing -indeed all along these directions. These findings shaped the team's overall research objectives, which aim to discern what is at stake for art, creativity, and human-technology interactions, leading up to tomorrow and also spell out that

responsible development and incorporation of AI into creative arts are critical.

### **Guidebook for Policies and Practice Aimed at Sustaining AI Art**

On the back of such a thorough investigation, the results supported efforts to develop policy guidelines and best practices in this new sphere of AI art that meets its special challenges as well as opportunities. It includes considerations related to things like authorship, intellectual property, ethical concern and responsible innovation in the field of production with AI-generated technology. The comprehensive analyses of survey data and case studies have revealed the potential risks and benefits that AI brought to the world of art. These include concerns about authenticity, rights to intellectual property and whether human artists may be entirely replaced. Through understanding the key points, the project team can from different angles in various voices and actions go on to make rational suggestions on policies, guidelines, and best practices for these issues. We hope that in this way it can be guaranteed the development and use of AI in Creative Arts served, as it should, both human creativity and artistic expression. Only by so cultural traditions were maintained.

### **Adding to the Canonical Work on the Intersection of Technology, Creativity and Human Experience**

Utilising research results published in academic journals, conference presentations and public participation, this work further discussed into today's creative arts and how it's experienced among people. The findings achieved here were part of discussions that involve artists, scientists, policymakers educators and those who consume art all around the world. Through dialogues like these, we define our future as a whole not simply single people acting on their own. The paper can give new perspectives and good evidence to society's general discussion about the relationship between computer technology, creativity art culture. Via research papers, contributions in collecting material, etc., it augured the results of this study in hopes that they enjoy wide circulation among artists and scientists as well as with government folk and also for people at large. For example, the case study viewed from their respective classical and practical implications of AI-generated art viewpoints --give us vivid stories with which we can vividly appreciate how we might bring together such instances. This fosters a richer conversation on technology's evolving place in creativity and its effects on man, an area likely eventually to necessitate responsible integration of AI into arts.

By integrating the quantified survey data with qualitative case study insights, the group of researchers filled out a detailed, multi-dimensional understanding of what impact AI-generated digital art has on human society. This paper feeds into future policy-making and evolving practice within this fast-changing field. In the design of AI-generated Digital Art, the research results can redefine, for example, the whole future direction of art and its relationship with technology within human society. If AI-generated digital

art is to truly enrich human life, then these findings provided a highly useful choice as a reference source for artists, art institutions, policymakers and the public to engage in more sophisticated conversation covering our interests and objectives in integrating technology with creative forms.

## CONCLUSION

The profound analysis of AI-generated art and its influence on society provides valuable insights into the potential future implications of this emerging discipline, beyond our initial expectations. The research team has utilised both quantitative survey data and detailed case studies to gain a comprehensive understanding of how AI art impacts individual perceptions, artistic practices within communities, and its role as a catalyst for broader artistic movements. The poll data revealed contrasting perspectives among the general public on AI-generated art, with some individuals expressing a sense of apprehension and pondering the source of originality when considering that machines are capable of creating such works. However, other individuals were content to temporarily embrace a replica or counterfeit that was not of the highest quality, as long as there was a continuous manufacture of these duplicates. The abovementioned findings underlie the need for a more sophisticated and sophisticated understanding of how people interact with this inanimate object in their artistic environment. Through in-depth case studies, the researchers were able to obtain a better understanding of the phenomenon by exploring contextual factors and stories that shape the lives of several stakeholders in the AI-furled art production world: the AI art makers, the backers, and the consumers. The team gained a holistic perspective on the different dimensions of the impact of AI art on people by using quantitative and qualitative data together. Awareness is crucial as it guides the recognition of the necessity for measures, such as regulations, that incorporate human values into human systems utilising new technologies to create artworks. The repercussions of integrating technological developments into communities, both legally and socially, can limit real-world implementations in several ways.

Although this research provides valuable insights, it is necessary to further analyse and explore the dynamic and rapidly evolving role of AI in art. Potential future study avenues could encompass longitudinal studies that document the enduring influence of AI art on artistic communities, cultural institutions, and society as a whole. Furthermore, engagement in multidisciplinary partnerships with people from other disciplines such as philosophy, cognitive scientists, and cultural studies would provide promising insights into the challenges that arise from the integration of AI in areas of creativity. Studying the interlinking areas of AI art and other emerging creativity-related developments such as virtual reality and augmented reality would provide an exciting endeavour for study as it offers new insight into how different technologies are transforming the relationship between technology, creativity, and human experience. Comparative studies of AI art from different cultural settings and areas of creative tradition would also establish a holistic picture regarding the global impact of AI on creativity. The

field of AI art study is very critical as it redefines how creativity should be viewed and challenges the very basic idea of authorship. The outcome of this research study would, therefore, create awareness of the need for informed and thoughtful conversation in this incorporation. An inclusive platform for discussion on technology and creativity development would be established through this discovery dissemination, a process that promotes the production of academic papers, conference presentations, and popularisation publishing. The forum would be critical in ensuring transformative and novel ideas are used to guide the ethical integration of AI in areas of creativity. Eventually, this would be a major factor in determining the fate of art in human society.

## REFERENCES

- Fiske, Amelia, Peter Henningsen, and Alena Buyx. (2019). Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research* Volume 21, No. 5.
- Gu, Li and Yong Li. (2022). Who Made the Paintings: Artists or Artificial Intelligence? The Effects of Identity on Liking and Purchase Intention. *Frontiers in Psychology* Volume 13.
- Ho, A. G. (2024a). Developing strategies for junior designers to manage the design and emotions. *New Design Idea* Volume 8, No. 1.
- Ho, A. G. (2024b). “‘Techspressionism’ of Digital Art in Contemporary China”, proceedings of the 10th International Symposium of Chinese Computer–Human Interaction (Chinese CHI 2022). Association for Computing Machinery, New York, NY, USA.
- Ho, A. G. (2024c). “The Theories and Impact of Mechanical Creation in Digital Art and Design Learning”, proceedings of the 2023 IEEE International Conference on Development and Learning (ICDL). Macau, China. <http://doi.org/10.1109/ICDL55364.2023.10364499>
- Lima, Gabriel, Zhunis, Assem, Manovich, Lev, Cha, Meeyoung 2021. On the Social-Relational Moral Standing of AI: An Empirical Study Using AI-Generated Art. *Frontiers in Robotics and AI* Volume 8, No. 2021.
- Liu, Bai. (2023). Arguments for the Rise of Artificial Intelligence Art: Does AI Art Have Creativity, Motivation, Self-Awareness and Emotion? *Arte Individuo Y Sociedad* Volume 35, No. 3.
- Yusa, I., Yu Yu, and Tetiana Sovhyra. (2022). Reflections on the Use of Artificial Intelligence in Works of Art.” in *Works of Art*, Yusa, IMM., Yu, Y., & Sovhyra, (Ed.), pp. 152–167.

# An AI-Driven User-Centric Framework Reinforced by Autonomic Computing: A Case Study in the Aluminium Sector

Ramon Angosto Artigues<sup>1</sup>, Andrea Gregores Coto<sup>1</sup>,  
Jonathan Josue Torrez Herrera<sup>2</sup>, Fernando Lou Tomás<sup>2</sup>,  
Sabrina Verardi<sup>3</sup>, Mattia Giuseppe Marzano<sup>3</sup>,  
and Andrea Fernández Martínez<sup>1</sup>

<sup>1</sup>AIMEN Technology Centre, 36418, O Porriño (Pontevedra), Spain

<sup>2</sup>Ibérica de Aleaciones Ligeras (IDALSA), 50668, Remolinos (Zaragoza), Spain

<sup>3</sup>Engineering Ingegneria Informatica S.p.A, 00144 - Roma (RM) - Italy

## ABSTRACT

The integration and deployment of AI in the industry faces several challenges, involving not only the need for robust and accurate AI models, but also their seamless integration with existing systems, while ensuring an intuitive user experience for workers. Furthermore, it is critical for AI solutions to be continuously managed for data governance, performance optimization, and the mitigation of risks, among other factors. This paper presents a service-oriented application that explores the integration of Machine Learning algorithms by adopting Human-in-the-Loop (HITL) strategies to enhance user-technology interactions in an Aluminium industrial environment. The proposed application exploits the use of data-driven Autonomic Computing techniques in AI Data Pipelines to promote the development of self-managed, adaptive systems that support dynamic interactions between technology and workers. Through the implementation of a web interface, workers are provided with seamless access to real-time data analysis and intelligent solutions within the user-empowered application.

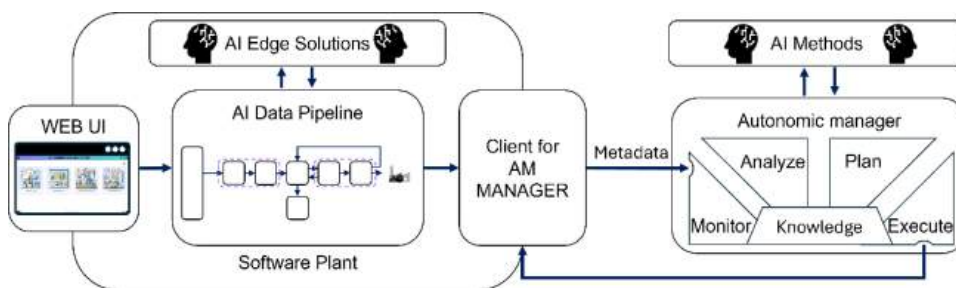
**Keywords:** Artificial intelligence, Autonomic computing, AI data pipeline, Human-centric design, Human-in-the-loop (HITL)

## INTRODUCTION

In today's rapidly evolving industrial landscape, the integration of Artificial Intelligence (AI) in the industry represents a paradigm shift, offering unprecedented opportunities to enhance efficiency, reduce environmental impact, and streamline operations, what is essential to maintain EU Industry competitiveness and sustainability (Mhlanga, 2022). Nevertheless, the complexities of modern industrial processes necessitate robust, flexible AI solutions that can be seamlessly integrated into existing infrastructures without disrupting current production workflows, while responding to the dynamics of the process, particularly under uncertain scenarios. To guarantee the successful adoption of AI technologies in industrial settings, it is essential

to keep the human in control and at the centre of all developments, actively involving workers in the design, implementation, and refinement of the technology. Indeed, AI systems have shown promising potential in supporting workers in their daily operational activities, enhancing their skills and knowledge of the process, and enabling faster-informed decisions (Rožanec et al., 2023), but their engagement with the technology is key to fully exploit the AI capabilities.

This paper presents a novel AI-driven application through the implementation of an AI Data Pipeline with Autonomic Computing capabilities to dynamically adapt to the needs of the industry, facilitating human-machine interactions and technology management, shown in Figure 1. The proposed user-centric application is integrated and evaluated in the Aluminium sector to demonstrate its ability to create a smart, human-centric production environment. By incorporating Human-In-The-Loop (HITL) strategies in a user-friendly interface, the proposed application enables a collaborative environment between humans and technology, ensuring that AI implementations are both effective and ethically sound.



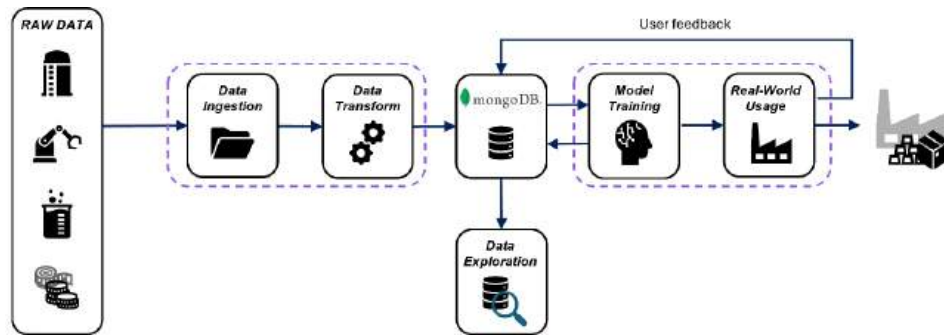
**Figure 1:** Diagram of the AI framework linking a data pipeline with an autonomic manager (AM) for industrial efficiency. The AM, based on the MAPE-K<sup>1</sup> model, leverages metadata to enhance autonomy.

## THE AI DATA PIPELINE

The integration of AI in industrial processes requires of methodical strategies to manage the flow of data through various computational tasks. Some of these tasks include the ingestion and transformation of data, as well as real-time data analysis. AI Data Pipelines are essential to ensure that data moves efficiently and reliably across different stages, transforming raw data into actionable information. Furthermore, this type of infrastructure must ensure the proper management and processing of datasets from different sources to enable AI applications managing industry factors—both internal and external—effectively (Krismentari et al., 2022). In this paper, we present an innovative AI Data Pipeline consisting of five main components that efficiently manages the end-to-end process of collecting, processing, and analysing data to train and deploy AI models effectively in industrial

<sup>1</sup>MAPE-K: (Monitor-Analyze-Plan-Execute-Knowledge)

scenarios. The pipeline has been designed and engineered with the aim of improving operational agility and performance in the industry. The AI Data Pipeline and the interrelations among components are depicted in Figure 2. The five components of the pipeline are described in more detail in the following subsections.



**Figure 2:** Workflow diagram illustrating the process from raw data to application, detailing components of pipeline: ingestion, transformation, MongoDB exploration, model training, and deployment.

**Data Ingestion:** The initial stage of the AI Data Pipeline involves collecting data from various sources, such as sensors, external databases, and/or third-party platforms and applications. At this stage, raw data is read and structured as necessary to prepare it for subsequent processing, ensuring its accessibility in a usable data format or schema. In industrial environments, customized ingestion procedures that cover the full spectrum of data might be required to handle the various data sources and formats.

**Data Transformation:** This component processes the ingested data, applying transformations such as normalization, filtering, and feature extraction to make the data suitable for data analysis. It is in this stage where data is merged and prepared to be stored in a proper data storage and management system e.g., MySQL, or MongoDB. Techniques to handle missing or incomplete data are implemented by this component as well e.g., inference and statistical methods from historical data.

**Data Exploration:** Data visualization and analysis are vital to enhance human knowledge of the process and its parameters, as well as elucidate intrinsic data patterns and analyze its evolution over time. This component can provide data insights in form of tables, charts, and/or pre-defined data analysis techniques e.g., via descriptive statistics and statistical inference methods.

**Model Training:** In this stage, ML models are trained, fine-tuned, and evaluated using the processed datasets from the Data Ingestion & Transformation components. Some of the functions comprised by the Model Training component includes the selection of appropriate ML algorithms according to the data nature and the problem addressed, the selection and fine-tuning of hyperparameters and architecture of the ML models, and

the evaluation of their performance to ensure that the specified criteria is accurately met.

**Real-World Usage (RWU):** Trained ML models can be deployed and integrated into production environments through various methods and platforms, such as cloud platforms, containers (Docker), or edge computing. In addition, monitoring ML models is essential to detect performance degradation or potential failures, leading to necessary re-training or model updates by the previous Model Training component. The RWU component often incorporates *feedback mechanisms* that enable continuous improvement for latest data inputs, as well as the integration of HITL strategies to further enhance the on-going refinement of AI models.

## **A HUMAN-CENTRIC SERVICE-ORIENTED APPLICATION**

For AI Data Pipelines to enable organizations extract insights and value from their data effectively, a seamless deployment and integration of the pipeline with the AI solutions is essential. A Service-Oriented Architecture (SOA) is an architectural approach wherein distinct, reusable software components, referred as services, communicate with each other over a network to achieve specific business functionalities (Schall et al., 2008). This architectural framework can facilitate the integration of AI Data Pipelines by presenting its various components as modular, reusable services. By following this approach, the system achieves loose coupling, thereby enhancing the flexibility and agility of the pipeline by facilitating independent development, deployment, and update of the individual components. The integration of the proposed AI Data Pipeline based on the SOA framework using APIs to expose services' functionalities provides a robust and scalable architecture that promotes modularity, interoperability, and the enforcement of security and monitoring mechanisms in the system. These characteristics are crucial to enable more effective HITL interactions through the simplification of complex processes into more manageable services. Moreover, this framework also enables the smooth integration of feedback mechanisms across the various services exposed, fostering continuous improvement and personalized user experiences.

In fact, HITL integration is imperative for any data-driven application to achieve full usability and functionality in real-world industrial settings, ensuring complete symbiosis between humans and Cyber-Physical Systems (CPS) (Adel, 2022). However, the transition towards the human-centric smart factory concept introduces challenges such as a prevalent skills gap where workers might lack the necessary IT expertise to effectively manage advanced systems (Tan et al., 2019). To this end, Autonomic Computing techniques can play a pivotal role closing this gap, providing technology with self-managing capabilities to mitigate the increasing complexity of computing systems (Gil et al., 2019). Additionally, the cognitive load imposed by evaluating large datasets, and the need for rapid decision-making under uncertain circumstances, can overwhelm workers without well-designed interfaces or adequate support tools. Hence, it is vital to prioritize human-centric strategies through the design and development of the technology to



ensure that it is not only intuitive and easy to use, but it is also tailored to the needs and preferences of users to foster their engagement, satisfaction, and productivity. Section *The AI-Driven Web Application In The Aluminium Use Case* provides more information about the proposed user-friendly web application.

## **AUTONOMIC COMPUTING FOR SELF-REGULATED AI SYSTEMS**

Autonomic Computing (AC) is a computing paradigm initiated by IBM in 2001 and originally inspired by the human autonomic nervous system that aims to develop computing systems with self-managing characteristics. The ultimate goal of this paradigm is to reduce human intervention in the rapidly growing complexity of software systems management, freeing humans from low-level management tasks, while still maintaining their central role in providing high-level guidance for their self-management (Parashar & Hariri, 2005).

A common framework to achieve self-management in AC is the MAPE (Monitor, Analyze, Plan, Execute) loop, which intends to provide systems with the necessary abilities to autonomously adapt to changing conditions, recover from failures, and optimize their performance. The Monitor module is responsible for continuously monitoring the environment and collecting relevant data about the system's state. This data is then processed, analyzed, and interpreted by the Analyze module to identify potential deviations or anomalies, as well as opportunities for the optimization of the system. Subsequently, based on the previous data analysis, strategies to maintain, adjust, and/or improve the system behaviour and attributes are determined by the Plan module. Ultimately, the planned actions are implemented as automated responses by the Execute module. The MAPE loop can be extended to the MAPE-K framework with the addition of the Knowledge component, which incorporates knowledge-driven reasoning and decision-making capabilities to the system to operate more intelligently and effectively. By incorporating the MAPE-K framework, the four main properties of AC systems can be realized as follows (Vizcarrondo et al., 2017).

**Self-configuration:** AC systems can automatically configure themselves based on their environment and requirements to adapt to changing conditions. For instance, AI algorithms can fine-tune their hyperparameters and model architectures based on the nature and characteristics of data, as well as the task to be addressed.

**Self-optimization:** AI systems can continuously monitor their performance and resource usage, adjusting their parameters to improve their efficiency. Optimization methods such as Bayesian optimization or evolutionary algorithms can be leveraged to fine-tune models. Other algorithms, such as in Reinforcement Learning, can directly learn, adapt, and optimize their strategies from their experiences in the environment. Ultimately, techniques such as re-training (continuous training) can be applied to existing models to incorporate information from new data without affecting the model architecture and parameters.

**Self-Healing:** Autonomic AI systems can be designed to automatically detect and diagnose failures and errors in real-time to execute corrective actions and restore the system functionality. Anomaly detection algorithms can detect unusual behaviour and data patterns by monitoring the system status. These algorithms cover statistical methods (e.g., z-score), ML methods like isolation Forest and one-class SVM, density methods (e.g., LOF), and state-of-the-art time series anomaly detection algorithms like Seasonal Hybrid ESD and Prophet.

**Self-Protection:** To protect systems from security threats and attacks, mechanisms to automatically mitigate risks and vulnerabilities can be implemented in AC systems. For this purpose, AI-powered cybersecurity systems and techniques such as adversarial training or anomaly detection algorithms can be deployed to detect, prevent, adapt, and recover from adversarial attacks. Regarding the performance of AI systems, adversarial training can also be used to improve AI robustness by preventing algorithms of being deceived by fake or perturbed data.

## **METADATA-DRIVEN AUTONOMIC MANAGER FOR AUTONOMIC COMPUTING CAPABILITIES IN THE AI DATA PIPELINE**

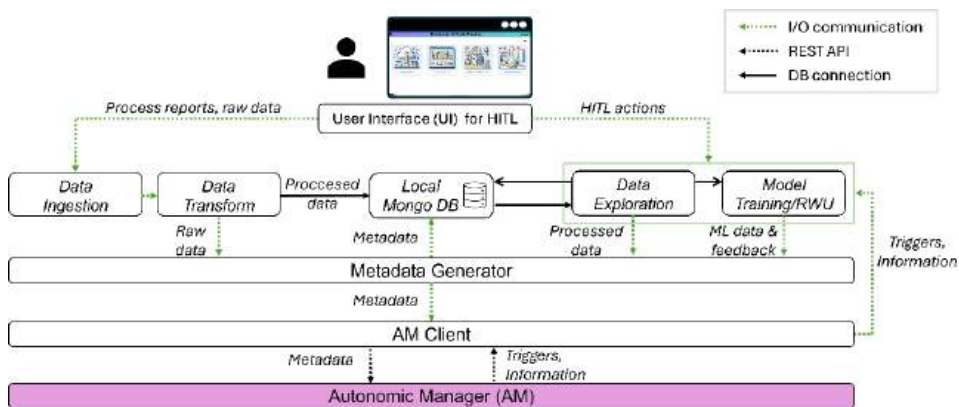
To effectively adopt the MAPE-K framework and implement the autonomous (self) abilities previously described -self-configuration, self-optimization, self-healing, and self-protection- the Autonomic Manager (AM) serves as the brain of the AC system. The AM can perform a variety of activities, including enforcing policies that govern the behaviour and operation of the AC system, monitoring and analysing the system to detect anomalies or opportunities for improvement, engaging in decision-making and real-time problem-solving processes, and consolidating learning strategies to enable continuous improvement and adaptation. In complex distributed systems, the AM can also coordinate and collaborate with other system components to exchange information, or coordinate actions (Vizcarrondo et al., 2017).

In the proposed framework, the AM assumes a central role serving as an autonomous AI Data Pipeline coordinator and decision-maker, thereby guaranteeing the AI Data Pipeline Governance. In this way, a continuous MAPE flow based on Knowledge of the AI Data Pipeline is implemented. The AM is responsible for continuously monitoring and analysing the state of the AI Data Pipeline to plan and execute corrective and optimization actions via the triggering of the autonomous (self) abilities. Thus, the implementation of the abilities is executed directly on the AI Data Pipeline for its self-management upon notification from the AM triggers.

As part of the practical implementation of our framework, the Monitoring module of the AM interacts with the AI Data Pipeline through an Orion Context Broker via APIs. The Analyze and Plan modules integrate a rule-based engine whose targets and goals are set by operators considering process requirements and specifications. Ultimately, the Execution module interacts with the AI Data Pipeline by triggering the autonomous (self) abilities when needed, as well as by sending notifications and alarms when abnormal behaviour is detected. To complement the rule-engine, AI methods based

on temporal series analysis combined with statistical measures are also available for anomalies detection. The temporal analysis considers weekly- and monthly- time windows of data for analysing the evolution of specific parameters over time. Figure 3 depicts the communication flow between the different elements of our framework.

For a proper analysis, planning, and execution of corrective and optimization actions, it is crucial to properly select the data that is shared from the AI Data Pipeline to the AM. One of the key innovations of the proposed framework is the use of metadata from the components of the AI Data Pipeline to better abstract their behaviour and performance.



**Figure 3:** Architecture of the HITL system for autonomous computing, illustrating the flow from data ingestion to model training and deployment, with a local MongoDB for storage and exploration.

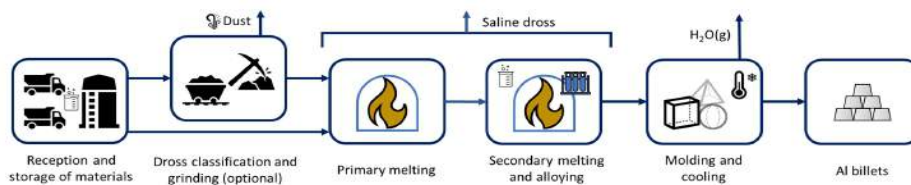
By providing metadata the AM can get additional context about the performance of the underlying components, enabling a higher-level conceptualization and understanding. Metadata is based on descriptive statistics of new raw data. Each component of the pipeline can share its own set of metadata with the AM, hence the distributed nature of the framework. As an example, the Data Exploration component might send the average and standard deviation of parameters, e.g., gas and oxygen consumption, whereas the Model Training component might send the errors obtained during training e.g., MAE. Furthermore, to guarantee that users are at the centre of all developments, human feedback is sent to the AM as part of the data. The rule-engine of the AM includes rules assessing the conformity of humans with the established AI solutions, serving as triggering events that activate the autonomous (self) abilities on these solutions.

### THE AI-DRIVEN WEB APPLICATION IN THE ALUMINIUM USE CASE

Although the proposed framework based on the AI Data Pipeline with Autonomous Computing abilities has been designed and engineered as a technology-agnostic, domain-independent solution, a preliminary integration in a real (recycled-based) aluminium industrial setting has been performed

as a validation of functionality and assessment of performance to enable its continuous improvement.

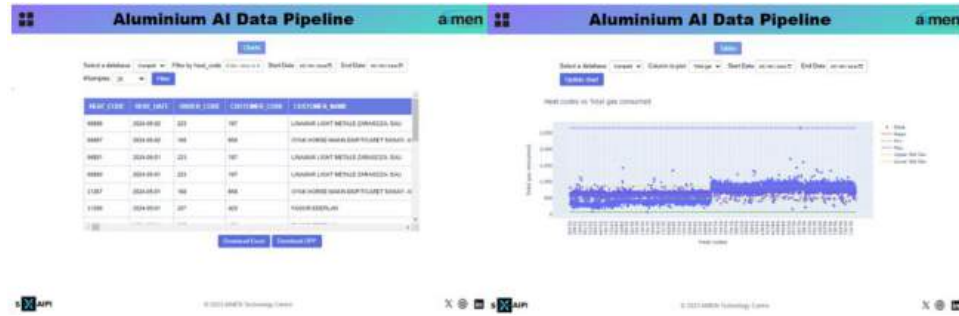
The recycled aluminium-making process starts with the reception of materials from different sources. Upon reception, materials are sampled for proper sorting and storage based on their chemical composition and characteristics. Based on customer orders on products, which are characterized by the product norms and their quantities, scraps are selected to be processed via primary and secondary melting. This selection is based on the chemical composition of scraps, their availability at the plant, and their cost. In secondary melting, alloys can be introduced to refine the melted mixture prior its moulding and cooling to produce the final aluminium products. The different stages of the process are depicted in Figure 4.



**Figure 4:** Schematic diagram of the recycled aluminium production process: from material reception to final product formation.

The AI solutions developed within the project aim to support operators in the decision-making process of scraps and alloys (aluminium recipe) via a Random Forest Regressor for the estimation of the chemical composition of aluminium mixtures, and a Generative Reinforcement Learning Framework to generate new recipes based on customer requests as a multi-criteria optimization problem. The details of the AI solutions are out of the scope of this paper.

To facilitate the interaction between users and the AI Data Pipeline and the Autonomic Manager, a web-based interface has been developed following a human-centric design strategy. By incorporating the targeted users in the design and development phases of the interface, the application was unequivocally built around the users' needs, preferences, and behaviours. There are two components of the AI Data Pipeline that hold special significance in terms of HITL interactions: Data Exploration and RWU components. The Data Exploration component can be accessed through two different web pages according to the user preferences. On one hand, users can access data in form of tables, filtering the available datasets according to their origin, heat process code, dates, and number of samples shown. In addition, tables can be downloaded in *.csv* or *.xlsx* format. On the other hand, users can visualize the evolution of a set of pre-defined parameters through charts. Currently, these charts also display statistical measures of the data shown for visualization e.g., mean and standard deviation. Figure 5 shows the two discussed pages of the web application for the Data Exploration component.



**Figure 5:** Data exploration interfaces in the web application: tabular and graphical representations for comprehensive analysis of industrial aluminium production parameters.

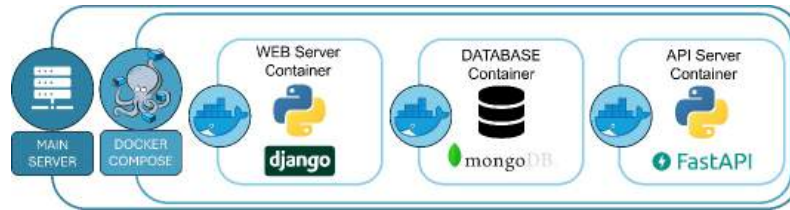
The RWU component can also be accessed through two different pages according to the AI solution to be invoked, as shown in Figure 6. To collect human feedback on the performance of those solutions, different strategies have been implemented. In the case of the AI solution that estimates the chemical composition of scrap-based mixtures, a thumbs-up/thumbs-down button is available for acquiring the overall satisfaction of users with the predictions. For the recipe generation algorithm, a multi-criteria evaluation via a 5-star scoring is integrated to allow users assess the recipes according to the final chemical composition estimated, the scraps used, and the overall cost. As previously described, human feedback is used as a satisfaction score that can trigger the activation of the self-abilities through the rule-engine of the AM to optimize the AI solutions in the pipeline.



**Figure 6:** Interactive AI solutions in aluminium production: incorporating user feedback for chemical composition prediction and recipe generation.

The front-end technologies for the web interface development are HTML, CSS, and JavaScript. For the back-end framework, Django is used as the backbone that handles user authentication, database management (MongoDB), and web-server logic. To isolate potential problems with the AI solutions, a FastAPI server has been deployed, which is also responsible for handling communications with the Orion Context Broker of the AM. The complete application is deployed on a private server configured with Docker

compose to guarantee that all software dependencies and requirements are met. The presented set-up, depicted in Figure 7, ensures a seamless interaction and easy deployment of all the elements of the application, enabling a smooth communication with the client side and real-time data processing and analysis.



**Figure 7:** Dockerized web application architecture: integrating Django, MongoDB, and FastAPI for efficient data management and real-time processing.

## CONCLUSION

This paper presents a novel AI-driven application embedding an AI Data Pipeline with Autonomic Computing capabilities to strategically integrate advanced AI technologies in real industrial settings. By employing a SOA that facilitates the modular integration of AI components, the various services can be developed and deployed independently, enhancing the flexibility and scalability of the system. This approach also enables the integration of human feedback loops to refine the AI functionalities, guaranteeing the system reliability and AI robustness. By facilitating an intuitive interface for workers following a user-centric design, the system simplifies human-machine interactions leveraging real-time data analysis for better decision-making support. The preliminary integration of the proposed framework through a web application in the Aluminium sector has proven the system to be robust and capable of real-time adaptation and optimization.

## ACKNOWLEDGMENT

This work has been supported by the project “self-X Artificial Intelligence for European Process Industry digital transformation” (s-X-AIPI), which has received funding from the European Union’s Horizon Europe research and innovation programme under grant agreement No. 101058715.

## REFERENCES

- Adel, A. (2022). Future of industry 5.0 in society: human-centric solutions, challenges and prospective research areas. *Journal of Cloud Computing* 2022 11:1, 11(1), 1–15. <https://doi.org/10.1186/S13677-022-00314-5>
- Gil, M., Albert, M., Fons, J., & Pelechano, V. (2019). Designing human-in-the-loop autonomous Cyber-Physical Systems. *Int. J. Hum. Comput. Stud.*, 130, 21–39. <https://doi.org/10.1016/J.IJHCS.2019.04.006>

- Krismentari, et al (2022). Data Pipeline Framework for AIS Data Processing. 2022 7th International Conference on Informatics and Computing, ICIC 2022. <https://doi.org/10.1109/ICIC56845.2022.10006941>
- Mhlanga, D. (2022). Human-Centered Artificial Intelligence: The Superlative Approach to Achieve Sustainable Development Goals in the Fourth Industrial Revolution. *Sustainability*, 14(13). <https://doi.org/10.3390/SU14137804>
- Parashar, M., & Hariri, S. (2005). Autonomic computing: An overview. *Lecture Notes in Computer Science*, 3566, 257–269. [https://doi.org/10.1007/11527800\\_20](https://doi.org/10.1007/11527800_20)
- Rožanec, et al. (2023). Human-centric artificial intelligence architecture for industry 5.0 applications. *International Journal of Production Research*, 2023(20), 6847–6872. <https://doi.org/10.1080/00207543.2022.2138611>
- Schall, D., Truong, H. L., & Dustdar, S. (2008). Unifying human and software services in web-scale collaborations. *IEEE Internet Computing*, 12(3), 62–68. <https://doi.org/10.1109/MIC.2008.66>
- Tan, Q., et al. (2019). Anthropocentric Approach for Smart Assembly: Integration and Collaboration. *Journal of Robotics*, 2019. <https://doi.org/10.1155/2019/3146782>
- Vizcarrondo, J., Aguilar, J., Exposito, E., & Subias, A. (2017). MAPE-K as a service-oriented architecture. *IEEE Latin America Transactions*, 15(6), 1163–1175. <https://doi.org/10.1109/TLA.2017.7932705>

# Survey of Research Issues and Proposed Solutions for Detecting Parameter Anomalies in System Logs

Hironori Uchida<sup>1</sup>, Keitaro Tominaga<sup>2</sup>, Hideki Itai<sup>2</sup>, Yujie Li<sup>1</sup>,  
and Yoshihisa Nakatoh<sup>1</sup>

<sup>1</sup>Kyushu Institute of Technology, 1-1 Sensuicho, Tobata-ku, Kitakyushu-shi, Fukuoka Prefecture, Japan

<sup>2</sup>Panasonic System Design Co., Ltd., 3-1-9, Shinyokohama, Kohoku-ku, Yokohama-shi, Kanagawa Prefecture, Japan

## ABSTRACT

In the ever-evolving field of software development, the demand for automation of fault analysis that is time-consuming and expertise-requiring is growing. One solution to this challenge is the study of anomaly detection using text logs, which has seen numerous research efforts. However, despite the variety of patterns that system anomalies can exhibit, many studies have predominantly focused on sequence anomalies. This is largely attributed to the limited availability of datasets, with the commonly used Loghub data being oriented towards sequence anomalies. This research addresses the current challenges in anomaly detection models and proposes several new methods for detecting parameter anomalies. Initially, due to the lack of datasets of parameter anomalies, we prepared common parameter anomaly scenarios and compared them with existing sequence anomaly detection models (including DNN models for sequence anomalies and DNN models using semantic information), and with a variety of proposed methods. The prepared parameter anomaly patterns include four Integer types and three String types. For instance, a parameter within a certain range (-100 to 100) is considered normal, while parameters outside this range are deemed anomalies. Our proposed method begins by extracting parameters using LogParser and determining whether they are of Int or String type. For Int types, we use Z-Score, IQR, K-NN and DBSCAN for evaluation, while for String types, we use a Bert-based positive-negative classifier. The experimental results showed that the DNN model for sequence anomaly had an F1 Score of less than 0.5 for all patterns. In contrast, our proposed methods achieved F1 Scores exceeding 0.9 or 0.8 for almost all methods, except for one anomaly pattern. It was found that the proposed methods are effective for common parameter anomaly problems. Furthermore, since our methods do not require prior training, they are particularly advantageous for ad-hoc learning in the context of continuously updated software development.

**Keywords:** Anomaly detection, Software log, Log analysis, Deep learning, Log generator, Parameter anomaly detection



## INTRODUCTION

In the constantly evolving field of software development, there is a pressing need for the automation of complex and time-consuming fault analysis tasks that require specialized knowledge. As one solution, anomaly detection using textual system logs has been extensively studied. However, despite the variety of anomaly patterns that can occur within systems, much of the research has focused on sequence anomalies. A significant reason for this focus is the limited number of dataset types available, with commonly used datasets like Loghub (He et al., 2020) being tailored for sequence anomalies. The issue of limited dataset variety has been reported in several studies (Le et al., 2022; Shilin et al., 2022, Uchida et al., 2023). Moreover, current log anomaly detection models face the issue of losing parameter information during preprocessing with LogParsers like Drain (He et al., 2017) and Brain (Yu et al., 2023), which template the logs. Consequently, some studies have explored the use of parameter sequence vectors in addition to sequence vectors (Zhu et al., 2020) and research utilizing sentence information without LogParsers (Le et al., 2021). While these studies focused on parameters, the datasets used were designed for sequence vectors and lacked specific anomaly labels for each parameter. In response to these challenges, this research aims to automatically generate datasets with typical parameter anomalies to create an environment suitable for evaluating parameter anomaly detection. Using these datasets, we will examine the current limitations of anomaly detection models and propose several new methods for detecting parameter anomalies. The parameter anomaly patterns prepared include four Integer types and three String types, such as considering parameters within a certain range (-100 to 100) as normal and those outside this range as anomalies.

### ■ Log Format

<Label> <Timestamp> <Date> <Node> <Time> <NodeRepeat> <Type> <Component> <Level> <Content>

### ■ Explanation of How to Create Log Formats

Format	How to Create	Format	How to Create
Label	"-" or "INTERR" or "STRERR"	NodeRepeat	Constant Value: "R02-M1-N0-C:J12-U11"
Timestamp	Constant Value: "1111111111"	Type	Constant Value: "RAS"
Date	Constant Value: Date Created	Component	Constant Value: "KERNEL"
Node	Constant Value: "R02-M1-N0-C:J12-U11"	Level	Constant Value: "INFO"
Time	16 lines of logging per 0.5 second	Content	Automatic Generation

### ■ Example of Generated Logs (only the variable portion is excerpted)

```
- 11111111 2024-03-10-11.46.46.050249 7wBfcbp2 r7fzP bknkj OKsGrT
- 11111111 2024-03-10-11.46.46.050249 pAxmWZ MKV7wpz VIBQuR jIT7NC 2UrYobZZaP eqv60i
- 11111111 2024-03-10-11.46.46.050249 e9JdJrsoS y5xiwbE AOuo4 KeMwOHL Oe2fHLR HmGYI
- 11111111 2024-03-10-11.46.46.050249 7wBfcbp2 r7fzP bknkj OKsGrT
- 11111111 2024-03-10-11.46.46.050249 y7mM eYgD9d0X cv5RRwva FRv15V UmU2rxY
INTERR 11111111 2024-03-10-11.46.46.050249 jPShEtW E3bPwaKYw OukP9NIR RXcupA oEnf -8159
- 11111111 2024-03-10-11.46.50.050249 jPShEtW E3bPwaKYw OukP9NIR RXcupA oEnf -36
- 11111111 2024-03-10-11.47.03.050249 jPShEtW E3bPwaKYw OukP9NIR RXcupA oEnf -53
```

**Figure 1:** Description of the log format used for log generation and some examples.

This research seeks to clarify the following:

RQ1: Accuracy of current anomaly detection models against typical parameter anomalies.

RQ2: Proposals for new methods tailored to parameter anomaly detection.

## DATASET FOR PARAMETER ANOMALY DETECTION

Due to the absence of readily available datasets for parameter anomaly detection, we created a custom algorithm for automatic generation. Please refer to Github for specific implementations <https://github.com/hiro877/LogGenerator/tree/ihiet2024>.

**Table 1.** Parameters used to generate logs.

Parameter	Description
Total number of logs	40000+1200(parameter logs)
Number of words of content	Range: 3~6
Number of words in content	Range: 4~10
Number of logs per second	15
Number of log types	16+1(parameter log type)
Histogram of logs	[14483, 1742, 1738, 1736, 1727, 1726, 1726, 1721, 1702, 1700, 1687, 1686, 1682, 1660, 1644, 1640]
Number of parameter logs	1200: Approx. 3% of total number of logs
Number of parameter logs	240: Approx. 20% of parameter logs

### Automatic Log Generation Algorithm

Figure 1 illustrates the log format used for automatic generation and provides examples of the generated logs. The log format is based on the BGL (Blue Gene/L supercomputer log) L from Loghub. Considering that only the content part of the logs is important for a dataset aimed at parameter anomaly detection, this log generator was designed to create logs with variables for labels and content parts.

Next, we introduce the log generation system. Table 1 presents the key variable parameters used within the log generation system. These parameters were selected based on research that investigated differences in complexity between research datasets and actual logs in development environments (Uchida et al., 2024). The parameters used in this study are as shown in Table 1, which are approximately 1/10th of the corresponding parameters in the BGL within Loghub.

The strings for the log content part were randomly generated.

### Parameter Log Pattern

In this experiment, we prepared seven commonly considered parameter anomaly patterns, including three Integer types and four String types. Figure 2 shows the pattern of parameter anomalies in this experiment.

## EVALUATED MODELS

In this experiment, we utilized NeuralLog and PLELog as representative models of sequence anomaly detection due to their high accuracy. The reason for their selection was based on the assessment that they have a higher potential for addressing parameter anomalies compared to other models. Both models were chosen because their use of semantic information was deemed likely to enable the detection of parameter anomalies, in contrast to other sequence vectors that exclude parameters. Furthermore, we proposed four methods capable of detecting parameter anomalies and compared the accuracy of each model. The following sections will introduce each model in detail.

### NeuralLog

NeuralLog is a novel approach that does not necessitate log parsing. It directly converts log messages into semantic vectors and employs a Transformer model to detect anomalies. This method extracts the semantic meaning of log messages using BERT and identifies anomalies through a Transformer-based classification model. However, there is a process that deletes data if the split strings are numerical during the operation of splitting the logs. Therefore, it is anticipated that NeuralLog will perform poorly in detecting integer type parameter anomalies in this experiment.

Int Type	Normal	Anomaly	
Param 1	Range: -100 ~ 100	Out of Range: -10000 ~ -101    101 ~ 10000	
Param 2	Range such as Registration ID	Out of Range: 30 ~ 10000	
Param 3	Random Specific Value: 30 values in the range 0 ~ 10000	Non-Specific Values in the Range 0 ~ 10000	
Param 4	Non-Specific Values in the Range 0 ~ 10000	Random Specific Value: 30 values in the range 0 ~ 10000	
String Type	Normal	Anomaly	Note
Param 5	"true"	"false", "null", "error"	Words for which positive negatives can be determined
Param 6	"connected"	"disconnected"	
Param 7	"MODE_0", "MODE_1", "MODE_30" ~ ..., "MODE_29")	"MODE_10000"	Words for which a positive negative cannot be determined
Param8	World Countries	World Cities	Same as above

**Figure 2:** Anomaly parameter patterns.

### PLELog

PLELog adopts a semi-supervised approach that uses only the labels of known normal log sequences to estimate the labels of mixed anomalous and normal log sequences through probabilistic label estimation. This allows for leveraging the advantages of a supervised approach while saving time on manual labeling. Additionally, it deals with unstable log data by using semantic embeddings and an attention-based (Gated Recurrent Unit) GRU neural network to efficiently and effectively detect anomalies. The steps for creating semantic embeddings include 1) log parsing, 2) word embedding,

and 3) aggregation based on TF-IDF. During the word embedding process, non-text tokens (such as delimiters, operators, punctuation, and numbers) are removed. Therefore, it is predicted that PLELog will perform poorly in detecting integer type parameter anomalies in this experiment.

### Proposed Methods

Initially, parameters are extracted from raw logs using Drain. Next, it is determined whether the extracted parameter strings are of type integer or string. If the type is Integer, anomaly detection is performed using Z-score, Interquartile range (IQR), k-Nearest Neighbour (k-NN), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). For String types, anomaly detection is conducted using a combination of BERT and Transformer for positive-negative judgment.

$$Z - \text{score} = (X - \mu) / \sigma \quad (1)$$

X represents an individual data point,  $\mu$  is the mean of the dataset, and  $\sigma$  is the standard deviation of the dataset. Data is considered anomalous when the Z-Score exceeds a threshold of 2.

IQR: The Interquartile Range (IQR) is a measure of variability based on dividing a dataset into quartiles. The IQR is defined as  $Q3 - Q1$ , and data points lying outside  $Q3 + 1.5 \times IQR$  or  $Q1 - 1.5 \times IQR$  are considered outliers.

**Table 2.** Experimental results.

F1-Score	PLELog	NeuralLog	Z Score	IQR	k-NN	DBSCAN
Param 1	0.385	0.000	0.655	0.998	0.821	0.947
Param 2	0.356	0.000	0.617	1.000	0.779	0.902
Param 3	0.177	0.000	0.000	0.000	0.818	0.912
Param 4	0.345	0.000	0.000	0.000	0.000	0.000
F1-Score	PLELog		NeuralLog		DistilBert	
Param 5	0.256		0.000		1.000	
Param 6	0.320		0.000		1.000	
Param 7	0.384		0.000		0.834	
Param 8	0.343		0.000		0.061	

k-NN: For each data point, the distance to the second nearest neighbour is obtained. If this distance is greater than twice the average distance to the second nearest neighbour, the data point is considered anomalous.

DBSCAN: In DBSCAN, points that do not belong to any cluster are considered anomalous.

DistilBert: DistilBERT is a smaller, faster, and lighter Transformer model based on the BERT architecture. We used the DistilBERT model fine-tuned on the Stanford Sentiment Treebank (SST-2) dataset for sentiment analysis tasks. Items judged as Negative by this model are considered anomalies.

## EXPERIMENTAL METHOD

### Sequence Anomaly Detection Models Input

The input utilized sequence data with Window = 20 and Slide = 1.

### Dataset Split Ratio

For NeuralLog and PLELog, which are learning-based methods, the dataset was divided into Training = 0.6, Validation = 0.3, and Test = 0.1.

On the other hand, the proposed method, which does not use learning, created distances and classes based on all logs and detected anomalous data.

### Accuracy Evaluation Method

In the accuracy comparison, the accuracy of anomaly detection on the test data is verified using each model after training. Each model is evaluated for classification performance using the F-measure value; The F-measure is an evaluation index that indicates the balance between detection accuracy and the number of anomaly detections. Here, the F-measure is computed as follows.

$$\begin{aligned}
 \textit{Precision} &= \frac{TP}{TP+FP} \\
 \textit{Recall} &= \frac{TP}{TP+FN} \\
 \textit{F-measure} &= \frac{2 \cdot \textit{Precision} \cdot \textit{Recall}}{\textit{Precision} + \textit{Recall}}
 \end{aligned} \tag{2}$$

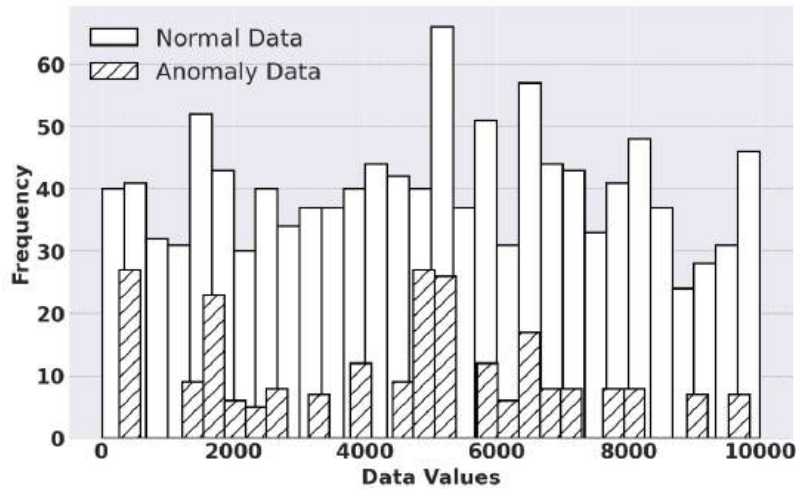
Where,

TP: Anomaly instances correctly classified by the model

TN: Normal instances correctly classified by the model

FP: Normal instances misclassified by the model

FN: Abnormal instances misclassified by the model



**Figure 3:** The histograms for the datasets used in Param 4.

## EXPERIMENTAL RESULTS

We conducted an accuracy evaluation of NeuralLog, PLELog, and the proposed method on the dataset with each parameter anomaly pattern shown in Figure 2. The experimental results are presented in Table 2. Looking at the results, it is evident that both NeuralLog and PLELog have low accuracy for both Int and String types. On the other hand, the proposed method shows high-accuracy results, except for the experiment with Param 4. Particularly, DBSCAN shows high accuracy for Integer types, and Bert shows high accuracy for String types. It is understandable that NeuralLog and PLELog would exhibit decreased accuracy due to their processes that exclude numerical parameters. However, the low accuracy with String-type data, which are not excluded, indicates that these models struggle with detecting parameter anomalies for each line. It is important to note, though, that these models are specialized for sequence anomalies.

Next, we consider the experiment with Param4. The histograms for the datasets used in Param 4 are shown in Figure 3. While the parameters for Normal and Anomaly in Param4 do not overlap, they are randomly selected from the same range of values. As a result, it is thought that classification based on clustering or distance did not function effectively.

Moreover, it is understandable why Param5 and Param6, which conform to positive-negative judgment, would have high accuracy. Additionally, it makes sense that Param8, which does not adhere to positive-negative judgment, shows low accuracy. Contrarily, the high accuracy observed in Param7 can be attributed to the influence of the value following “MODE\_”. However, it is necessary to note that in this instance, since anomalies are defined as cases with high values, the method will not function in situations that oppose this definition. However, it should be noted that this time, since high values were defined as anomalies, the approach will not work in cases that contradict this definition.

## CONCLUSION

In this study, we created a dataset for parameter anomaly detection using our custom log generator and conducted an investigation into the challenges of parameter anomaly detection for current anomaly detection models, which had not been thoroughly explored before, and examined proposed methods. As a result, we found that sequence anomaly detection models struggle to detect parameter anomalies on a per-line basis. On the other hand, we demonstrated the potential to detect anomalies by extracting parameters and using parameter-specific machine learning methods, as done with the proposed methods. However, we also discovered anomaly patterns that cannot be addressed with simple machine learning methods, indicating the need to explore various approaches, including specialized DNN models for parameter anomalies. We hope that there will be an increase in research using parameter anomaly log generators like the one we have created.

## ACKNOWLEDGMENT

This work is supported by a grant from Panasonic System Design.

This work was supported by JST, Kyutech Research Fellowship, Grant Number JPMJFS2133.

## REFERENCES

- Bei Zhu, Jing Li, Rongbin Gu, and Liang Wang., (2020) “An Approach to Cloud Platform Log Anomaly Detection Based on Natural Language Processing and LSTM”, ACAI ‘20: Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, pp. 1–7, <https://doi.org/10.1145/3446132.3446415>.
- He. P, Zhu. J, Zheng. Zm and Lyu. M, (2017) “Drain: An Online Log Parsing Approach with Fixed Depth Tree”, 2017 IEEE International Conference on Web Services (ICWS).
- He. S, Zhu. J, He. P, R. M, and Lyu. M, (2020) “Loghub: A Large Collection of System Log Datasets towards Automated Log Analytics”, Arxiv Website: <https://arxiv.org/pdf/2008.06448.pdf>.
- Le. V and Zhang. H, (2022) “Log-based anomaly detection with deep learning: how far are we?”, ICSE ‘22: Proceedings of the 44th International Conference on Software Engineering, pp. 1356–1367.
- Le. V-H., Zhang. H., (2021) “Log-based Anomaly Detection Without Log Parsing”, in Proceedings of the 2021 36th IEEE/ACM International Conference on Automated Software Engineering (ASE), pp. 492–504, doi: 10.1109/ASE51524.2021.9678773.
- Shilin. H., Xu. Z., Pinjia. H., Young. X., Liqun. L., Yu. K., Minghua. M., Yuning. W., Yngnong. D., Sarabanakumar. R., & Qingwei. L., (2022 November 9) “An empirical study of log analysis at Microsoft”, ESEC/FSE 2022: Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 1465–1476, <https://doi.org/10.1145/3540250.3558963>.
- Tree, IEEE Transactions on Service Computing, vol. 16, issue 5, pp. 3224–3237, doi: 10.1109/TSC.2023.3270566.
- Uchida. H, Tominaga. K, Itai. H, Li. Y, and Nakatoh. Y., (2023) “Verification of Generalizability in Software Log Anomaly Detection Models”, Anomaly Detection - Recent Advances, AI and ML Perspectives and Applications, doi: 10.5772/intechopen.111938.
- Uchida. H, Tominaga. K, Itai. H, Li. Y, and Nakatoh. Y., (2024) “Differences between research log datasets and development field logs and creation of complexity evaluation index”, Pertanika Journal, under review.
- Yu. S, He. P, Chen. N, and Wu. Y, (2023) “Brain: Log Parsing with Bidirectional Parallel”.

# Predicting Key Substance Levels in Aquaculture Through AI-Based Water Quality Monitoring

Ji-Yeon Kim<sup>1</sup>, Ki-Hwan Kim<sup>2</sup>, Young-Jin Kang<sup>2</sup>,  
and Seok-Chan Jeong<sup>1,2,3</sup>

<sup>1</sup>Department of Artificial Intelligence, Dong-Eui University, Busan, 47340, South Korea

<sup>2</sup>AI Grand ICT Research Center, Dong-Eui University, Busan, 47340, South Korea

<sup>3</sup>Department of e-Business, Dong-Eui University, Busan, 47340, South Korea

## ABSTRACT

Land-based aquaculture farms use seawater transported from nearby seas instead of large amounts of freshwater. A seawater recirculating filtration system is essential for sustainable fish farming; however, this system has limitations in improving the levels of ammonia, nitrite, and nitrate, which are directly linked to fish mortality. Therefore, most land-based aquaculture farms periodically exchange a certain amount of seawater to maintain optimal water quality. Despite these efforts, managing water quality remains a significant challenge due to the fluctuating levels of these harmful substances. This study aims to address this challenge by predicting the levels of ammonia, nitrite, and nitrate—the primary causes of fish mortality in land-based aquaculture—using AI models. The training data were collected from various sensors installed in the farms, including those measuring water temperature, dissolved oxygen, dissolved solids, pH level, oxidation-reduction potential, salinity, nitrate, and ammonia. By leveraging this comprehensive dataset, we evaluated the performance of multiple models, such as Random Forest (RF) and K-Neighbors Regressor (KNN). The study demonstrated that these models could achieve remarkable performance metrics, with the Random Forest model recording an MAE of 0.0150, MSE of 0.0008, RMSE of 0.0289,  $R^2$  of 0.9999, RMSLE of 0.0039, and MAPE of 0.0024. Such high accuracy levels indicate that AI-based water quality prediction models have significant potential for effectively monitoring and predicting fish health in aquaculture farms. Implementing these AI models can lead to more proactive and precise management of water quality, ultimately reducing fish mortality rates and enhancing the sustainability and profitability of aquaculture operations.

**Keywords:** Water quality prediction, Aquaculture, Fish, AI, Monitoring

## INTRODUCTION

Aquaculture in underwater farms is increasingly gaining importance globally as a response to the rising demand for seafood. However, securing optimal conditions for fish growth and health within aquaculture facilities continues to be a challenging task. This is particularly true in land-based facilities, where there is an absolute dependence on seawater circulation and filtration systems.



Land-based aquaculture requires monitoring and measuring various water quality parameters such as pH, temperature, and dissolved oxygen (DO) (Tziortzioti et al., 2019; Mwegoha et al., 2010). Dissolved oxygen, the most critical factor for fish survival, does not dissolve well in water, and its solubility decreases sharply as temperature and salinity increase (Timmons et al., 2018; Chumkiew et al., 2019). DO affects fish growth and feed efficiency (Buentello et al., 2000; Zhang et al., 2011). Fluctuations or suboptimal temperatures can affect feed utilization and cause stress in fish when grown at temperatures higher or lower than ideal, potentially leading to disease (Timmons et al., 2018).

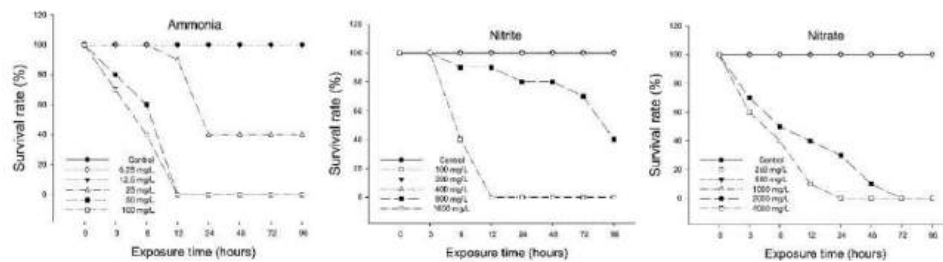
The Recirculating Aquaculture System (RAS) provides advantages for efficiently managing fish growth by maintaining appropriate water temperatures and optimal oxygen saturation levels, as well as purifying impurities. Physical and biological filters in the RAS system reduce concentrations of nitrogen compounds such as ammonium and nitrites, which are the existing alternatives to diluting seawater. Ammonium and nitrites, being toxic to fish (Timmons et al., 2018), are critical components that must be managed in aquaculture. Ammonia, a key element of ammonium and nitrites, occurs naturally from the decomposition of fish waste and leftover feed but is highly toxic (Wicks et al., 2002), weakening the gills and potentially causing death. RAS converts ammonia into less toxic nitrites and further into nitrates, thus increasing the water reuse rate (Suurnäki et al., 2020). The standards for recycled water vary by fish species, but maintaining nitrates below 100 mg/L for Atlantic salmon has been shown not to significantly affect their growth and health (John Davidson et al., 2017). Chronic exposure to NO<sub>3</sub> in turbot has been found to cause persistent toxicity leading to death, particularly causing methemoglobinemia, ion homeostasis disruption, lipid peroxidation, and abnormal cell apoptosis in flounder (Jiachen Yu et al., 2021). Additionally, attention needs to be paid to the accumulation of metabolites such as cortisol and testosterone, particularly steroids (Mota et al., 2017). It has been indicated that highly toxic hydrogen sulfide (H<sub>2</sub>S) can be produced during the RAS filtration process, and early detection is crucial (Salim et al., 2023).

Efforts to predict water quality using AI are continuously being made. Experiments using convolutional neural networks (CNN) and long short-term memory (LSTM) have been conducted to predict basic water quality information (Haq et al., 2022). These experiments have forecasted fundamental water quality data. Common basic water quality parameters such as temperature, pH, and dissolved oxygen (DO) can be measured using portable devices like the Multi 3410 (WTW GmbH) (Chun et al., 2018). While portable alternatives are available, nitrates, nitrites, and phosphates are often more accurately measured in laboratories using ion exchange chromatography and suppressed conductivity detectors (e.g., Dionex DX-500, Dionex ICS1600, Dionex Integration HPIC; Chun et al., 2018; Lindholm-Lehto et al., 2020, 2021).

This paper investigates an AI model that infers the presence of substances such as ammonia, nitrites, and nitrates, which significantly impact fish mortality, using basic sensor data.

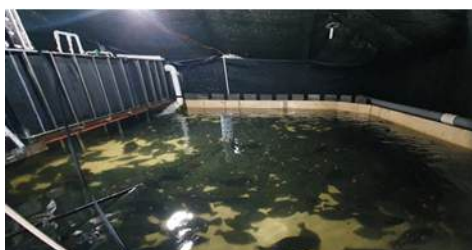
## EXPERIMENTS ENVIRONMENT

The experimental species is the olive flounder, and it has been observed that there is no mortality rate when ammonia levels are up to 12.5 mg/L, but survival rates begin to decrease when exposed to 25 mg/L for more than 12 hours (Kim et al., 2019). Similarly, nitrite and nitrate concentrations also impact survival rates. For nitrite, a significant decrease in survival rate was observed at 800 mg/L (40% survival), and no fish survived 12 hours later at 1600 mg/L. There was no mortality under nitrite concentrations of 100, 200, and 400 mg/L. For nitrate, survival rates decreased at concentrations above 2000 mg. No fish survived after 72 hours at this concentration. However, no mortality was observed at control concentrations and at levels below 1000 mg/L.



**Figure 1:** Survival rate (%) of flatfish, *paralichthys olivaceus* exposed to ammonia, nitrite, and nitrate for 96 h (Kim et al., 2019).

An experiment was conducted to observe changes in water chemical concentrations in an indoor aquaculture facility raising olive flounder. Approximately 300 fish were cultured (see Figure 2(a)), and significant variations in water concentrations were evident with fish growth. The recirculating aquaculture system (RAS) utilized in the facility includes a module combining a 40-ton drum filter, a skimmer, and a biological filtration unit, as depicted (see Figure 2(b)).



(a)



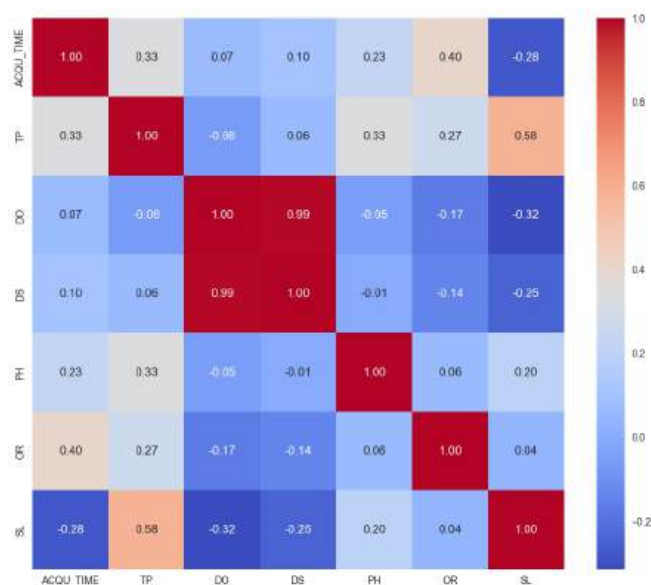
(b)

**Figure 2:** Integrated overview of aquaculture density and recirculating aquaculture system components.

Sensors were installed in the aquaculture facility to facilitate real-time monitoring of water temperature, dissolved oxygen, total dissolved solids, acidity, oxygen redox potential, and salinity, while ammonia levels were measured manually. Experiments were conducted to identify the optimal water quality conditions for the growth of flatfish, considering factors such as temperature, salinity, and pH.

## EXPERIMENTS

The heatmap illustrates the correlation coefficients between various water quality parameters. Each cell represents the correlation between two variables, with color intensity indicating the strength of the relationship, ranging from blue (negative correlation) to red (positive correlation). Subsequently, various water quality data such as Total Phosphorus (TP), Dissolved Oxygen (DO), Total Dissolved Solids (DS), acidity (pH), Oxygen Redox Potential (OR), and Salinity (SL) were collected, and the correlation analysis results (see Figure 3) revealed several characteristics. DO and DS parameters demonstrate a very strong positive correlation with a coefficient of 0.99, indicating that increases in dissolved oxygen are closely associated with increases in dissolved solids. A moderately strong positive correlation of 0.58 is observed between TP and SL, suggesting that higher concentrations of total phosphorus are generally accompanied by higher salinity levels. OR and ACQU\_TIME exhibit a positive correlation with a coefficient of 0.40, implying a potential relationship between the measurement duration and changes in redox potential. A negative correlation of -0.28 between SL and ACQU\_TIME is observed, indicating a decrease in salinity over time.



**Figure 3:** Correlation heatmap of key water quality parameters in aquaculture settings.

## CONCLUSION

In conclusion, this study highlights the efficacy of AI-based models in predicting levels of critical substances such as ammonia, nitrite, and nitrate in land-based aquaculture environments. Through extensive data collection from various sensors and subsequent analysis using advanced AI algorithms, including Random Forest and K-Neighbors Regressor, the research has demonstrated that these models can achieve high accuracy. The Random Forest model, in particular, showed exceptional performance metrics, with a mean absolute error (MAE) of 0.0150, mean squared error (MSE) of 0.0008, root mean square error (RMSE) of 0.0289,  $R^2$  of 0.9999, root mean square logarithmic error (RMSLE) of 0.0039, and mean absolute percentage error (MAPE) of 0.0024. These results indicate that AI-driven approaches can significantly enhance the monitoring and prediction of water quality parameters, crucial for maintaining fish health and reducing mortality rates in aquaculture settings.

## ACKNOWLEDGMENT

This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government (MSIT) (IITP-2024-2020-0-01791).

## REFERENCES

- Buentello, J. A., Gatlin III, D. M., & Neill, W. H. (2000). "Effects of water temperature and dissolved oxygen on daily feed consumption, feed utilization and growth of channel catfish (*Ictalurus punctatus*)", in: *Aquaculture*, Volume 182, Issues 3–4, pp. 339–352.
- Chumkiew, S., Jaroensutasinee, M., Tina, F. W., Jaroensutasinee, K., & Koad, P. (2019). "Physical factors affecting oyster diversity and distribution in Southern Thailand", in: *Journal of Environmental Biology*, Volume 40, No. 1, pp. 3–8.
- Chun, S. J., Cui, Y., Ahn, C. Y., & Oh, H. M. (2018). "Improving water quality using settleable microalga *Ettlia* sp. and the bacterial community in freshwater recirculating aquaculture system of *Danio rerio*", in: *Water Research*, Volume 135, pp. 112–121.
- Davidson, J., Good, C., Williams, C., & Summerfelt, S. T. (2017). "Evaluating the chronic effects of nitrate on the health and performance of post-smolt Atlantic salmon *Salmo salar* in freshwater recirculation aquaculture systems", in: *Aquacultural Engineering*, Volume 79, pp. 1–8.
- Kim, J. H., Kang, Y. J., Kim, K. I., Kim, S. K., & Kim, J. H. (2019). "Toxic effects of nitrogenous compounds (ammonia, nitrite, and nitrate) on acute toxicity and antioxidant responses of juvenile olive flounder, *Paralichthys olivaceus*", in: *Environmental Toxicology and Pharmacology*, Volume 67, pp. 73–78.
- Lindholm-Lehto, P. C., Pulkkinen, J. T., Kiuru, T., Koskela, J., & Vielma, J. (2021). "Efficient water treatment achieved in recirculating aquaculture system using woodchip denitrification and slow sand filtration", in: *Environmental Science and Pollution Research*, Volume 28, pp. 65333–65348.
- Lindholm-Lehto, P., Pulkkinen, J., Kiuru, T., Koskela, J., & Vielma, J. (2020). "Water quality in recirculating aquaculture system using woodchip denitrification and slow sand filtration", in: *Environmental Science and Pollution Research*, Volume 27, pp. 17314–17328.

- Mota, V. C., Martins, C. I. M., Eding, E. H., Canario, A. V. M., & Verreh, J. A. J. (2017). "Water cortisol and testosterone in Nile tilapia (*Oreochromis niloticus*) recirculating aquaculture systems", in: *Aquaculture*, Volume 468, pp. 255–261.
- Mwegoha, W. J. S., Kaseva, M. E., & Sabai, S. M. M. (2010). "Mathematical modeling of dissolved oxygen in fish ponds", in: *African Journal of Environmental Science and Technology*, Volume 4, No. 9, pp. 625–638.
- Rasheed Abdul Haq, K. P., & Harigovindan, V. P. (2022). "Water Quality Prediction for Smart Aquaculture Using Hybrid Deep Learning Models", in: *IEEE Access*, Volume 10, pp. 60078–60098.
- Salim, M. S. (2023). "Early Warning Signs and Dynamics of Hydrogen Sulfide Production in Recirculating Aquaculture System (RAS): With Nitrate and Iron (III) Addition", Master's thesis, Norwegian University of Science and Technology (NTNU).
- Suurnäkki, S., Pulkkinen, J. T., Lindholm-Lehto, P. C., Tiirola, M., & Aalto, S. L. (2020). "The effect of peracetic acid on microbial community, water quality, nitrification and rainbow trout (*Oncorhynchus mykiss*) performance in recirculating aquaculture systems", in: *Aquaculture*, Volume 516, 734534.
- Timmons, M. B. (2002). *Recirculating aquaculture systems*.
- Tziortzioti, C., Amaxilatis, D., Mavrommati, I., & Chatzigiannakis, I. (2019). "IoT sensors in sea water environment: Ahoy! Experiences from a short summer trial", in: *Electronic Notes in Theoretical Computer Science*, Volume 343, pp. 117–130.
- Wicks, B. J., Joensen, R., Tang, Q., & Randall, D. J. (2002). "Swimming and ammonia toxicity in salmonids: the effect of sub lethal ammonia exposure on the swimming performance of coho salmon and the acute toxicity of ammonia in swimming and resting rainbow trout", in: *Aquatic Toxicology*, Volume 59, Issues 1–2, pp. 55–69.
- Yu, J., Xiao, Y., Wang, Y., Xu, S., Zhou, L., & Li, X. (2021). "Chronic nitrate exposure cause alteration of blood physiological parameters, redox status and apoptosis of juvenile turbot (*Scophthalmus maximus*)", in: *Environmental Pollution*, Volume 283, 117103.
- Zhang, S. Y., Li, G., Wu, H. B., Liu, X. G., Yao, Y. H., Tao, L., & Liu, H. (2011). "An integrated recirculating aquaculture system (RAS) for land-based fish farming: The effects on water quality and fish production", in: *Aquacultural Engineering*, Volume 45, No. 3, pp. 93–102.

# Exploring the Use of ChatGPT4 API in Approaching Math Word Problems

M. G. R. Harsavardhan<sup>1</sup>, Joe Wong<sup>1</sup>, Kenneth Y. T. Lim<sup>2</sup>,  
and Swee Ling Leong<sup>3</sup>

<sup>1</sup>Anglo Chinese Junior College, Singapore

<sup>2</sup>National Institute of Education, Singapore

<sup>3</sup>Ministry of Education, Singapore

## ABSTRACT

With the evolving educational landscape precipitated by the COVID-19 pandemic, online education has become increasingly prevalent. Much help is needed to provide innovative solutions to address the challenges faced by both students and teachers during this time of crisis. This paper describes an independent research project conducted by a pair of high school students between April 2023 and February 2024, under the mentorship of a senior research scientist at the National Institute of Education in Singapore. The project investigates various methods of Tesseract OCR text recognition, OpenCV image processing, Flask web development and OpenAI's Large Language Models to improve mathematics-solving applications. Our program extracts text using Tesseract OCR, utilising it as input for the GPT-4 API, enabling a conversational presentation of mathematics problems. Users interact by inputting the image address of the math problem that they would like the AI to solve, and GPT-4 provides solutions with detailed step-by-step explanations. OpenCV improves the provided image's quality such as making the text or diagrams more distinct to reduce the possibility of them being misinterpreted. Through rigorous evaluation by testing with different types of maths problems of varying difficulty, our findings underscore the potential for advanced language models in educational tools, offering interactive and intuitive maths problem-solving experiences. There were a few limitations encountered during experimentation, such as challenges with the extraction of non-Latin alphabets and accuracy of the OpenAI's Large Language Modules when solving more complex diagram problems, highlighting the need for further refinement to enhance the system's robustness and adaptability. Future work involves addressing these limitations to broaden the system's applicability for educational purposes and beyond.

**Keywords:** Education, Mathematics-solving, Tesseract OCR, OpenCV, GPT4-API

## INTRODUCTION

### Background and Purpose of the Research Area

In the dynamic tech landscape, Python shines as a versatile language for various projects, from simple web apps to complex machine learning with deep-learning algorithms. Our project leverages Python's potential, merging Tesseract OCR, OpenCV, and the OpenAI API for innovative math problem-solving. Designed to assist students, our approach provides quick, step-by-step answers to both word and diagram problems, showcasing the

power of Python in creating comprehensive solutions for diverse educational challenges.

### Hypothesis of the Research

Our hypothesis centres on the idea that the integration of optical character recognition (Tesseract OCR) and image processing (OpenCV) with advanced natural language processing (OpenAI API, specifically GPT-3.5 and GPT-4 Vision) can significantly enhance the accuracy and efficiency of maths problem-solving. This amalgamation of technologies will provide solutions and detailed step-by-step explanations, revolutionising how users interact with mathematical content.

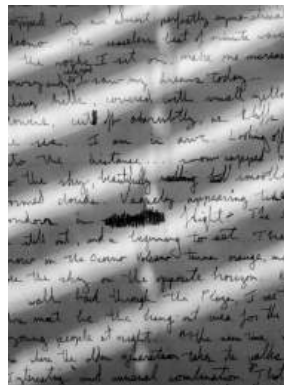
### METHODOLOGY/MATERIALS

#### Tesseract OCR

Employing Tesseract OCR, we extracted text, focusing on mathematical expressions and laying the foundation for subsequent processing. Tesseract OCR employs algorithms for line and word detection, character identification, and accommodates variations in spacing due to different fonts (Smith, 2007). The versatility of Tesseract OCR, with algorithms like fixed pitch detection and proportional word finding, is crucial for reading user handwriting with varying fonts (NYU, 2024). These features enhance the accuracy of our program, making it resilient to diverse font styles and attributes.

#### OpenCV Image Processing

Utilising OpenCV, we captured images from the webcam and sent them to the Tesseract OCR to extract text. However, many challenges existed, such as shadows and varying lighting conditions (Computer vision engineer, 2023), as shown in Figure 1. This affects the Tesseract OCR significantly and prevents it from extracting data accurately (Klippa, 2022).

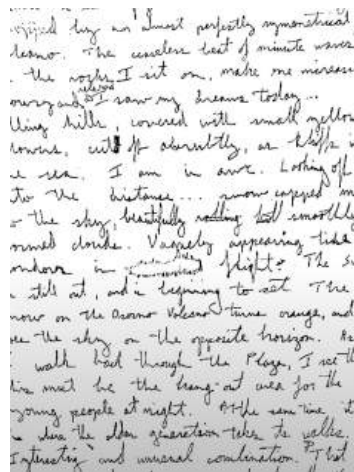


**Figure 1:** Original image with bad lighting.

To solve this problem, an adaptive threshold is used instead. The algorithm determines the threshold required based on the surrounding pixels (Kumar, 2019). There will be different thresholds calculated for different regions of the image, allowing the text which is darker to become more visible regardless if there is a shadow on the image (Rosebrock, 2021).

```
gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
thresh =
cv2.adaptiveThreshold(gray_image, 255, cv2.ADAPTIVE_THRESH_GAUSSIAN_C,
cv2.THRESH_BINARY, 21, 30)
```

This code snippet changes the image to grayscale, then uses adaptive thresholding to make the darker areas more visible, as shown in Figure 2.



**Figure 2:** Image after adaptive thresholding.

Through OpenCV's webcam and adaptive thresholding feature, the image can become clearer and users would not need to worry about poor lighting or blocking text when scanning their images.

### OpenAI GPT-3.5 API

We integrated OpenAI's GPT-3.5 API to process prompts derived from extracted data. Users could pose mathematical queries, and GPT-3.5 provided detailed, conversational solutions. Our utilisation involved a straightforward API call to GPT-3.5, enabling prompt processing and solution generation within the program. GPT-3.5 is a learning language model that uses transformer neural networks. Neural networks have multiple layers of algorithms where they would recognise the underlying relationships in the input, and the transformer model takes the context into more detail, improving performance on NLP tasks. It analyses natural language queries to predict the best possible response based on their understanding of language. The transformer neural network architecture also uses self-attention mechanisms to focus on different parts of the input text during each processing step. These GPT models rely on the knowledge they gain after they are trained with hundreds of billions of parameters on massive language datasets (AWS, n.d.).



## OpenAI GPT-4 Vision API

In addition to GPT-3.5, we explored the capabilities of GPT-4 Vision for image analysis.

```
response = client.chat.completions.create(

    model="gpt-4-vision-preview",
    messages=[
        {
            "role": "user",
            "content": [
                {"type": "text", "text": "What's in this image?"},
                {
                    "type": "image_url",
                    "image_url": {
                        "url":
"https://mindyourdecisions.com/blog/wp-content/uploads/2016/08/harde
st-easy-geometry-thumb-social.png",
                    },
                },
            ],
        }
    ],
    max_tokens=300,
)

print(response.choices[0].message.content)
```

This code snippet shows a GPT-4 Vision API call.

This multimodal model integrated image inputs into text, enabling a more comprehensive understanding of the mathematical content presented in images.

## Integration of Flask for Interactive Problem Solving

Flask is a simple Python Web framework that has useful tools and features that make website development easier. It provides developers with flexibility and is easily accessible as a website can be easily set up with only one Python file. Flask also offers various customisations with different elements like text boxes and buttons and a wide selection of colours. We leveraged Flask to design an intuitive web interface where users can effortlessly submit math problems for our custom solver. The website features a user-friendly input box where individuals can paste the URL of an image containing a maths problem. This seamless interaction is made possible through Flask's capabilities in handling HTTP requests and rendering dynamic HTML content. Upon submission of the image URL, Flask orchestrates the communication between the user interface and our Python maths problem-solving code. Using Flask routes, we directed form submissions to the appropriate function, initiating the process of solving the maths problem extracted from the provided image.

## Dynamic Rendering With Jinja2

Flask's integration with the Jinja2 template engine allowed us to dynamically render HTML templates. The result of the math problem-solving process

is seamlessly displayed on the website, providing users with immediate feedback. This dynamic rendering capability enhances the user experience and ensures a fluid interaction with the system.

## Scalability and Future Development

The modular structure of Flask ensures scalability and facilitates future enhancements to our system. As our project evolves, Flask provides a solid foundation for incorporating additional features, refining the user interface, and optimizing the overall user experience (Saini, 2023).

In essence, Flask serves as the backbone of our web-based math problem-solving platform, seamlessly connecting users with the power of our custom Python code and OpenAI API integration.

```
from flask import Flask, render_template, request

app = Flask(__name__)

@app.route('/')
def index():
    return render_template('index.html')

@app.route('/', methods=['POST'])
def solve_problem():
    url = request.form['url']
    # Call your Python math problem solver function with the provided URL
    # Solve the math problem and get the result
    result = solve_math_problem(url)
    return render_template('index.html', result=result)

if __name__ == '__main__':
    app.run(debug=True)
```

This code snippet shows how Flask is used to set up the website and to call functions when a button is clicked. Figure 3 shows the user interface of the website, in which the user can paste an image URL and submit it in order to get a response.

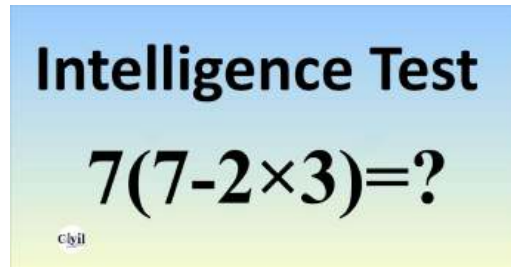


**Figure 3:** The user interface of the website.

## RESULTS/DISCUSSION

We have developed a model using Tesseract OCR to extract text from images, OpenCV to modify these images for the computer to read, and GPT-3.5 to generate answers to the questions from the text extracted. Hence, we investigated the difference between our model and the GPT-4 Vision model.

### Test 1 (GPT-4) | Extract Equations From Image | Fail



**Figure 4:** The image (Engineering discoveries, 2020) we tested our model with.

```

Enter the path to the image containing the math problem: C:\Users\Harsavardhan\Downloads\Can-You-Solve-This-Math-Problem-scanned.jpg
Solution:
It seems like there is some misunderstanding with the problem provided. The equation  $7(7 - 2x3)$  can be simplified, like so:
Simply expand the formula  $7(7 - 2x3)$ 
=  $7(7 - 6x)$ 

However, "Cly" seems to be out of the context due to the structure of the problem you've provided. It would be helpful if you could provide more information about the "Intelligence Test" or, potentially, confirm if "Cly" is to be considered part of the equation.

If "Cly" is a typo and should instead say "Simplify," then the simplification process would look like this:
1. Simplify inside the parentheses first, according to the standard order of operations (PEMDAS/BODMAS - parentheses, exponents, multiplication and division, addition and subtraction). Here you have 2 times 3, which equals 6.
2. So your equation simplifies to  $7(7 - 6x)$ .
3. This equation can be distributed further to  $7^2 - 67x$ .

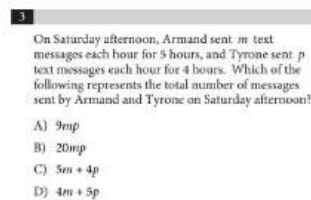
If "Cly" is part of the equation, we'll need more information to proceed.
PS C:\Users\Harsavardhan>

```

**Figure 5:** The result from our model.

*Comments:* Heavily inaccurate. Misinterpretation of the equation due to font type. A potential solution involves training a custom machine learning algorithm with various fonts, which was not feasible within the project's time constraints.

### Test 2 (Vision) | Text Problem | Success



**Figure 6:** The image (Wells, 2019) we tested the vision model with.

The image displays a math word problem followed by multiple-choice answers. The problem reads:

On Saturday afternoon, Armand sent  $m$  text messages each hour for 5 hours, and Tyrone sent  $p$  text messages each hour for 4 hours. Which of the following represents the total number of messages sent by Armand and Tyrone on Saturday afternoon?

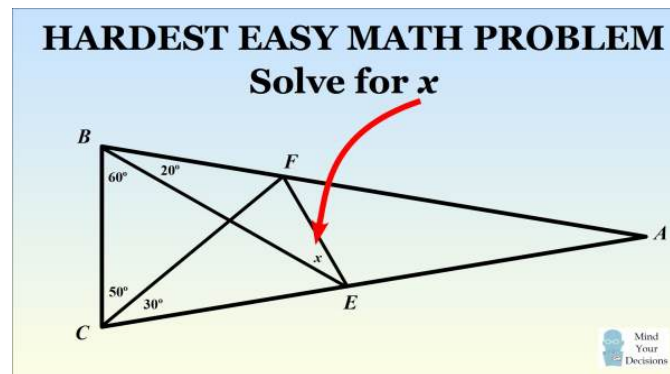
A)  $9mp$   
 B)  $20mp$   
 C)  $5m + 4p$   
 D)  $4m + 5p$

The correct choice that represents the total number of messages sent by Armand and Tyrone is C)  $5m + 4p$ , since Armand sent messages for 5 hours and Tyrone for 4 hours.

**Figure 7:** The result from the model.

*Comments:* GPT-4 Vision demonstrated superior capabilities in solving various problems compared to Tesseract OCR. Notable successes were observed in both text and diagram-based problems.

### Test 3 (Vision) | Diagram Problem | Success



**Figure 8:** The image (Talwalkar, 2016) we tested the vision model with.

The image displays a geometric math problem consisting of a large triangle labeled A, B, C, with smaller triangles inside it. We're asked to solve for the angle  $x$ , which is at point F where lines inside the triangle intersect. There are various other angles given:

- Angle at point B is 60 degrees.
- There's another angle next to it marked as 20 degrees inside the triangle.
- The angle at point C is 50 degrees, with another angle next to it inside the triangle marked as 30 degrees.

There's a large "HARDEST EASY MATH PROBLEM" caption at the top, and "Solve for  $x$ " in the middle with a red arrow pointing to angle  $x$ . It's a puzzle designed to test knowledge of geometry, specifically the properties of angles within triangles and intersecting lines.

To solve for  $x$ , one would typically use the fact that the sum of angles in a triangle is 180 degrees, along with the principles of supplementary and vertically opposite angles. By systematically applying these principles and calculating unknown angles step by step, it's possible to determine the value of  $x$ . However, without the ability to perform calculations or mark on the image, I'm unable to provide the exact solution here.

**Figure 9:** The result from the model.

*Comments:* GPT-4 Vision was able to identify angles but unable to name angles and solve for  $x$  despite knowing the steps to the solution.

## CONCLUSION

### Future Plans

Unfortunately, due to the limited time constraints, we were unable to implement everything we had hoped to previously. Firstly, we would like to utilise Dall-E, an image generation module developed by OpenAI, to help draw diagrams or models for our users to better understand the working presented by GPT-4. Secondly, we would like to format the large chunk of text provided by GPT-4 into steps that users can easily read and follow. Thirdly, we hope that our app could have been more interactive by asking the user for the next step to follow or to fill in the working after telling the user the next step to solve the question. Lastly, our web application requires some refining in its User interface to be more readable and user-friendly.

### Conclusion

In conclusion, our project successfully integrated Tesseract OCR, OpenCV, and the OpenAI API, showcasing the synergy of diverse tools for comprehensive math problem-solving. Tesseract OCR excelled in text extraction, while OpenCV addressed image quality issues, and GPT-3.5 and GPT-4 Vision provided insightful solutions for both text- and diagram-based problems. Our findings underscore the potential of these technologies in creating interactive learning experiences.

However, our experimentation revealed certain limitations. GPT-4 Vision faces challenges with non-Latin alphabets, impacting global optimization. Enlarging small text improves readability but risks cropping essential details. Rotation of text or images may lead to misinterpretation, and struggles with diverse visual elements and spatial reasoning tasks were observed. Future work involves refining our approach to enhance robustness and adaptability, aiming to broaden the system's applicability for educational purposes and beyond.

## REFERENCES

- AWS (n.d.). *What is GPT AI? - Generative Pre-Trained Transformers Explained*. Available at: <https://aws.amazon.com/what-is/gpt/#:~:text=More%20specifically%2C%20the%20GPT%20models,on%20their%20understanding%20of%20language> (Accessed: 4 January 2024).
- Computer vision engineer (2023). *OpenCV tutorial for beginners | FULL COURSE in 3 hours with Python*. Available at: <https://www.youtube.com/watch?v=eDIj5LuIL4A> (Accessed: 4 January 2024).
- Engineering Discoveries (2020). 'A trip back to high school: Can you solve this math problem?', *Engineering Discoveries*, 13 October. Available at: <https://engineeringdiscoveries.com/a-trip-back-to-high-school-can-you-solve-this-math-problem/> (Accessed: 4 January 2024).
- Klippa (2022). 'Tesseract OCR: What is it, and Why Would You Choose it in 2023?', *Klippa*, 20 October. Available at: <https://www.klippa.com/en/blog/information/tesseract-ocr/> (Accessed: 4 January 2024).
- Kumar, S. (2019). 'A straightforward introduction to Image Thresholding using python', *Spinor*, 2 October. Available at: <https://medium.com/spinor/a-straightforward-introduction-to-image-thresholding-using-python-f1c085f02d5e> (Accessed: 4 January 2024).

- NYU (2024). *Tesseract OCR Software Tutorial*. Available at: <https://guides.nyu.edu/tesseract> (Accessed: 4 January 2024).
- Rosebrock, A. (2021). 'Adaptive Thresholding with OpenCV (cv2.adaptiveThreshold)', *PyImageSearch*, 12 May. Available at: <https://pyimagesearch.com/2021/05/12/adaptive-thresholding-with-opencv-cv2-adaptivethreshold/> (Accessed: 4 January 2024).
- Saini, A. (2023). 'An Easy introduction to Flask Framework for Beginners', *Analytics Vidhya*, 23 August. Available at: <https://www.analyticsvidhya.com/blog/2021/10/flask-python/> (Accessed: 4 January 2024).
- Smith, R. (2007). 'An Overview of the Tesseract OCR Engine', *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. Curitiba, Brazil, 23–26 September. Los Alamitos: IEEE Computer Society, pp. 629–633. doi: 10.1109/ICDAR.2007.4376991.
- Talwalkar, P. (2016). 'The Hardest Easy Geometry Problem – Sunday Puzzle – Mind Your Decisions', *Mind Your Decisions*, 4 September. Available at: <https://mindyourdecisions.com/blog/2016/09/04/the-hardest-easy-geometry-problem-sunday-puzzle/> (Accessed: 4 January 2024).
- Wells, J. (2019). 'The power of plugging numbers in', *Wells Academic Solutions*, 20 September. Available at: <https://www.wellsacademics.com/the-power-of-plugging-numbers-in/> (Accessed: 4 January 2024).

# An Exploration of Machine Learning and Reinforcement Learning for Emotional Well-Being

Kenneth Y. T. Lim<sup>1</sup>, Ryan Thaddeus Shi Eng Chio<sup>2</sup>,  
and Zhi Wen Lim<sup>3</sup>

<sup>1</sup>National Institute of Education, Singapore

<sup>2</sup>National Junior College, Singapore

<sup>3</sup>Independent Scholar, Singapore

## ABSTRACT

With the high levels of stress in Singapore, mental and emotional well-being is an important health and social issue today. Research has shown the positive effects of pet ownership on mental and emotional well-being, however challenges of owning a pet in Singapore such as pet licensing restrictions, high costs, fear of losing a pet, a busy lifestyle and even allergies may deter pet lovers from owning a pet. Thus, we propose a technology-driven solution to emulate the useful effects of pets while mitigating the challenges of pet ownership. This project focuses on designing emotion recognition and reinforcement learning models as a stepping stone to individualise responses to a person's emotions. Our approach utilises the output from the emotion recognition model as an input in the proposed reinforcement learning algorithm. Hence, the paper first compares pre-trained and custom trained facial recognition models, and postulates the use of physiological signals via hardware sensors to further enhance the emotion recognition model. This is inspired from the ability of pets to perceive and respond to different emotions based on facial expressions and physiological signals like heart rate. The paper then outlines the development of novel K-Bandit algorithms in reinforcement learning tested on simulated reward functions, with the aim of optimising parameters for individualised responses to a person's emotions. Since reinforcement learning is typically used in simulation scenarios, this paper works towards developing a model that will eventually learn a person's preferences in real time by monitoring their emotional changes. To conclude, this project has showcased the feasibility of facial expressions and physiological signals for emotion recognition, and established the effectiveness of our proposed parameter optimisation functions in the K armed bandit reinforcement learning model to customise responses based on an individual's emotions. We hope this paper can act as a basis for future works in creating a human-friendly prototype to emulate man's best friend.

**Keywords:** Human systems integration, Machine learning, Reinforcement learning, Empathetic robotics, Human robot interaction

## INTRODUCTION

Studies have shown that Singaporeans face high levels of stress (Cigna Singapore, 2022), so measures to improve mental and emotional

well-being are becoming increasingly important. Singapore's recent launch of a national mental health strategy highlights this issue's significance (Teo, 2023). A study conducted by Boehringer Ingelheim Animal Health Singapore found that pets had a positive impact on the mental well-being of the majority of participants (Goh, 2021). Moreover, research has found that the presence of pets contributed to a lower resting heart rate as well as a faster recovery to resting heart rate when feeling stressed (Allen et al., 2002). However, Singaporeans may be deterred from pet ownership due to pet licensing restrictions for Housing & Development Board (HDB) flats (Housing & Development Board, n.d.), the high cost of owning a pet, estimated to be around \$102,000 over a pet's lifespan (Wong, 2022), and the rising cost of veterinary fees and services by at least 10 - 20% due to inflation (Ee, 2023). Moreover, allergy considerations could discourage pet ownership, with allergic rhinitis typically triggered by pet dander affecting up to 13.1% of the Singapore population (Pharmaceutical Society of Singapore, 2023).

## **AIMS & OBJECTIVES**

We propose a solution to incorporate artificial intelligence and robotics to emulate the useful effects of pets while mitigating the challenges of pet ownership, focusing on designing emotion recognition and reinforcement learning models as a stepping stone to individualise responses to a person's emotions. This paper aims to (1) contribute to existing literature regarding emotion recognition and (2) explore the use of reinforcement learning in individualising responses based on emotions. The paper will first delve into the emotion recognition details before investigating the use of the K Armed Bandit reinforcement learning model in individualising actions based on emotions detected. Given time and resource constraints, our research is focused on the theoretical aspects of the larger idea. Since the hardware can be developed from pre-existing frameworks, we scoped our research to the computational approach of emotion recognition and reinforcement learning taken for simulating pet ownership, which should be taken as the primary focus in this research endeavour.

## **METHODOLOGY**

In the development of an emotional recognition model, no active participation was required from those asked to contribute to the dataset. Hence, there were no practices constituting "direct human participation" or involving their associated safety concerns in this stage of the investigation. Friends and family who consented, submitted images of their interpretations of the seven pre-identified basic emotions (anger, fear, happiness, sadness, surprise, neutral and disgust). The use of custom training data was hypothesised to allow for a wider range of possible expressions to be classified as the same emotion for more accurate emotion recognition in practical implementations.

The reinforcement learning aspect of research aimed to create an algorithm that allowed our robot prototypes to effectively and efficiently adapt to



feedback provided on the effectiveness of its selected emotional response in real-time constraints. This focus was selected due to a gap of knowledge identified in current RL algorithms, whose training times were not emulative of actual learning rates. (The design of many RL algorithms and hence what was considered “efficient” in current literature was done for systems where training and interaction processes were done separately, as opposed to concurrently in our proposed design application.) We first create a testing ground for our algorithm by parameterising human-pet companion interactions. A basic three stage framework for an interaction was proposed (E.g. Approach > Interaction > Retreat) with each stage being composed of various mutable parameters (E.g. Approach considerations may be the speed of the approach, or the distance from the human a companion stops at) Background research regarding how interaction parameters generally evolved with respect to varying conditions allowed constraints to be made for the Reinforcement Learning Reward equation (the function to be maximised during the training process) allowing for more optimisations in our algorithms to be proposed. Each algorithm would then be tested with the metrics of accuracy and efficiency. Environmental conditions (noise and error in the input information) was also simulated to assess the stability of our proposed algorithm in applied conditions.

## **MACHINE LEARNING IN EMOTION RECOGNITION**

A study conducted by the University of Veterinary Medicine Vienna indicates that dogs can distinguish and respond to 6 basic facial expressions: anger, fear, happiness, sadness, surprise, and disgust (Siniscalchi, d’Ingeo and Quaranta, 2018). Hence, we chose to explore facial expressions as a form of emotion recognition. Psychologist Dr Paul Ekman (1992) identified the same 6 emotions as universal emotions. Another study corroborated Ekman’s claim of universal facial expressions, finding that people from different cultures share approximately 70% of facial expressions (Cowen et al., 2021). However, Jack et al. (2012) shed light on the differences in facial expressions across cultures, such as Chinese participants conveying emotion through the eyes or Western Caucasian participants expressing emotions with the eyebrows and mouth.

The FER2013 facial expression dataset comprises mainly of Caucasian and African Americans (Lukac et al., 2023). A custom dataset for Singapore was created through mass data collection to test the accuracy of a pre-trained model DeepFace against a custom trained Teachable Machine model (Teachable Machine, n.d.). The Teachable Machine model was neural network based, trained using the Tensorflow-Keras library in python. A 50/50 split was used for training and testing data. The custom trained model had 75 images for the training dataset of each emotion to ensure fair distribution amongst the different classes of emotion.

## **PHYSIOLOGICAL SIGNALS IN EMOTION RECOGNITION**

Pets such as dogs are able to detect the rise or fall of heart rate (Deangelis, 2020), and can be trained to monitor blood oxygen saturation (SPO<sub>2</sub>)

levels (Leomiti and Ausman, 2023). Hence, we chose to explore the use of physiological signals in addition to facial expression to increase the accuracy of emotion recognition. A study on the reliability of physiological responses induced by basic emotions concluded that physiological features, such as skin conductance level (SCL), heart rate (HR) and blood volume pulse (BVP), would be important for emotion recognition in the human computer interface area (Jang et al., 2019). Other studies have found that blood oxygen levels (SPO2) were also a possible physiological feature to classify emotions (Alkawaz et al., 2015). Stress, typically associated with fear and anger (Gu et al., 2019), as well as sadness which may lead to crying (Raypole, 2020), can lower blood oxygen level (Cox, 2022). Emotions can lead to differences in blood oxygen levels (Alkawaz et al., 2015). When interfacing the MAX30102 heart rate sensor on the Arduino, output data of SPO2 levels and heart rate was obtained. This shows the feasibility of including physiological signals to enhance emotion recognition.

From the accuracy results of both pre-trained and custom trained emotion recognition models, it is shown that the custom trained Teachable Machine model has a higher average accuracy than that of the pre-trained DeepFace model (as shown in Table 1).

**Table 1.** Accuracy of emotion recognition models.

Emotion Recognition	Deep Face	Teachable Machine
Happy	45.3%	78.7%
Sad	28.0%	73.3%
Surprise	42.7%	57.3%
Angry	28.0%	53.3%
Fear	57.3%	74.7%
Disgust	8.0%	12.0%
Neutral	77.3%	48.0%
Average	40.9%	56.8%

It should be noted that disgust is the emotion with the lowest accuracy score. An article exploring the use of a convolutional neural network trained on the same FER2013 dataset also had poor accuracy in classifying disgust, citing the smaller number of disgust facial expressions in the training dataset as the limitation (Tang, 2019). However, the custom trained Teachable Machine model had the same number of images for all classes of emotion, and still had a significantly low accuracy score for disgust. Thus, further research could be undertaken in this area to investigate the difficulty in classifying this emotion.

## REINFORCEMENT LEARNING ALGORITHMS TO SIMULATE ORGANIC LEARNING BEHAVIOUR

Reinforcement learning algorithms are fundamentally maximisation problems: they evolve policy algorithms to select actions that optimise reward functions derived from intended outcome behaviours. We aim to

make this process more efficient for practical simulation of organic learning behaviour. With the parameter framework in our proposed robot design, our proposed use of reinforcement learning intends to tune the action parameters based on a Reward function that reflects positive changes in sensor data when a specific action configuration is executed.

For the proposed inputs into the standard reinforcement learning model, the State is determined through the facial detection model described above. The Reward is determined by the reward function  $R(x)$  which encompasses the “positive changes” detected in sensors. Determining the actual reward functions and sensor response correlations are not discussed in this paper and are proposed as future work to be completed when progressing towards eventual testing and deployment.

We discuss assumptions and constraints that can be made to the reward function  $R(x)$  where optimisations to the action selection algorithm may be proposed. It is asserted that there is only one optimal parameter value at any given state of companion owner interaction. While the optimal parameter itself may shift with a greater number of encounters, multiple configurations of a robot’s parameters cannot be equally preferred. It is assumed that for reward functions where parameters may appear to have two or more similarly preferred values, that only one of such is eventually optimised. In a reward function that has only one stationary maxima, and when provided with the range where this point may lie, the optimum point can be estimated through analysing the function value at regular intervals  $k$  within the specified range. The subsequent removal of values that return lower values of the function, and addition of values between those who return higher values of the function, would shift the range in a manner such that the exact maxima point would eventually be found. This principle of successive halving is often applied in the hyper-parameter optimization of neural network configurations. This algorithm can reliably determine the optimum point in single maxima functions within the time constraints suitable for authentic pet-human interactions. Research on human-robot interaction has emphasised the need for individualised technology development (Søraa et al., 2022), with a study investigating living with robots noting that participants became increasingly comfortable with the robot approaching them closely over time (Mehta and Losey, 2023). The assumption made is that there is a single optimising value for each mutable parameter in the robots action array, resulting in a reward function with a single maximum. This allows for parameter optimization through a sequential decision-making algorithm.

## THE K ARMED BANDIT MODEL

The Stochastic K-Armed Bandit is a Single State Markov Decision Process that frames *sequential decision-making under uncertainty problems*. It abstracts the action selection as: within a  $K$  number of slot machines that return Reward  $R$  drawn randomly from a fixed probability distribution, and maximises the cumulative rewards for  $N$  slot machines and an array of actions  $\{a_i = 1, a_i = 2, \dots a_i = N\}$ , where the action selected at a time step

$t$  is denoted as  $A_t$ . This results in reward  $R_t$ , the mean reward of the action number  $i$ ,  $v_i$  is denoted as:<sup>28</sup>

$$v_i = v(a_i) = E[R_k | A_k = a_i] \quad (1)$$

$v_i$  is initially unknown and is approximated with greater time steps  $t$  as:

$$\hat{v}_{i,t} = \hat{v}_t(a_i) = \frac{\sum_{j=1}^t R_{i,j}}{n_{i,t}} \quad (2)$$

where  $R_{i,j}$  is the reward from action  $a_i$  at step  $t$  and  $n_{i,t}$  is the selection count of the action  $a_i$  prior to  $t$ . Rearranging (2), a recurrence relation is obtained, to update the estimated value  $\hat{v}_t$  with each step of the algorithm:

$$\hat{v}_k = \frac{1}{t + 1} (R_{t-1} + t \hat{v}_{k-1}) \quad (3)$$

The selection of an action uses the epsilon greedy approach: for specified number  $\epsilon$  and randomly generated number  $p$ , if  $p \leq \epsilon$ , an action is selected by taking the largest reward  $A_t = \text{argmax}_{a_i} \{v_{i,t}, v_{i+1,t}, \dots\}$ , whereas otherwise randomly selected from the action array. In this way, a balance is struck between the maximisation of historically rewarded actions, and the exploration of new potential strategies. Such an approach can be applied to the algorithm described in 3.1 to approximate Reward function maximas, with the range divided into  $k$  intervals being taken as the action array and with the Reward value  $R$  taking up the value of the reward function  $R(x)$  instead of a static fixed value. The subsequent section describes this implementation. For the Reward Function  $R(x)$  with no local maximas and given an initial search range  $[a, a_n]$ , this range is divided into a  $k$  array of ‘‘arms’’  $A = \{a_1, a_2, \dots, a_n\}$ , then fed into a  $k$  number of slot machines, which returns their corresponding input from the reward function e.g.  $\{R(a_1), R(a_2), \dots, R(a_n)\}$ . The K bandit model in 3.2 selects from the array of slot machines and ‘‘plays’’ until the proportion of the most historically selected actions from all choices exceeds a specified confidence value  $C$ . Thereafter, the average rewards for each ‘‘arm’’ would be sorted, discarding ‘‘arms’’ from a lower performing half and pruning the array down to a size  $\lfloor \frac{k}{2} \rfloor$ . The array is then resorted, with new arms being inserted between neighbours as their mean, obtaining a new set of arms  $A$  of original length  $k$ . The K bandit model then plays and prunes recursively, until the range of  $A$  falls below a specified convergence value,  $\mu$ . As such, an accurate range of values for where the Reward maxima lies is obtained. The *successive halving* algorithm was tested: Provided with an initial range from 1 to 60, the hidden maxima value of 31.5 was reached by the algorithm after 6 interactions (translated as 6 different companion, pet encounters), demonstrating the capability of *successive halving* in tuning parameters within a relatively few number of iterations.

## ALGORITHM 2 GRADIENT ASCENT

While *Successive Halving* is an efficient and effective algorithm for determining maximas of reward functions that remain unchanged throughout the duration of interaction, it is unable to adjust for changes should the optimal value of the reward function shift during the course of the interaction. This presents an important drawback to be considered, as the optimum parameter value seldom remains unchanged with respect to the time step of the model. For example, it can be reasonably asserted that human companion interactions become more trusting with an increased number of encounters. In such a case, the optimum value for parameters such as the distance of approach would vary with the extent of trust. As such we draw from various gradient descent techniques commonly applied in machine learning to allow for our model to dynamically shift along with the Reward equation when a change is detected, thus taking into account that human preferences may change over time. For the Reward Function  $R(x)$  including an additional parameter  $t$ ,  $R(x, t)$  whereby for all parameters  $t$  in  $R(x, t)$  the corresponding  $R(x)$  contains no local maximas, and provided an initial array of parameters  $P$ , gradient ascent can be used to recursively shift the parameter selection range towards the new function maxima:  $R_{x+1,t} = R_{x,t} + \alpha \frac{\partial R}{\partial x}$  (where  $\alpha$  is the “learning rate”). The value of  $\frac{\partial R}{\partial x}$  is approximated by taking the gradient between the Reward function at the maximum and minimum values in the parameter array  $P$ :  $\frac{\partial R}{\partial x} \approx \frac{R(\text{argmax}(P), t) - R(\text{argmin}(P), t)}{\text{argmax}(P) - \text{argmin}(P)}$ . The model would default to responding either as the current maximum or minimum in the parameter array, depending on the remaining value required to approximate the gradient.  $\alpha$  is taken as the  $\text{range}(P)$ . The bounds of  $P$  are then shifted by the value  $\alpha \frac{\partial R}{\partial x}$ , with an additional step added in the step direction (positive or negative and taken as the opposite sign to the initial gradient:  $\frac{\partial R}{\partial x_1}$ ). As such values of  $P$  progresses towards the maxima, resizing in the direction of search to favour values closer to convergence at later recursions. When  $\frac{\partial R}{\partial x}$  falls below a “find condition value”  $\lambda$  ( $\frac{\partial R}{\partial x} \rightarrow 0$  approaching the maxima), *successive halving* with input  $P$  is then used to determine the final new maxima range.

Models utilising sensor data as input are subject to significant amounts of noise and distortion, owing to minute uncontrolled variations in the environment of choice, and inherent uncertainties within our chosen instruments, as such it is imperative that the proposed algorithms maintain a high level of accuracy outside of the assumption of a smoothed reward function. We select the reward equation for successive halving arbitrarily as a distribution function with a maxima of 0.40 at parameter 51.5. Noise was simulated by randomly scattering inputs about the range of the true function value, whereby distortion could be varied through changing the range variable  $r$ . The convergence and error scatter plots for 5 trials,  $r = [0.01, 0.02, \dots, 0.05]$  are below. The range of  $R(x)$  in our test selection was  $[0.00, 0.40]$  with a peak distortion  $r = 0.05$  translating to a 12.5% error overall. Despite this, successive Halving maintained a high accuracy range between 99.2 - 99.8% when varying the distortion values, with all configurations converging after 4–7 interactions. This demonstrates the resilience in the

algorithm to a reasonable level of distortion within input data. Gradient Ascent was similarly tested where a quadratic reward function  $R(x)$  was switched after the 10th stable interaction from a maxima of 31.5 to 11.5, with a similar 99.0% - 99.8% accuracy range, converging 10–14 interactions after the change in reward function, demonstrating also a resilience to a reasonable level of distortion within input data.

Variations in the reward function's shape were initially hypothesised to have significant effects on the number of interactions required for models to converge towards maximas: this is significant given the amount of variability that reward functions may take between individuals in our selected use case, and may present a significant design flaw to be imperatively addressed. We similarly arbitrarily select a distribution function with a maxima of 0.40 at parameter 51.5, but include additional parameters to allow variations in the equation shape, in the dimensions of translation and scaling. Notably scaling parallel to  $x$  and  $y$  axis had no effect on the number of interactions required to converge to stability, when translation between the range of the function was performed, an interesting pattern emerged, indicating that the number of interactions required for convergence increases as the maxima approaches the bounds of the specified function range but also contains periodic regions at certain translation values where an increase in the number of interaction is noted.

## **EVALUATION AND CONCLUSION**

The single maxima constraints may result in reward functions being unsupported by the gradient ascent algorithm, as there is currently a possibility of approaching a local maxima. Additionally, both algorithms are susceptible to butterfly effect conditions, particularly the convergence value in Algorithm 1 and “find condition” in Algorithm 2 as values greater than a specified value would often result in highly inaccurate final parameters. While a solution was found in setting both variables to be extremely small, there was a significant tradeoff in the number of interactions required. Tests however demonstrate that successive halving is capable of converging parameter values towards an unknown optimal value within a reasonable number of iterations. Additionally, further optimisations to the gradient ascent algorithm are proposed to minimise its susceptibility and improve its performance. Further research can be conducted to address limitations in this project. First, further experimentation to determine the extent of effectiveness of incorporating physiological signals in emotion recognition. Second, more would need to be done for the development of reward equations with sensor data as input in real world conditions. Finally, further research of a physical prototype for mass testing on human participants.

To conclude, this project has achieved its goals of showcasing the feasibility of facial expressions and physiological signals for emotion recognition. This project has also established the effectiveness of our proposed parameter optimisation functions in the  $K$  armed bandit reinforcement learning model to customise responses based on an individual's emotions. We hope that this project has provided a new perspective as to how artificial intelligence and

robotics can be used to emulate the positive effects of pets, and acts as a basis for future works in creating a human-friendly prototype to emulate man's best friend.

## REFERENCES

- Alkawaz, M. H., Mohamad, D., Basori, A. H., Saba, T. and Rehman, A. (2015). 'The Correlation between Blood Oxygenation Effects and Human Emotion towards Facial Skin Colour of Virtual Human', *3DR Express*, 6(2), article no: 13. doi: 10.1007/s13319-015-0044-9.
- Allen, K., Blascovich, J. and Mendes, WB. (2002). 'Cardiovascular reactivity and the presence of pets, friends, and spouses: The truth about cats and dogs', *Psychosomatic Medicine*, 64(5), pp. 727–39. doi: 10.1097/01.psy.0000024236.11538.41.
- Cigna Singapore (2022). *Stressed in Singapore*. Available at: [https://www.cigna.com.sg/static/docs/pdf/Cigna\\_Singapore\\_360\\_Well-being\\_Report\\_2022.pdf](https://www.cigna.com.sg/static/docs/pdf/Cigna_Singapore_360_Well-being_Report_2022.pdf) (Accessed: 5 January 2024).
- Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H. and Prasad, G. (2021). 'Sixteen facial expressions occur in similar contexts worldwide', *Nature*, 589(7841), pp. 251–257. doi: 10.1038/s41586-020-3037-7.
- Cox, J. (2022). 'Can Stress Cause Low Oxygen Levels?', *PsychCentral*, 12 August. Available at: <https://psychcentral.com/stress/can-stress-cause-low-oxygen-levels> (Accessed 5 January 2024).
- Deangelis, S. (2020) 'What Can Dogs Detect?', *Wag!*, 6 April. Available at: <https://wagwalking.com/sense/what-can-dogs-detect> (Accessed: 5 January 2024).
- Ee, A. (2023). 'More Abandoning Pets as Costs to Take Care of Them Rise', *The Straits Times*, 11 May. Available at: <https://www.straitstimes.com/singapore/more-abandoning-pets-as-costs-to-take-care-of-them-rise> (Accessed: 05 January 2024).
- Ekman, P. (1992). 'An argument for basic emotions', *Cognition and Emotion*, 6(3–4), pp. 169–200. doi: 10.1080/02699939208411068.
- Goh, T. (2021). '89% of owners surveyed feel pets help improve mental health during Covid-19 pandemic', *The Straits Times*, 24 December. Available at: <https://www.straitstimes.com/singapore/health/89-of-owners-surveyed-feel-pets-help-improve-mental-health-during-covid-19-pandemic> (Accessed: 05 January 2024).
- Gu, S., Wang, F., Patel, N. P., Bourgeois, J. A. and Huang, J. H. (2019). 'A Model for Basic Emotions Using Observations of Behavior in *Drosophila*', *Frontiers in Psychology*, 10(781). doi: 10.3389/fpsyg.2019.00781.
- Housing & Development Board. (n.d.). *Keeping Pets*. Available at: <https://www.hdb.gov.sg/residential/living-in-an-hdb-flat/keeping-pets> (Accessed: 5 January 2024).
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012). 'Internal representations reveal cultural diversity in expectations of facial expressions of emotion', *Journal of experimental psychology. General*, 141(1), pp. 19–25. doi: 10.1037/a0023463.
- Jang, EH., Byun, S., Park, MS. and Sohn, JH. (2019) 'Reliability of Physiological Responses Induced by Basic Emotions: A Pilot Study', *Journal of Physiological Anthropology*, 38(1), article no: 15. doi:10.1186/s40101-019-0209-y.
- Leomiti MD, Ausman C. (2023). 'EMS Canine Evaluation of Oxygen Saturation (SpO2)', in StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK597366/> (Accessed 5 January 2024).

- Lukac, M., Zhambulova, G., Abdiyeva, K. and Lewis, M. (2023). ‘Study on Emotion Recognition Bias in Different Regional Groups’, *Scientific Reports*, 13 (1), article no: 8414. doi:10.1038/s41598-023-34932-z.
- Mehta, S. A. and Losey, D. P. (2023). ‘Unified Learning from Demonstrations, Corrections, and Preferences during Physical Human-Robot Interaction’, To be published in *ACM Transactions on Human-Robot Interaction* [Preprint]. doi: 10.1145/3623384 (Accessed 5 January 2024).
- Pharmaceutical Society of Singapore (2023) Allergic Rhinitis. Available at: <https://www.healthhub.sg/a-z/diseases-and-conditions/allergic-rhinitis> (Accessed: 5 January 2024).
- Raypole, C. (2020). ‘6 Surprisingly Practical Reasons We Cry’, *Healthline*, 28 August. Available at: <https://www.healthline.com/health/why-do-we-cry> (Accessed: 5 January 2024).
- Siniscalchi, M., d’Ingeo, S. and Quaranta, A. (2018) ‘Orienting Asymmetries and Physiological Reactivity in Dogs’ Response to Human Emotional Faces’, *Learning & Behavior*, 46(4), pp. 574–85. doi: 10.3758/s13420-018-0325-2.
- Søraa, R. A., Tøndel, G., Kharas, M. W. and Serrano, J. A. (2022) ‘What Do Older Adults Want from Social Robots? A Qualitative Research Approach to Human-Robot Interaction (HRI) Studies’, *International Journal of Social Robotics*, 15(3), pp. 411–424. doi: 10.1007/s12369-022-00914-w.
- Tang, T. (2019). ‘How Are You Feeling? Answered with A Real-Time Emotion Detector’, *Medium*, 25 December. Available at: <https://medium.com/analytics-vidhya/building-a-real-time-emotion-detector-towards-machine-with-e-q-c20b17f89220> (Accessed: 5 January 2024).
- Teachable Machine (n.d.) *Teachable Machine*. Available at: <https://teachablemachine.withgoogle.com/> (Accessed: 5 January 2024).
- Teo, J. (2023) ‘S’pore Launches National Mental Health and Well-Being Strategy’, *The Straits Times*, 6 October. Accessed: <https://www.straitstimes.com/singapore/health/s-pore-launches-national-mental-health-and-well-being-strategy> (Accessed: 5 January 2024).
- Wong, S. (2022) ‘Want a Dog? Be Prepared to Set aside over \$100k’, *The Straits Times*, 27 November. Available at: <https://www.straitstimes.com/business/invest/i-ve-always-wanted-a-dog-but-it-could-cost-over-100k> (Accessed: 5 January 2024).



# SME-Capable Innovations-Management-System as a Service: Artificial Intelligence by Click

**Robin Bakir, Maximilian Müller, and Dennis Bakir**

Eskalator c/o Debon Industries SIA, Nachbarsweg 25, 45481 Mülheim an der Ruhr, Germany

## ABSTRACT

The rapid advancement of Artificial Intelligence (AI) is reshaping the landscape of innovation management, especially regarding Small and Medium-Sized Enterprises (SMEs). This paper explores the integration of AI technologies into SMEs' innovation processes, demonstrating how AI can automate complex tasks and enhance operational efficiency and innovation outcomes. Aligned with the structured innovation management processes of DIN EN ISO 56002, encompassing five critical stages— [1] Idea Generation & Evaluation, [2] Concept Development, [3] Development, [4] Prototype Building & Testing, and [5] Production & Market Launch—this study introduces the developed service framework “eskalator.io.” By leveraging Large Language Models (LLMs) and APIs, this innovative approach streamlines data analysis and project evaluation, facilitating a nuanced analysis of customer feedback, technical specifications, and market research data to optimize decision-making. The study addresses challenges in adopting AI technologies, such as security and privacy concerns, emphasizing the importance of ongoing developments for secure and ethical AI integration within SMEs' innovation ecosystems. It aims to contribute to the broader discourse on AI's transformative role in enhancing SMEs' innovation capabilities while proposing future research directions. Common barriers to AI adoption and effective innovation management in SMEs, including lack of technical expertise, administrative burdens, and skepticism about tangible benefits, underscore the need for tailored, user-friendly solutions to encourage broader adoption.

**Keywords:** Artificial intelligence, SME, Innovation-management, Funding autopilot, Data analysis, Workflow automation, Administrative effort

In the landscape of digital transformation, a new horizon is emerging for Small and Medium-Sized Enterprises (SMEs) to enhance their innovation processes. At the forefront of this shift is Artificial Intelligence (AI), which plays a pivotal role in fostering the development of autonomous decision-making systems, thereby revolutionizing innovation management through automation (Keicher et al., 2022). This paper introduces an advanced service framework, designed to streamline the complexities of innovation management, like DIN ISO 56002, for SMEs, leveraging workflow automation. Recognized for its innovation, the system integrates a

highly customizable automation framework aimed at ensuring unparalleled flexibility and scalability.

Through the incorporation of Application Programming Interfaces (APIs) and the utilization of Large Language Models (LLMs), the framework facilitates a deep analysis of customer data, meticulously processing initial data to extract and categorize vital information from customer feedback, technical specifications, and market research. Each component is managed through dedicated processing pathways, setting the stage for an in-depth discussion on how the framework's automation transcends mere expert opinion preparation, covering broader processes such as preliminary funding reviews and the nuanced evaluation of innovative projects.

Participating in this study was a diverse sample of German SMEs from various industries, providing a broad perspective on the impacts of AI integration in innovation management. Prior to the implementation of the AI framework, indicators such as the number of innovations, time to market, and project processing efficiency provided a baseline for evaluating the effectiveness of the AI-driven approach to enhancing the innovation management process. Table 1 showcases the significant improvements observed in key innovation metrics after implementing the AI framework, including a 120% increase in the number of innovations per quarter, a 33% reduction in time to market, and a 75% improvement in the success rate of innovation projects.

**Table 1.** Increased efficiency and improvement in innovation management through AI framework.

Indicator	Before implementation	After implementation	Improvement
Number of innovations per quarter	5	11	+120%
Average time to market launch (days)	540	360	-33%
Number of projects processed	22	39	+77%
Customer satisfaction (1-10 scale)	6.5	8.2	+26%
Matched funding opportunities per idea	0.75	5	+566%
Time saved in project evaluation (hours)	-	272	-
Success rate of innovation projects (%)	40%	70%	+75%

## METHODOLOGY

In this study, we adopted a mixed-methods research approach to meticulously evaluate the deployment and effectiveness of an advanced Artificial Intelligence (AI) service framework in enhancing the innovation management processes within Medium-Sized Enterprises (SMEs) (Makowski & Kajikawa, 2021).

The methodology commenced with a detailed articulation of the research problem, emphasizing the challenges SMEs face in innovation management and the transformative potential of AI to address these issues (Meng et al., 2021). A comprehensive, dual-faceted strategy was employed for data collection and analysis (Tanev & Blackburn, 2022). Quantitative

data and efficiency improvements was systematically gathered through the platform's built-in analytics, while qualitative insights were derived from semi-structured interviews with SME personnel to capture nuanced experiences and perceptions of the AI impact.

This blend of quantitative and qualitative analysis allowed for a rich, multi-dimensional understanding of the effectiveness, ensuring both the statistical rigour of measurable outcomes and the depth of insight into user experiences. Throughout the research process, the study navigated various challenges, such as ensuring the relevance and comprehensiveness of data collected, by implementing standardized protocols for data processing and adapting interview schedules to accommodate participant availability.

The choice of methodology, underpinned by a rigorous review of existing literature on AI applications in innovation management, was justified by the need to achieve a holistic evaluation of the AI framework integration into SMEs' innovation ecosystems, thereby ensuring the reliability, validity, and replicability of the findings.

### **DATA COLLECTION THROUGH AI-ENHANCED AUTOMATION**

The initial phase of integrating the "eskalator.io" framework was focused on the meticulous collection of project data, setting a solid foundation for a transformative ISO 56002 innovation management strategy. Utilizing an advanced Application Programming Interface (API), the framework was designed to efficiently gather detailed information on each project submission. This included customer feedback, technical specifications, and market research data, which were crucial for the subsequent analytical processes. The methodical approach to data collection was not merely about aggregation but ensuring that the data aligned with the strategic needs of the SME, thereby facilitating a targeted analysis that would directly inform the innovation process. This phase was pivotal in ensuring that all necessary information was captured accurately, laying the groundwork for the sophisticated analysis that followed.

The data collection process was significantly enhanced using prompt engineering, which enabled direct and effective as well as GDPR-compliant communication with the AI capabilities, see Figure 1. By precisely defining the criteria and parameters for data collection, the platform could autonomously identify and gather the relevant information required for a comprehensive analysis. This approach not only streamlined the data collection phase but also ensured that the subsequent analysis was based on complete and accurate data. The efficiency and precision in collecting project-related information underscored the company's commitment to leveraging cutting-edge technology to optimize its innovation management practices, thereby enabling a more strategic allocation of resources towards high-potential innovation projects.

### **DATA PROCESSING AND EVALUATION ENHANCEMENT**

Following the collection of project data, the process transitioned into the critical phase of data processing and evaluation, as illustrated in Figure 1,

which showcases an exemplary data analysis module using AI to process the data. Within this stage, the “eskalator.io” framework employed sophisticated AI algorithms to analyze the collected data, leveraging integrated Artificial Intelligence capabilities to autonomously categorize project submissions based on predefined criteria. This process involved an in-depth analysis of the gathered information, where AI-driven insights identified key patterns and deviations. The objective was to distill the vast array of data into actionable intelligence that could inform strategic decision-making and project evaluation. The result was a comprehensive evaluation report that synthesized the separate analysis results in a coherent and sequential order, reflecting the ability to deliver insightful assessments through an automated validation process.

This phase was significantly bolstered by a structured workshop series designed to establish clear interfaces and responsibilities, effectively supporting the operational deployment and integration of the AI system. The initial phase began with setting up the system and conducting a kick-off workshop to align all stakeholders with the framework’s capabilities and expected outcomes. This was followed by three focused work phases during the innovation stage, where iterative reviews and adjustments were made to optimize the framework’s performance based on real-time data and feedback. The final learning phase involved evaluating the experiences and results achieved, allowing the SME to refine its approach and prepare for future projects.

The impact of this advanced data processing capability was profound, dramatically improving the company’s operational efficiency and the effectiveness of its innovation management process based on ISO 56002. By automating the categorization and evaluation of projects, the framework not only reduced the time and resources traditionally required for these tasks but also enhanced the accuracy and reliability of the evaluations. This enabled the SME to quickly adapt to changing market demands and customer needs, fostering a dynamic environment conducive to continuous strategic development. The deployment marked a significant milestone in the company’s pursuit of innovation excellence, showcasing the potential of AI to transform the core aspects of innovation management from data collection to the nuanced evaluation of projects.

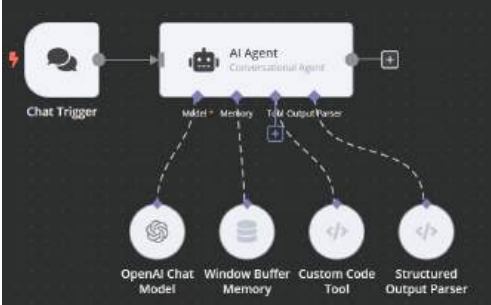
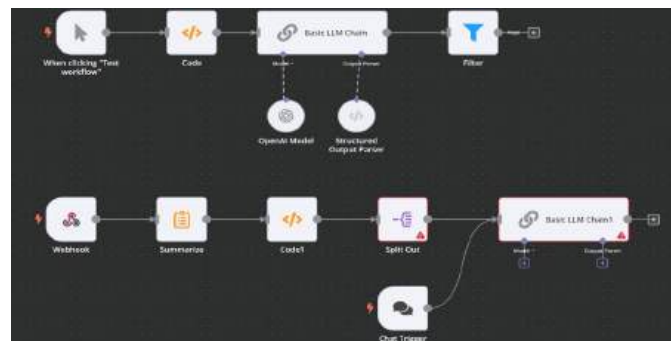


Figure 1: Data analysis module.

## ENHANCING OPERATIONAL EFFICIENCY AND EFFECTIVENESS

The introduction of the automation framework fundamentally transformed the way Medium-Sized Enterprises (SME) are approaching evaluation and management of its innovation projects. At the heart of this transformation was the capability to meticulously process and analyze project data, including customer feedback, technical specifications, and market research. This initial data processing facilitated a structured analysis, enabling the identification of key trends and deviations which were then meticulously categorized according to predefined criteria. Such a detailed evaluation process not only streamlined the assessment of projects but also allowed for the aggregation of individual evaluation elements into a comprehensive evaluation report. This report, generated through an automated validation process, ensured the integrity and completeness of the assessment, setting the foundation for informed decision-making regarding project viability and alignment with innovation goals.

Moreover, the utility extended beyond the initial assessment phase, encompassing critical aspects such as preliminary funding review and the facilitation of funding applications, payment requests, and interim reports. By automating these processes, the framework significantly reduced the administrative workload, freeing up valuable resources to focus on core innovation management activities. The AI framework also supported the enterprise in submitting detailed project information through an API, streamlining the creation of applications for funding and other financial supports. This capability not only optimized the company's operational procedures but also enhanced its ability to communicate efficiently with project holders, ensuring a seamless flow of information and facilitating a more agile response to funding opportunities. Through the strategic implementation of this framework, the SME effectively elevated its operational efficiency and effectiveness, reinforcing its commitment to innovation and securing a competitive edge in the dynamic business landscape. An example of an automated workflow facilitated by this transformation is depicted in Figure 2.



**Figure 2:** Example for an automated workflow.

## **ENHANCING COLLABORATION**

The implementation within the Medium-Sized Enterprise (SME) significantly transformed the collaborative dynamics between the company and us, establishing a highly interactive and responsive framework. One of the key features facilitating this dynamic collaboration was the capability to send real-time notifications and push updates, ensuring both parties remained informed of any changes, missing information, or new insights relevant to ongoing projects. This level of interactivity ensured that the process of innovation management became more collaborative, with an efficient flow of information enabling rapid responses to emerging needs or adjustments required in the project evaluations.

Moreover, the framework was designed with functionalities that allowed for shared access to project data and evaluation reports, enabling both parties to engage in joint project appraisals. This shared access fostered a transparent environment where feedback could be exchanged freely and efficiently, directly contributing to refining innovation strategies, and aligning them more closely with market demands and customer expectations. The integration of these functionalities not only streamlined the communication process but also empowered us to play a more active role in the company's innovation management process, enhancing the collective capability to drive forward strategic initiatives.

## **INTEGRATION AND STRATEGIC DEPLOYMENT**

The strategic implementation of the “eskalator.io” framework fundamentally transformed the production and market introduction phases within the SME. This phase not only supported the internal management and evaluation of innovation projects but also facilitated the seamless introduction of products to the market. By optimizing communication paths and accelerating the time-to-market, the framework significantly enhanced the company's capability to respond dynamically to market demands. Utilizing sophisticated AI algorithms, the framework effectively streamlined the transition from prototype testing to full-scale production, ensuring that projects aligned with strategic goals were rapidly brought to market. This approach not only solidified the company's position in a competitive business environment but also underscored the potential of integrated AI solutions to drive continuous innovation and growth. The operational efficiencies gained through this AI-driven process reflect a significant leap in the company's ability to innovate at scale, proving that strategic AI integration is not just about technological enhancement but also about aligning technological capabilities with broader business objectives and market opportunities.

## **DISCUSSION**

The introduction of an automation framework “eskalator.io” into SME operations illuminates the nuanced interplay between human innovation management and artificial intelligence systems, catalyzing a reevaluation of traditional processes within the framework of digital transformation.

The capacity of AI to autonomously manage and critically evaluate innovation projects, facilitated through advanced API integrations and prompt engineering, heralds a significant advancement towards streamlining the innovation management lifecycle. This fusion of technology and strategy not only democratizes the assessment process but also elevates the efficiency and effectiveness of innovation practices within SMEs. However, this evolution prompts a critical discussion on the integration's breadth, probing into the extent and manner in which AI systems could complement or substitute human roles in the innovation spectrum.

The crux of integrating AI into innovation management hinges on the balance between leveraging AI's capabilities for enhanced decision-making and maintaining the indispensable human insight within the innovation process. While AI can process and analyze vast datasets with unparalleled speed and accuracy, the nuanced judgment and strategic foresight of human managers remain irreplaceable. The automation framework design to incorporate AI-driven analysis and evaluation underscores a collaborative model where technology augments human expertise rather than supplanting it. This collaborative approach not only ensures the alignment of innovation projects with strategic objectives but also fosters a dynamic environment conducive to adaptive and responsive innovation management.

To enhance the acceptance of this framework and align with DIN EN ISO 56002 standards, the system is designed to mitigate common barriers such as technical complexity and administrative burdens. By simplifying the integration and operation processes, the framework aims to improve accessibility for SMEs, encouraging wider adoption and adherence to structured innovation management processes. This is crucial for SMEs that may lack the resources to engage deeply with complex AI systems or ISO standards independently.

Nevertheless, the deployment of such advanced systems within SMEs is not devoid of challenges, particularly concerning security, privacy, and the ethical implications of AI decisions. The integration of robust encryption protocols, access control mechanisms, and continuous security monitoring addresses these concerns, establishing a secure foundation for managing sensitive innovation data. However, the ethical considerations surrounding AI's decision-making processes and transparency necessitate ongoing vigilance and adherence to evolving standards and regulations. Additionally, the collaborative dynamics facilitated spotlight the importance of human-AI interaction, suggesting that the future of effective innovation management lies not only in technological advancement but also in fostering an environment where technology and human expertise coalesce to drive innovation forward.

## **CONCLUSION**

This investigation reveals the pivotal role of Artificial Intelligence (AI) in redefining innovation management within Small and Medium-Sized Enterprises (SMEs). By integrating an AI-enhanced framework, we have illuminated the pathway for SMEs to enhance their management of innovation activities with unprecedented efficiency and effectiveness.

This initiative underscores the capacity of AI to simplify and refine complex innovation processes, catalyzing a shift towards a more vibrant and inventive operational paradigm. Yet, this exploration also brings to the forefront the intricate challenges that accompany the deployment of AI technologies—namely, issues surrounding security, privacy, and the ethical dimensions of AI application. These concerns underscore the necessity of a nuanced approach to integrating AI into the innovation management framework, ensuring that technological advancements are harmonized with rigorous security and ethical standards.

Looking to the future, the potential for expanding the application of AI-enhanced frameworks across varied industry sectors, coupled with the incorporation of more advanced AI functionalities, presents a promising research direction. Such explorations could further elucidate the ways in which AI can be harnessed to provide even deeper insights into innovation projects, thus amplifying its utility within SME innovation ecosystems. Additionally, the examination of user feedback mechanisms and their influence on the iterative refinement of AI frameworks may offer critical insights into how user-centric design principles can amplify the impact of AI systems in supporting organizational innovation. This study advocates for a strategic synthesis of AI capabilities and human expertise as a cornerstone for evolving the domain of innovation management. It posits that a judicious integration of AI not only addresses the limitations inherent in traditional, human-centric approaches but also leverages digital transformation as a conduit for enhancing the creativity, responsiveness, and strategic agility of SMEs. Through this lens, we envision a future where AI acts not as a replacement for human ingenuity but as a powerful ally, propelling SMEs towards a horizon of boundless innovation and growth.

## **ACKNOWLEDGMENT**

The authors would like to thank the Mission Top 5 initiative for their support in advancing Germany's position among Europe's top competitive economies and promoting a digitalized future. Their cooperation has been invaluable in our research on the impact of AI in SMEs.

## **REFERENCES**

- Haefner, Naomi. Wincent, Joakim. Parida, Vinit and Gassmann, Oliver. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda, *Technological Forecasting and Social Change*.
- Hutchinson, Philip (2020). Reinventing Innovation Management: The Impact of Self-Innovating Artificial Intelligence. <https://dx.doi.org/10.1109/TEM.2020.2977222>
- Keicher, Lukas; Ardilio, A.; Nawroth, G. (2022). Boosting Innovation by Artificial Intelligence (AI): Effects and Potentials of AI in the Process of Innovating. <https://dx.doi.org/10.23919/PICMET53225.2022.9882591>
- Makowski, Piotr Tomasz and Kajikawa, Yuya. (2021). Automation-driven innovation management? Toward Innovation-Automation-Strategy cycle, *technological Forecasting and Social Change*.



- 
- Meng, Fansheng; Han, B. (2021). Research on the Impact of Innovation Input on Artificial Intelligence Technology Empowerment and Performance of Manufacturing Enterprises. <https://dx.doi.org/10.1109/EIMSS53851.2021.00134>
- Tanev, Stoyan and Blackburn, Helena. (2022). Artificial Intelligence And Innovation Management, World Scientific, Series on Technology Management Volume 38.

# AI-Based Chatbot Coaching for Interdisciplinary Project Teams: The Acceptance of AI-Based in Comparison to Rule-Based Chatbot Coaching

Vanessa Mai<sup>1</sup>, Johanna Nickel<sup>1</sup>, Anna Gähl<sup>1</sup>,  
Rebecca Rutschmann<sup>2</sup>, and Anja Richert<sup>1</sup>

<sup>1</sup>TH Köln/University of Applied Sciences, Cologne, Germany

<sup>2</sup>evoach, Karlsruhe, Germany

## ABSTRACT

Project-based work is integral in corporate and academic settings, where coaching plays a crucial role in enhancing team performance and project success. To streamline this process and improve scalability, we developed a coaching chatbot at TH Köln/University of Applied Sciences to assist interdisciplinary teams. Utilizing a systemic coaching approach, the chatbot prompts self-reflection through solution-focused questions. We collaboratively created it with student facilitators and lecturers and tested it during a University-wide Interdisciplinary Project Week in November 2023. The pilot study involved two versions of the chatbot: a rule-based system and a hybrid model incorporating generative AI capabilities. As part of the field test, we analysed its acceptance: How effective is the chatbot in supporting projects groups and facilitating reflection processes? Are there differences in acceptance between the two chatbots? Half of the project groups in the one-week course used the rule-based chatbot, while the other half of the project groups were provided with the AI-based chatbot. 134 students participated and used the chatbots at the end of each day of the project week. The results of this study indicate that our test subjects accepted both types of chatbots with moderate to good scores in acceptance. However, the AI-based chatbot fared significantly worse in terms of performance expectancy and effort expectancy. This is possibly due to the fact that hybrid coaching chatbots are neither widely developed nor researched. We conclude that regardless of the technical basis of such a chatbot, conversation design and prompting is an essential part of chatbot development and contributes significantly to acceptance. This study demonstrates the potential of chatbots in supporting group coaching, not only in educational settings but also in corporate environments where they can aid agile project teams. This research marks one of the initial explorations into the acceptance of group coaching through chatbots.

**Keywords:** Chatbot, Coaching, AI-based coaching, Rule-based coaching, Group coaching, Interdisciplinary project work

## **INTRODUCTION: CHALLENGES IN INTERDISCIPLINARY TEAMWORK**

Project-based interdisciplinary work has become a crucial component of the corporate landscape, necessitating robust support systems for teams. Effective teamwork hinges on the provision of continuous guidance and assistance. In this context, coaching emerges as an essential tool, aiding in the reflection on team dynamics and work processes within project settings. In many organizations, process facilitators or internal coaches are employed to help teams address and navigate work-related challenges effectively (Lippmann, 2013).

Navigating interdisciplinary team dynamics in a corporate environment presents its own set of challenges, particularly in digital workplaces. Teams must overcome initial obstacles, gel as a cohesive unit, manage tasks autonomously, and engage collectively in reflective practices. While some teams naturally excel with minimal intervention, others benefit from structured support, appreciating the guidance offered. However, some teams may be hesitant to engage in reflection with an outside facilitator present.

Also, a significant challenge is making reflective processes scalable across large numbers of employees. Digital tools and AI-based technologies, such as chatbots, are helpful in this regard, providing personalized, on-demand support that promotes self-coaching and prepares employees for human to human reflection sessions (Kanatouri, 2020; Mai and Rutschmann, 2023). Studies suggest that chatbots are often perceived as non-judgmental, which can facilitate more open and effective reflection among team members (e.g. Lee et al., 2020).

## **DEVELOPING A COACHING CHATBOT FOR INTERDISCIPLINARY TEAMWORK**

### **Use Case and Concept**

At TH Köln/University of Applied Sciences, our students already train during a Bachelor's degree course to work in interdisciplinary project teams and to meet the challenges described above. Within the "Hochschulweite Interdisziplinäre Projektwoche" (University-wide Interdisciplinary Project Week) students work in interdisciplinary teams for one week. They make joint decisions and develop an understanding of the methods and ways of thinking of other disciplines. The aim of the Interdisciplinary Project Week is to bring the importance and function of interdisciplinary work processes to the fore and to develop an awareness of them.

To accompany the students within this week, we have developed a coaching chatbot as a virtual process facilitator for interdisciplinary project teams. The aim of the coaching chatbot is to provide low-threshold coaching for student project groups in their team and work processes. Based on the systemic coaching approach, it asks solution- and resource-orientated questions that stimulate self-reflection (Berninger-Schäfer, 2018). The concept of a coaching chatbot for students is based on previous

developments of a coaching chatbot on the topic of exam anxiety (Mai et al., 2021; 2023).

We developed the chatbot in a co-creation process with student process facilitators and lecturers and first used it as a prototype in November 2023 as part of the one-week University-wide Interdisciplinary Project Week at TH Köln. The chatbot concept is based on the tasks the student process facilitators fulfil during the project week: Throughout the week, a process facilitator accompanies each group and provides support. Their tasks include supporting the project's start, conducting daily stand-ups and reflection sessions in the afternoon, and evaluating the lecturer's feedback together with the project team. Our coaching chatbot serves as a virtual process companion for the afternoon reflection round. It was used during the project week at the end of each project day from Monday to Thursday. This project is unique in that it involves group reflection with a chatbot.

To ensure an authentic chatbot design and a strong user experience, we collaborated with former process facilitators. Their experience allowed us to integrate cooperative principles into the conversation design to make it authentic. Cooperative principles date back to Paul Grice and describe how contributions to a conversation should align with the conversation's purpose (Grice, 1975). To enable the chatbot to conduct a productive dialog, it utilizes elements of the conversation from the training guide for the process facilitators of the Interdisciplinary Project Week.

The coaching interaction comprises of the following phases: Onboarding, review of the day, goal setting, solution finding, and farewell. Onboarding is an essential component of the chatbot concept and involves welcoming and getting to know the user (Kohne et al., 2020). The coaching phase focuses on reflection methods that follow a solution- and resource-oriented approach (Berninger-Schäfer, 2018). Here, the coaching chatbot uses interventions and methods that the student process facilitators also use, such as scaling questions to assess satisfaction with teamwork.

### **Design Approach: Rule-Based vs. AI-Based**

For our accompanying study, we developed two variants of the coaching chatbot: a rule-based chatbot and a hybrid chatbot that had an interface to generative AI (here: chatGPT) in addition to rule-based processes. We developed the rule-based chatbot using the no coding platform Landbot and the hybrid chatbot in cooperation with the start-up evoach.

Rule-based chatbots are programmed so that the responses generated follow a predefined structure. They therefore follow defined decision paths, which means that they always have the same process. Users primarily control them through selection options and buttons (Stucki et al., 2020). Rule-based chatbots have a less flexible flow of conversation, which often seems somewhat rigid. However, these guardrails also offer advantages: Rule-based bots are secure and predictable; by providing buttons and a clear path for coachees, the communication flow is smoother (Mai and Rutschmann, 2023). We have developed our rule-based chatbot in Landbot. Landbot is a chatbot platform that can be used to develop a chatbot without any

programming knowledge. It can be used to design and implement rule-based chatbots with which users can interact via click-based buttons (Landbot, 2023).

For the hybrid chatbot, we concentrated on a rule-based approach for the onboarding part to ensure consistency and reliability, followed by an AI generated part for the actual self-reflection to ensure a natural conversation flow with the aim to appear less scripted. The chatbots were created on the evoach platform, which can be used to design either scripted, hybrid or fully AI integrated chatbots through the use of a state machine enabling a fully no-code experience to create these chatbots. In order to access the chatbots, participants had to sign-up with the evoach platform. For the onboarding part of the chatbot, we ensured to follow a proposed design framework to create AI coaches (Terblanche, 2020). We especially focused on providing clarity on the process, confidentiality and data policies as well as on providing transparency about the underlying technology used and the risk of generative AI delivering inaccurate or misleading information. This was important in order to manage expectations of the AI's capabilities by being clear on its limitations. We also ensured that all communication generated by AI was labeled visually for transparency and full disclosure, as suggested by (Lee and Choi, 2017).

For the prompting of the AI generated self-reflection parts we designed different prompts for the reflections of each of the project days (day 1, day 2, day 3 and the final reflection on the last project day). The prompts consisted of two parts, the role and the task part, where the role stayed consistent for all prompts while the tasks differed according to the project phases:

1. Role (consistent for all prompts): ensuring the personality and role is defined the AI should incorporate for this conversation by providing a name, role (supporting students in reflecting on their group collaboration), personality (supportive, empathetic, showing concern for the team during the conversation and encouraging them to fully express their feelings).
2. Task (different for each prompt): outlining the clear task the AI should follow for the self-reflection of that day/phase, describing the questions to ask and defining a maximum amount of questions to ask.

## **RELATED WORK AND RESEARCH QUESTION**

In chatbot research, the study of acceptance is central to investigating the extent to which users accept chatbots. To date, there has been little research into the acceptance of chatbot coaching. However, studies on chatbot coaching indicate that users' acceptance of chatbots as coaching tools is highly dependent on the productivity of the chatbot system and the users' performance expectations (Brandtzaeg and Følstad, 2017; Mai et al., 2023; Terblanche and Cilliers, 2020). Crucial factors include performance expectancy, defined as "the extent to which an individual believes that the chatbot will help him or her to achieve performance improvements" (Venkatesh et al., 2003, p. 450) and effort expectancy, described as the

degree of ease associated with using the system (Venkatesh et al., 2003). Davis et al. (1989) found that people are more likely to use an application if they perceive it to be easy to use. This involves not only ensuring the chatbot platform operates smoothly but also that the coaching conversations are rooted in validated coaching methods and dialogue processes to maximize user benefits (Terblanche, 2020). Developers of coaching chatbots need to craft a cohesive concept and bot persona, anchoring their design in clear expectation management. Terblanche and Kidd (2022) show for a goal-attainment coaching chatbot that the willingness to use it depends not only on how much the chatbot supports them in achieving their goals, but also on how much other people are willing to use a coaching chatbot and how easy and convenient it is to use.

Goal achievement is a focus area of coaching that leads to increased performance, progress and achievement of outcomes (Grant, 2012). The purpose of the chatbot in the present study was specifically to guide group reflection and facilitate the process. Since the construct of performance expectancy measures perceived performance growth (Terblanche, 2020), and since daily reflection with the chatbot is related to the achievement of the project goal, we hypothesize that performance expectancy has a significant impact on individual intention to use the chatbot for group reflection. Moreover, we hypothesize that effort expectancy influences the individual intention of the students to use a chatbot for group reflection.

As part of our field test, we therefore analyzed the acceptance of our coaching chatbot – measured with the items performance expectancy and effort expectancy – for interdisciplinary project work and formulated the following research questions:

- How effective is the chatbot in supporting projects groups and facilitating reflection processes?
- Are there differences in acceptance between the two chatbots (rule-based vs. AI-based)?

## RESEARCH DESIGN

The experimental design of this study consists of a combination of chatbot coaching with the developed and programmed coaching chatbots and a survey. A questionnaire was used to capture the student's perceived acceptance of the chatbot coaching. Half of the project groups in the one-week course interacted with the rule-based chatbot, while the other half of the project groups were provided with the AI-based chatbot. The participants were asked to interact with the coaching chatbot at the end of every project day for the daily stand-up reflection session; however, they were free to decide whether they wanted to use the coaching chatbot. In addition, there was support from human process facilitators – independent of the use of the chatbot. The students were further asked to fill out the questionnaire at the end of the first and the last day of the project week (Monday and Thursday).

Frameworks such as the Technology Acceptance Model (TAM) or the Unified Theory of Acceptance and Use of Technology (UTAUT) are valuable

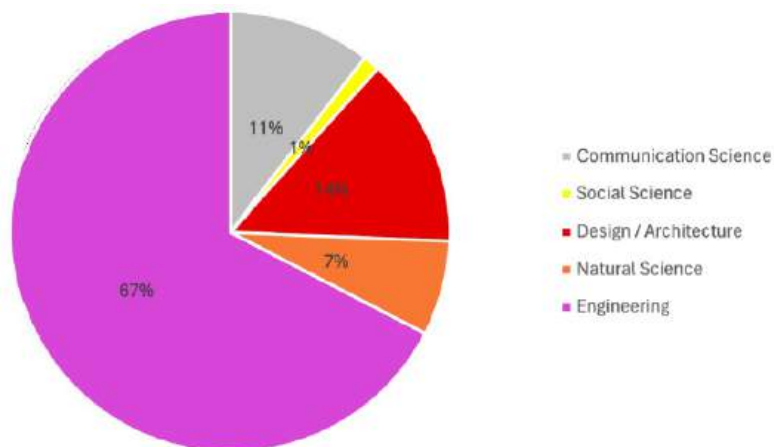
methods for assessing technology acceptance (Venkatesh et al., 2003). From the UTAUT, we selected the two constructs performance expectancy and effort expectancy, which were measured with a total of nine items. Demographic data such as gender, age, and field of study were collected at the end of the questionnaire.

The answers to all items on acceptance were given on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Reliability according to Cronbach's alpha is given for both constructs: For performance expectancy (PE).89 and for effort expectancy (EE).61. The data were analysed using descriptive and inductive statistical methods. For each construct (PE, EE), we conducted a t-test to measure significant differences in the experimental groups (rule-based vs. AI-based chatbot). The significance level for the evaluation of the data in this study is 5%.

## RESULTS AND DISCUSSION

### Sample Description

A total of 134 students interacted with the chatbots. The individual project groups consisted of an average of 10 students from different disciplines. The following diagram (Figure 1) shows the disciplines represented in the project groups. They were formed in such a way that students from every discipline were represented in each group. The youngest study participant was 19 and the oldest 39 years old. The average age of the study participants was 21.5. 14.3 % of students were female, 43.7 % were male and 1.3 % were other/both/neither/interchangeable.



**Figure 1:** Disciplinary affiliation of the participants in the University-wide interdisciplinary project week. The project groups were formed in such a way that students from every discipline were represented in each group.

### Results: (AI-Based) Chatbot Coaching for Interdisciplinary Project Teams Is Accepted

Table 1 illustrates the results of the questionnaire survey on acceptance, with the constructs performance expectancy (PE) and effort expectancy (EE), conducted at the end of the project's first day. Both constructs indicate moderate to high levels of acceptance. In terms of acceptance, the rule-based chatbot outperforms the AI-based chatbot for both PE and EE, with mean scores of 3.26 and 3.81 respectively, compared to 3.10 and 3.60 for the AI-based chatbot. However, only the differences in EE are statistically significant, with the rule-based chatbot showing significantly higher scores than the AI-based chatbot.

**Table 1.** Statistical analysis of the questionnaire results at the end of the first project day (acceptance) (N = 134).

First project day (Monday)	Performance Expectancy (PE, 5 items) – Mean (Standard Deviation)	Effort Expectancy (EE, 4 items) – Mean (Standard Deviation)
Rule-based Chatbot (N = 56)	3.26 (.86)	3.81 (.44)
AI-based Chatbot (N = 78)	3.10 (.90)	3.60 (.58)
p-value (p<.05)	.31 (not significant)	.02 (significant)

Table 2 illustrates the results of the questionnaire survey on acceptance, with the constructs performance expectancy (PE) and effort expectancy (EE), conducted at the end of the project's last day. Here as well, both constructs indicate moderate to high levels of acceptance. Overall, the results of the last project day confirm the results of the first day.

The acceptance values for the rule-based chatbot have slightly decreased, with mean values of 3.01 for PE and 3.74 for EE. However, for the AI-based chatbot, these values have significantly dropped, particularly for PE, indicating significant differences compared to the rule-based chatbot. In terms of acceptance, the rule-based chatbot outperforms the AI-based chatbot for both PE and EE statistically significant, with mean scores of 3.01 and 3.74 respectively, compared to 2.22 and 3.38 for the AI-based chatbot.

Notably, the PE values for the AI-based chatbot plummeted from 3.10 on the first day to 2.22 by the project's end. Additionally, the data shows that the number of participants in the survey decreased from 134 to 86 from the beginning to the end of the project.

**Table 2.** Statistical analysis of the questionnaire results at the end of the last project day (acceptance) (N = 86).

End of last project day (Thursday)	Performance Expectancy (PE, 5 items) – Mean (Standard Deviation)	Effort Expectancy (EE, 4 items) – Mean (Standard Deviation)
Rule-based Chatbot (N = 51)	3.01 (.99)	3.74 (.69)
AI-based Chatbot (N = 35)	2.22 (.90)	3.38 (.49)
p-value (p<.05)	.0006 (significant)	.01 (significant)



## DISCUSSION

The results enable us to infer levels of acceptance and user needs. They reveal that chatbot coaching is well-received by interdisciplinary project teams and suggest guidelines for designing such chatbots.

Survey outcomes indicate both types of chatbots (rule-based and AI-based) are accepted with moderate to good scores in performance expectancy (PE) and effort expectancy (EE). This aligns with findings that the acceptance of chatbots, particularly for coaching, relies on a coherent concept and clear expectation management (Mai et al., 2022; 2023; Terblanche, 2020; Terblanche and Cilliers, 2020). For our chatbots, we assessed the needs of the target group beforehand, collaboratively developed the chatbot concept with participants (student process facilitators), and refined it through several feedback cycles.

We encountered unexpected aspects in our study. Firstly, contrary to our expectations, the AI-based chatbot did not receive higher acceptance than the rule-based one, despite its capability for more flexible, personalized responses via AI-generated answers. This may be due to its hybrid design, where AI-generated responses were only partially used, which did not significantly affect acceptance. Nonetheless, this underscores the adequacy of the chatbot concept and design for our purposes, effectively supporting students in reflecting on their project work. Our first research question, regarding the chatbot's effectiveness in supporting project groups and facilitating reflection processes, can be affirmatively answered.

As for our second question about differences in acceptance between the two chatbots, the answer is also partially affirmative. Surprisingly, the AI-based chatbot fared significantly worse in terms of effort expectancy, which measures ease of interaction. At the end of the project week, the AI-based chatbot significantly underperformed also in terms of performance expectancy. Although both chatbots shared the same underlying concept and script, their differing user interfaces might explain this variance. The rule-based chatbot was more accessible, requiring no registration and being directly usable via a URL. In contrast, the AI-based chatbot required registration and login via a platform, complicating access. This may also explain the drop in the number of participants between the first and last day of the project – which fell in particular for the AI-based chatbot (78 vs. 35 participants in the survey, a drop-out rate of more than 50%).

These findings echo research suggesting that a chatbot's ease of use significantly impacts user willingness to engage (Terblanche and Kidd, 2022). Another reason could be that there has been little development and research into hybrid coaching chatbots yet – consisting of a rule-based foundation and interfaces to generative AI – as used in this study. This is therefore a very early prototype which might have affected its acceptance.

## LIMITATIONS, FUTURE RESEARCH AND DESIGN IMPLICATIONS

This study is subject to several limitations regarding its research design. Our study focused on performance expectancy (PE) and effort expectancy (EE), and as a result, we did not collect other acceptability items. Consequently, no comparisons can be made between this study and other studies that have

used the full UTAUT construct. To enable comparisons, it would be necessary to use the entire UTAUT construct in the next study.

Another limitation is access to the study: the study participants used the chatbot in an online environment. One group member interacted with the chatbot, while the other group members observed the chat display via Zoom. In other words, the entire group collaborated with the chatbot, but only one person operated it. There may be differences in the acceptance of the chatbot depending on whether someone interacted with it themselves or “just” watched. This should be taken into account in future studies.

For subsequent studies, it is also essential to pay particular attention to the conversations histories with the chatbot. These could be utilised to draw further valuable conclusions about user acceptance. The study also shows limitations with regard to the chatbot design. The chatbot interaction was quite short, which may have had an impact on the acceptance values.

Moreover, the findings offer insights into the technical and conceptual design of coaching chatbots for interdisciplinary teamwork and guide future research directions. Next, we plan to assess a more developed prototype of the AI-based chatbot and explore its impact on acceptance. We will also evaluate other variables like effectiveness with regard to relationship design (e.g. working alliance, social presence). Further, analyzing the conversations histories of the chats will help identify challenges in dialogue and suggest improvements for conversational design. Insights into the content will shed light on the bot’s effectiveness in fostering reflection.

Future research should also investigate how groups utilize such a chatbot: whether it’s typically one person responding on behalf of the group, if its use promotes group dialogue, and how it compares to interaction with a human facilitator in terms of helpfulness and limitations.

## **CONCLUSION**

The results of this study show that coaching chatbots for interdisciplinary teamwork are accepted. They also show that regardless of the technical basis of such a chatbot (rule-based vs. AI-based), conversation design and prompting is an essential part of chatbot development and contributes significantly to acceptance.

This is one of the first studies to show the acceptance of group coaching with a chatbot. This means we will keep refining and researching our chatbot. Beyond its application in student project settings, there are numerous other potential uses, such as aiding and assisting agile project teams in the industrial sector.

## **ACKNOWLEDGMENT**

The teaching project of this study was supported by the Foundation for Innovation in Higher Education as part of the REDiEE project. The aim of REDiEE is to promote the development of hybrid approaches to teaching and learning at Cologne University of Applied Sciences (TH Köln) that combine “future skills” and specialist expertise in order to strengthen the profile of its alumni.

The authors would furthermore like to acknowledge for the support in conceptualizing our chatbots provided by our student assistants and process facilitators, Rebecca Broich and Melisa Cansiz. We also extend our thanks to our colleague Christian Deggelmann for his assistance in analyzing the data.

## REFERENCES

- Berninger-Schäfer, E. (2018). Online-Coaching. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-10128-2>
- Brandtzaeg, P. B., & Følstad, A. (2017). Why People Use Chatbots. In I. Kompatsiaris, J. Cave, A. Satsiou, G. Carle, A. Passani, E. Kontopoulos, S. Diplaris, & D. McMillan (Eds.), *Lecture Notes in Computer Science. Internet Science* (Vol. 10673, pp. 377–392). Springer International Publishing. [https://doi.org/10.1007/978-3-319-70284-1\\_30](https://doi.org/10.1007/978-3-319-70284-1_30)
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
- Grant, A. M. (2012). An integrated model of goal-focused coaching: An evidence-based framework for teaching and practice. *International Coaching Psychology Review*, 7(2), 146–165. [https://www.academia.edu/18048251/An\\_integrated\\_model\\_of\\_goal\\_focused\\_coaching](https://www.academia.edu/18048251/An_integrated_model_of_goal_focused_coaching)
- Grice, H. P. (1975). *Logic and conversation*. Academic Press, New York.
- Kanatouri, S. (2020). *The Digital Coach*. Routledge. <https://doi.org/10.4324/9780429022753>
- Kohne, A., Kleinmanns, P., Rolf, C., & Beck, M. (2020). Conversation Design. In A. Kohne, P. Kleinmanns, C. Rolf, & M. Beck (Eds.), *Chatbots: Aufbau und Anwendungsmöglichkeiten von autonomen Sprachassistenten* (pp. 83–97). Springer Vieweg. [https://doi.org/10.1007/978-3-658-28849-5\\_5](https://doi.org/10.1007/978-3-658-28849-5_5)
- Landbot. (2023). Landbot Knowledge Center. <https://landbot.io/knowledge-center>
- Lee, Y.-C., Yamashita, N., Huang, Y., & Fu, W. (2020). “I Hear You, I Feel You”: Encouraging Deep Self-disclosure through a Chatbot. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguy, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). ACM. <https://doi.org/10.1145/3313831.3376175>
- Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human Computer Studies*, 103, 95–105.
- Lippmann, E. (2013). *Coaching: Angewandte Psychologie für die Beratungspraxis*. Springer. <https://doi.org/10.1007/978-3-642-35921-7>
- Mai, V., & Rutschmann, R. (2023). Chatbots im Coaching. Potenziale und Einsatzmöglichkeiten von digitalen Coaching-Begleitern und Assistenten. *Organisationsberatung, Supervision, Coaching*, 30(1), 45–57. <https://doi.org/10.1007/s11613-022-00801-3>
- Mai, V., Bauer, A., Deggelmann, C., & Richert, A. (2023). Acceptance and User Needs of Coaching Chatbots: An Empirical Analysis of a StudiCoachBot’s Conversation Histories. In P. Zaphiris, A. Ioannou, R. A. Sottolare, J. Schwarz, F. Fui-Hoon Nah, K. Siau, J. Wei, & G. Salvendy (Eds.), *Lecture Notes in Computer Science: Vol. 14060, Hci International 2023 – Late Breaking Papers: 25th International Conference on Human-Computer Interaction, HCII 2023, Copenhagen, Denmark, July 23–28, 2023, Proceedings, Part VII* (1st ed. 2023, pp. 176–190). [https://doi.org/10.1007/978-3-031-48060-7\\_14](https://doi.org/10.1007/978-3-031-48060-7_14)

- Mai, V., Neef, C., Richert, A. (2022). "Clicking vs. Writing": The Impact of a Chatbot's Interaction Method on the Working Alliance in AI-based Coaching. *Coaching: Theorie & Praxis* (8). 1–17. <https://doi.org/10.1365/s40896-021-00063-3>
- Mai V., Wolff A., Richert A., Preusser I. (2021). Accompanying Reflection Processes by an AI-based StudiCoachBot: A Study on Rapport Building in Human-Machine Coaching Using Self Disclosure. In *HCI International 2021 – Late Breaking Papers: Cognition, Inclusion, Learning, and Culture*. Lecture Notes in Computer Science. 439–457. [https://doi.org/10.1007/978-3-030-90328-2\\_29](https://doi.org/10.1007/978-3-030-90328-2_29)
- Stucki, T., D'Onofrio, S., & Portmann, E. (2020). Chatbots gestalten mit Praxisbeispielen der Schweizerischen Post: HMD Best Paper Award 2018. Springer eBooks Computer Science and Engineering. Springer Vieweg. <https://doi.org/10.1007/978-3-658-28586-9>
- Terblanche, N. (2020). A design framework to create Artificial Intelligence Coaches. Advance online publication. <https://doi.org/10.24384/B7GS-3H05>
- Terblanche, N., & Cillers, D. (2020). Factors that influence users' adoption of being coached by an Artificial Intelligence Coach. *Philosophy of Coaching: An International Journal*, 5(1), 61–70. <https://doi.org/10.22316/poc/05.1.06>
- Terblanche, N., & Kidd, M. (2022). Adoption Factors and Moderating Effects of Age and Gender That Influence the Intention to Use a Non-Directive Reflective Coaching Chatbot. *SAGE Open*, 12(2). <https://doi.org/10.1177/21582440221096136>
- Venkatesh, Morris, & Davis (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425. <https://doi.org/10.2307/30036540>

---

# Unravelling the Aesthetics and Emotion: Exploring the Artistic Value of AI-Generated Artworks

**Amic G. Ho**

Hong Kong Metropolitan University, School of Arts & Social Sciences, Department of Creative Arts, Hong Kong, 00852

## ABSTRACT

The debut of AI art, a new cultural realm, has given rise to discussions about its artistry and aesthetic value. On the other hand, since the new path of the art-creational movement does not allow applying traditional components of creativity to evaluate the creative and aesthetic value, art coming from AI creates new difficulties concerning its evaluation. What are the pillars of the art genre – composition, technique, visual phenomena – no longer look the same? New tools are required to give a scientifically sound assessment of the esprit and visual components of AI art. How do we assess emotions driven by works from AI? They may show the extent of faithfulness to the traditional genre due to the same or diverse components from conventional art. This paper tries to compose and define the base on which an assessment of emotions driven by such works is to be done. When the artist agrees, AI uses machine learning algorithms, and the experience is always learning.

**Keywords:** AI-generated artworks, Artistic value, Aesthetic value, Evaluation criteria, Emotional impact

## INTRODUCTION

The emergence of artificial intelligence has introduced a novel mode of artistic expression known as AI art, which prompts investigations into its creative essence and aesthetic worth. The evaluative analysis of creative and aesthetic value in AI-generating works becomes challenging due to the uniqueness of its development stages. The current criteria for assessing art such as the composition, technique, and visual appeal may need a form of revision to analyse the artistic creativeness of the artworks created through artificial intelligence. A comprehensive understanding of the creative and aesthetic value depicted by art in the traditional realms provides a framework that necessitates the development of new criteria. The real question is whether the artworks made through artificial intelligence can elicit emotional feelings similar to those experienced in traditional art forms. The importance of emotions in artistic performances is the creation of emotional connections and the development of further choices. The identification of possibility of artificial intelligence replicating or developing more engaging emotions than traditional art can be identified based on analysing the emotional reaction in

AI-generated art. The collection of data from any persons interacting with AI-generated artworks, through the use of both qualitative and quantitative methods such as surveys and aesthetic standard compilation will enable the real potential of the artwork. This research will support the identification and analysis of emotional feelings in AI-generated art and enable the field to analyse the possible differences between traditional forms of artworks that generate emotional feelings. Both the possible commonalities and differences between AI-generated artworks are determined.

### **AI-GENERATED ART AND ITS RISE AS A NEW FORM OF ARTISTIC EXPRESSION**

The advent of a new kind of art, AI art, prompted debates on its creativity and artistic value. The novelty of creativity and aesthetic value assessment in works of AI generation is associated with the evolutionary instability of the species. To evaluate the alien subtlety in AI works, one needs a new system that includes criteria and assessment tools. This system may retain the traditional parameters – such as composition, drawing and visual appeal – because creation is always creation. The ability of an alien to cause feeling is a sensation – human from aliens – extraterrestrial. The importance of these feelings can be as a means of relationship and as a uniqueness of both interacting parties. A study of how people “perceive” more creative and specific works in the form of a picture. However, to do this, one needs an idea of how AI art will attract people. One can evaluate revelation through the same commission; focus depends on the criterion. We can get a view of the replication or deviation of the emotions for old ones from AI visual art. To do this, I collect data from people who interact with AI works of art. Data collection is through both qualitative and quantitative research methods, including question and answer response and aesthetic analysis. The purpose of this research is to identify and study the feelings of A in exposure to AI work and enable a comparative investigation of the feelings that old get from them. This comparative criterion can be the same for acceptance and incompatibility in feeling.

### **CHALLENGES IN EVALUATING CREATIVE AND AESTHETIC VALUE IN AI ART**

The assessment of creative and aesthetic value in AI art is a complex phenomenon, demanding the discussion of ethical, social, and artistic implications. The recent debate on the creativity and beauty of AI was sparked by the artwork that is generated by AI (Liu, 2023). While others claim that AI-created art has aesthetic efficacy (Ullrich & Trump, 2022), the investigation of things like computer-generated art using conventional assessment techniques is insufficient to understand and appreciate them, respectively. Ethical implications, the creative procedure, aesthetics, and audience acceptance are also explored when AI is utilised in art (Ho, 2024b). Finally, there is evidence indicating that comparing human and AI-made

art improves human creativity (Hwang, 2022), which could inflate the worth of human labour (Horton et al., 2023). The study conducted by Bellaiche et al. (2023) demonstrates that people generally have a negative tendency towards artworks created by artificial intelligence. This discovery implies that the participation of humans in the creative process has a positive effect on the assessment of art, indicating that AI may not entirely replace human creativity. The advancement of AI technology presents both advantages and difficulties in the development, encounter, and recognition of art (Lyu et al., 2021). Moreover, as it demonstrated, AI can provide new perspectives on aesthetic, personality, and social components of cultural creation, which are less focused on the commodity side of cultural production and intellectual property (L'Yi & Gehlenborg, 2022). In addition, the evaluation of AI art assumes creating computational aesthetic judgment to motivate artists and graphic designers, which would be able to argue that the automatic generation of art is relevant. The evaluation, purchasing intention, and collecting intention commerce of the artwork can also be evidenced in the personality of the creator (Zhang et al., 2023), who can be human or artificial. The identity of the creator changes the perception of art, hence this factor is important for the evaluation of creative and aesthetic value (Gu & Li, 2021). As a result, ethical, social, and artistic implications should be taken into during the creative and aesthetic evaluation of AI art. Therefore, as the role of AI in creative production increases, new conditions and values emerge, which necessitates a more detailed and full evaluation of AI art on creative and aesthetic grounds.

### **NECESSITY FOR A NOVEL FRAMEWORK TO ASSESS AI-GENERATED ARTWORKS**

A recent study argues that there is a need for a new framework to assess AI-generated artworks. The study acknowledges that the existing assessment methods are complex, but fails to address the inherent prejudice against computer-generated art. Therefore, the study suggests that the bias identification paradigm should be included in the review process (Hosseini, Resnik & Holmes, 2023). However, this work also raises the question of difference and whether the recognition of whether art is done by AI if it belongs to the same cultural spectrum has an incidence on perception. Moreover, a recent analysis with division handling and a prospective study on AI and art explains the need to develop an ethical frame and a collaborative paradigm to evaluate the AI art influence on preference and action intention sighting productivity framework which is also related to the creator's identity (Chamberlain et al., 2018; Gangadharbatla, 2021). A need for a broad aesthetic analysis (Bellaiche et al., 2023) regarding various AI types of generation is also highlighted in a work involving AI-generated dances and image-to-text (Darda, Carre & Cross, 2022). AI-created paintings and experiments based on measuring the cognitive aspect of audience perception of art style transference. The work states that a cue framework measuring

would not be enough to analyse how people perceive and experience AI art and that a new framework should be developed.

### **CONVENTIONAL CRITERIA USED FOR ASSESSING TRADITIONAL ART AND THE NEED FOR REVISION**

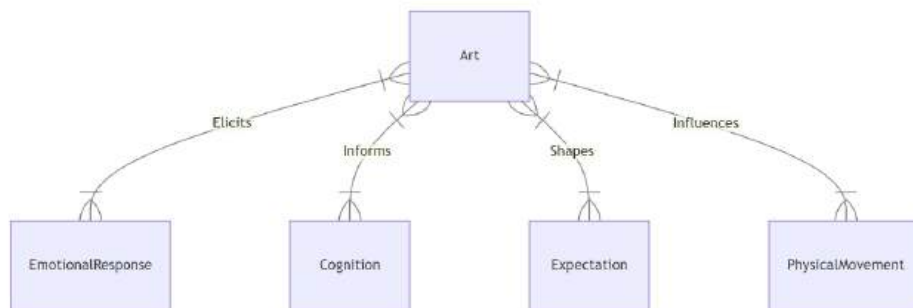
In the past years, traditional art metrics have come under scrutiny. With AI-generated art gaining momentum, there has been a call to reconsider what these criteria mean, specifically in the context of art that is created by artificial intelligence. Studies have uncovered biases against computer-generated art, arguing that a framework that accounts for these biases - and knowledge about AI-attribution impact on artwork evaluation - is necessary and urgent. The previous review and forward-looking evaluation of AI and art underscores the need for the establishment of ethical guidelines coupled with a further exploration of collaborative practices in the evaluation of AI-based art. The work on how human expertise and AI-created artefacts could come together also emphasises the necessity for a framework that also considers the identity of the creator. Furthermore, A previous piece of work seems to suggest that a framework that captures the nuances of the audience reception and experience in AI art will be able to pick up on the effects of audience expertise on the evaluation of artwork that was generated through style transfer by machine learning. The examination of paintings produced by a 'text-to-image system' (Darda, Carre & Cross, 2022) as well as the squid game and the poet's game previously discussed, suggests the need for a framework that enables us to contemplate models of viewer engagement and aesthetic responses across various AI-art domains. The works in the existing literature call for a novel framework to evaluate AI-generated artworks that can consider biases, AI attribution knowledge, ethical considerations, creator identity, and audience perception, should be comprehensive enough to account for various forms of AI-generated art and comprehensive enough to capture the complexities of human-AI coproducing these works.

### **EXPLORATION OF EMOTIONAL RESPONSES TO ART AND ITS IMPORTANCE IN CREATIVE EXPERIENCES**

The subject of emotional responses to art is vast and diverse. Ho (Ho, 2024a) has recently suggested that emotional responses are primarily responsible for positive emotions in response to art. At the same time, other studies (Menninghaus et al., 2019) prove the significance of the negative emotional dimensions to the aesthetic experience. Furthermore, emotions are affected by many cognitive evaluations that prompt the emotional experience and expression due to the work or portrayal, highlighting the connection between cognition and emotions responding to art (Miu et al., 2016; Schindler et al., 2017). Several other studies have also indicated the differences between art forms on the grounds of what emotional responses they may evoke (Menninghaus et al., 2019). Imitation of these emotions is also consistent with the distancing-embracing model (Pelowski et al., 2020).



Various scholarships have also examined the factors influencing positive emotions in response to music and painting (Silvia, 2012), highlighting the differences between various art forms. In addition, cognitive appraisals have also been found to impact one's interest in visual arts (Ho, 2014), suggesting that cognitive processes are also responsible for the emergence of aesthetic emotions. Finally, the proposed research has inspired the formulation that emotional responsiveness to art is diverse, spanning many dimensions of emotions, cognitive evaluations, and responsiveness. Thus, to thoroughly understand the emotional part of creativity, it is essential to acknowledge the vast number of emotional responses to art. Different levels of exposure to these factors explain new emotions. Furthermore, Ho's (2014) finding reveals that there are also some cognitive assessments which determine one's interest in visual arts. Therefore, cognition is another important factor in determining the roots of aesthetic production. Hence, in seeking to acquire a more comprehensive view of emotional development, it is essential to take into account the varied and multidetermined nature of emotional responses to art.



**Figure 1:** Entity relationship diagram to explain the interplay between art, emotional responses, cognition, expectations, and physical movement.

## RESEARCH METHODS

### Research Study Emotional Engagement With AI-Generated Art vs. Traditional Art

This study aimed to contrast emotional engagement in two types of art forms -- AI-generated art and traditional art. It provides an understanding of the emotional responses each evokes and how each type affects people's feelings – both at the individual level and as a whole society. We will use a combination of qualitative (through the SAM self-assessment matrix) (Chen, Chang & Liang, 2019). and quantitative approaches (questionnaire) to get at people's feelings about AI-generated art and traditional artworks. Surveys will be used for collecting quantitative data on emotions while qualitative

methods such as interviews and participant observation yield deep insight into what participants experienced. A wide variety of artworks from both sources were chosen to give as complete an analysis as possible. The study uncovered differences and similarities in emotional engagement between AI-generated art and traditional art through the SAM self-assessment matrix. It means to provide a more complete understanding of the different ways in which AI-generated art induces emotional connection or engagement compared to traditional art. This knowledge will also help build up a picture of how AI-generated art affects an individual's emotional experience and the transformation that can be expected to occur to emotion within that context of art.

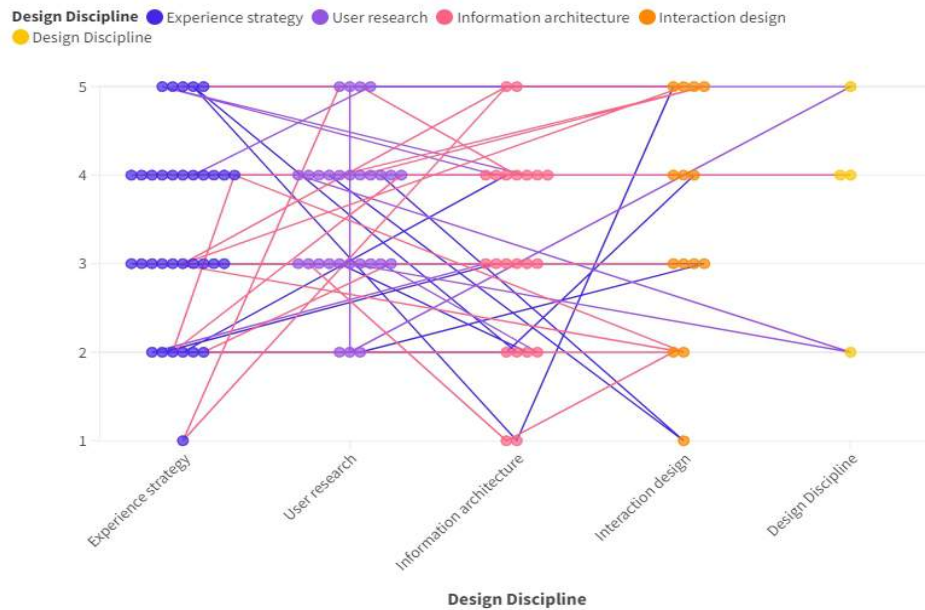
### **Participants**

For both research studies, the participants should meet the following criteria. Participants must be at least 18 years old and possess the ability to give informed permission. Additionally, they should have a sophisticated understanding of art and emotional experiences. They would be randomly invited through the online platform's computer art discussion groups. Participants with broad, cultural, educational, and artistic backgrounds ought to be involved to capture diverse perspectives and experiences. To that end, professionals, amateur artists, and those with no artistic background shall be represented. To fulfil this existential objective, it is necessary to include people who have varying levels of artistic experience. The minimum technology literacy to ensure that subjects will be able to interact with AI-generated art without hesitation and assist them put things in perspective.

### **FINDINGS**

A comprehensive examination of descriptive analysis depicts a slightly higher mean for all dimensions from traditional art as compared to AI art. Regarding emotional engagement, descriptive analysis revealed that the mean was 4.5 standard deviation in Figure 2. The highest mean for subjective preferences was 6.3 and the lowest for perceived emotional impact and artistic appeal have means of 4.2 and 6.1, respectively. However, participants reported a connection of 4.1, a total emotional impact of 6.4, and a mean of 4.2. On the other hand, all the mean scores of subjective parameters of AI art were slightly lower than the mean scores of traditional art. The emotional engagement mean was 4.2, subjectiveness was 5.8, and the perceived impact values were 3.9 with the 5.6 artistic appeal and a connection of 3.9 and an overall emotional value of 6.2. Significantly, a comparative analysis using a paired t-test revealed no substantial variations in the average values between the two categories. Irrespective of slightly different standard means, the dataset's overall trend suggests a similar emotional response towards traditional and AI-generated artwork. Further analysis through the cluster can differentiate participants providing differing emotional engagement patterns to AI art and traditional art. Such analyses will provide

more insights into the complex emotional experiences among different participants.



**Figure 2:** The impact of emotional design on communication effectiveness through different design disciplines.

## DISCUSSION

### Advancement in Understanding the Emotional Impact of AI-Generated Art

The emotional dimension of AI-generated art has also been at the centre of recent research. For example, studies on the affective responses to computer-generated art have also been done because of how creative acts, which have always been valued highly since they were believed to be unique to human beings, can be valued when they are produced by a machine (Chamberlain et al., 2018). Hence, employing features inspired by psychology and art theory for affective picture classification can significantly aid in conducting an affective analysis of images, specifically in identifying the ability of images to evoke emotions in humans (Machajdik & Hanbury, 2010). Recent research also provided the impact of AI-generated art concerning the upcoming era of artificial intelligence AI and how it will affect normal societal norms and aesthetics (Rodgers et al., 2023). Others include the evidence on the psychological effect of AI awareness on employee depression, showing that changes in AI technology result in different emotions in employee practices (Chen et al., 2024). In addition, AI art is increasingly

interesting on its own accord, since it represents humans creating devices that can create images never seen before, researchers citing how the real picture only can be made because of emotional outcomes (Chatterjee, 2022). McGuirt et al. discuss how AI works in images and everyone is very sure about the picture taken as the image, but they found that there is emotion every time used AI-generated image. They also documented how the AI attribution knowledge affected the evaluation of artwork that humans produce, suggesting there are emotions that humans bring in creating images that machines cannot attain (Gangadharbatla, 2022). The finding meant to show how AI was associated with the social-relational moral standing that suggests the normative debates on robot rights (Lima et al., 2021). Further, there were investigations into the emotional actions of AI service measures as there were twelve studies that examined emotions arising from AI service measures, and explained how the emotions can be divided into three categories (Bagozzi et al., 2022). These reviews suggest there were various emotional and cognitive processes included in the recent research on the emotional dimension of AI-generated art.

### **Facilitation of a Comparative Analysis of Emotional Experiences Between AI-Generated Artworks and Traditional Art**

The ability to conduct a comparative analysis of the emotional experience related to AI-generated artworks and traditional art is contingent on the understanding and consideration of multiple factors and perspectives. The nature of emotional responses to art is multi-faceted and complex, conditioned by numerous psychological, social, and cultural factors. The respective references present several factors affecting the emotional experience related to AI-generated art and traditional art that enable a reasonable comparison. Thus, the reflection based on the SAM self-assessment matrix and the factorial related to the aesthetic judgment, identified via the feedback from the participants, enlightened the underlying biases towards AI artworks and the challenges of identifying them properly in blind comparison. It is particularly relevant in light of the unique forms and patterns of emotional response identified for AI art. Additionally, the work of Cotter et al. (2023) provides valuable information about the emotional experience and flourishing during the visitation of art museums, thus accentuating the diverse patterns of emotional response to traditional art. The study conducted by Gangadharbatla et al. (2021) played a crucial role in the present evaluation by examining how awareness of AI attribution affects the assessment of artwork across individuals from different age groups, specifically Gen Z and Millennials. Finally, the study of Yusa et al. (2022) comprehensively analysed the concept of AI-generated art in all its dimensions. Thus, combining these references enables the comparative analysis of multiple factors, such as biases, attribution knowledge, and the specifics of emotional responses to both forms of art.

## Recommendations for Future Research in the Field of AI Art and Emotions

Given the above references from the data, AI art and emotions may be focused in several key directions. For example, such work as Suhaimi et al. (2020) could be supplemented if innovative methods such as EEG-based emotion recognition able to capture an individual's emotional responses to AI-generated art were developed; this would make it clear how an individual works with his emotions and help to form emotionally intelligent AI systems. Also, it could be the investigation of the emotional consequences of human-AI interaction in various spheres from the provision of services and marketing to entertainment and recreation use of leisure time (Huang & Rust, 2018). One more future research might be the development of emotion profiling tools using psychological lexicons, artificial intelligence and network science to investigate not only the emotional states but also the emotional content and effect of AI-generated content (Hua et al., 2024). Another possible research question might be whether people prefer human- or AI-created artwork in the whole creative field (Bellaiche et al., 2023) would give insights into how people evaluate creativity and what they feel about evaluating AI-generated art. All these research studies provide insights into the emotional effect of AI-generated art per se and on third-party areas such as man-machine interaction, personal emotional health and AI technology implementation.

## CONCLUSION

The implications of the assessment of artistic and aesthetic value in AI-generated artworks refer to how art is understood and evaluated. Aesthetic response to computer generated art: A weak link to art history background (Chamberlain et al., 2018) refers to how art educational background provides a reliable source of variance in aesthetic judgments” This part of the references helps to understand the variety of factors that determine the assessment for aesthetic value. For instance, the difference in art educational backgrounds appears to provide a reliable source of variance in aesthetic value judgment. On the other hand, the influence of national identity on art evaluation was explored in the art appreciation heuristic (Mastandrea et al., 2021), adding national identity as a new one. This study captures the different heuristics of appreciation to use in the assessment of artistic value (Gangadharbatla, 2022). Based on the results of interaction influences between the factors addressed, the findings reflect on the implications of AI-generated art in the industry.

## REFERENCES

- Bagozzi, Richard P., Michael K. Brady, and Ming-Hui Huang. (2022). AI service and emotion. *Journal of Service Research* Volume 25, No. 4.
- Cetinic, Eva, and James She. (2022). Understanding and creating art with AI: Review and outlook, *ACM Transactions on Multimedia Computing, Communications, and Applications*, Volume 18, No. 2.
- Chamberlain, Rebecca, Mullin, Caitlin, Scheerlinck, Bram Wagemans, Johan (2018). Putting the art in artificial: Aesthetic responses to computer-generated art, *Psychology of Aesthetics, Creativity, and the Arts*, Volume 12, No. 2.

- Chatterjee, Anjan. (2022). Art in an age of artificial intelligence, *Frontiers in Psychology*, Volume 13, 1024449.
- Chen, J. M., Chang, P. C., & Liang, K. W. (2019, December). Speech emotion recognition based on joint self-assessment manikins and emotion labels, In 2019 IEEE International Symposium on Multimedia (ISM). IEEE, pp. 327–3273.
- Chen, Jing., Yuan, Dongfeng, Dong, Ruotong, Cai, Jingyi, Ai, Zhongzhu, & Zhou, Shanshan (2024). Artificial intelligence significantly facilitates development in the mental health of college students: A bibliometric analysis. *Frontiers in Psychology* Volume 15, 1375294.
- Cotter, K., Rodriguez-Boerwinkle, R., Silver, S., Hardy, M., Putney, H., Crone, D., & Pawelski, J. (2023). “Emotional Experiences and Flourishing During Art Museum Visits: A Latent Class Analysis”, in *Emotions And Flourishing In Art Museums*, pp. 1–27.
- Darda, Kohinoor, Marion Carre, and Emily Cross. (2023). Value attributed to text-based archives generated by artificial intelligence, *Royal Society Open Science* Volume 10, No. 2.
- Gangadharbatla, Harsha. (2022). The role of AI attribution knowledge in the evaluation of artwork, *Empirical Studies of the Arts* Volume 40, No. 2.
- Gu, Li, and Yong Li. (2022). Who made the paintings: Artists or artificial intelligence? The effects of identity on liking and purchase intention, *Frontiers in Psychology* Volume 13.
- Ho, Amic G. (2014). The New Relationship between Emotion and the Design Process for Designers. *Archives of Design Research* Volume 27, No. 2.
- Ho, Amic G. (2024a). Developing strategies for junior designers to manage the design and emotions. *New Design Idea* Volume 8, No. 1.
- Ho, Amic G. (2024b). “Exploration of the Emotion-Adaptive Engine in the Computer-aided Design Process.” In: Chan, S. W. (Ed). *Routledge Encyclopedia of Technology and the Humanities*. Routledge.
- Horton Jr, C. Blaine, Michael W. White, and Sheena S. Iyengar. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports* Volume 13, No. 1, 19001.
- Hosseini, Mohammad, David B. Resnik, and Kristi Holmes. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts, *Research Ethics* Volume 19, No. 4.
- Hua, Y., Niu, S., Cai, J., Chilton, L. B., Heuer, H., & Wohn, D. Y. (2024). “Generative AI in User-Generated Content”, In *CHI Conference on Human Factors in Computing Systems (CHI EA '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA.
- Huang, Ming-Hui, and Roland T. Rust. (2018). “Artificial intelligence in service”, *Journal of Service Research* Volume 21, No. 2.
- Hwang, A. H. C. (2022, April). “Too late to be creative? AI-empowered tools in creative processes”, In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1–9.
- Lima, Gabriel, Zhunis, Assem, Manovich, Lev, & Cha, Meeyoung (2021). On the social-relational moral standing of AI: An empirical study using AI-generated art, *Frontiers in Robotics and AI* Volume 8, 719944.
- Liu, Xiaojia, Chen, Xi, Huang, Yan, Wang, Weihong, Zhang, Mingkan, & Jin, Yang (2023). Landscape Aesthetic Value of Waterfront Green Space Based on Space–Psychology–Behavior Dimension: A Case Study along Qiantang River (Hangzhou Section), *International Journal of Environmental Research and Public Health* Volume 20, No. 4.

- Lucas Bellaiche, Rohin Shahi, Martin Harry Turpin, Anya Ragnhildstveit, Shawn Sprockett, Nathaniel Barr, Alexander Christensen & Paul Seli (2023). Humans versus AI: whether and why we prefer human-created compared to AI-created artwork, *Cognitive Research: Principles and Implications*, Volume 8, No. 1.
- L'Yi, Sehi, & Gehlenborg, Nils (2022). Multi-view design patterns and responsive visualization for genomics data, *IEEE Transactions on Visualization and Computer Graphics* Volume 29, No. 1.
- Lyu, Yanru, Wang, Xinxin, Lin, Rungtai, & Wu, Jun. (2022). Communication in human–AI co-creation: Perceptual analysis of paintings generated by text-to-image system, *Applied Sciences* Volume 12, No. 22.
- Machajdik, J., & Hanbury, A. (2010, October). “Affective image classification using features inspired by psychology and art theory”, In *Proceedings of the 18th ACM international conference on Multimedia* (pp. 83–92).
- Mastandrea, S., Wagoner, J. A., & Hogg, M. A. (2021). “Liking for abstract and representational art: National identity as an art appreciation heuristic”, *Psychology of Aesthetics, Creativity, and the Arts* Volume 15, No. 2.
- Menninghaus, Winfried, Wagner, Valentin, Wassiliwizky, Eugen, Schindler, Ines, Hanich, Julian, Jacobsen, Thomas, & Koelsch, Stefan (2019). What are aesthetic emotions?, *Psychological Review* Volume 126, No. 2.
- Miu, Andrei C., Simina Pițur, & Aurora Szentágotai-Tătar. (2016). “Aesthetic emotions across arts: A comparison between painting and music”, *Frontiers in Psychology* Volume 6.
- Pelowski, Matthew, Specker, Eva, Gerger, Gernot, Leder, Helmut, & Weingarden, Lauren S. (2020). Do you feel like I do? A study of spontaneous and deliberate emotion sharing and understanding between artists and perceivers of installation art, *Psychology of Aesthetics, Creativity, and the Arts* Volume 14, No. 3.
- Rodgers, Waymong, Murray, James M., Stefanidis, Abraham, Degbey, William Y., & Tarba, Shlomo Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes, *Human Resource Management Review* Volume 33, No. 1.
- Schindler, Manuel, Maihöfner, Christian, Bolwerk, Anne, & Lang, Frieder R. (2017). Does participation in art classes influence performance on two different cognitive tasks?, *Aging & Mental Health* Volume 21, No. 4.
- Silvia, Paul J. (2012). Human emotions and aesthetic experience, in: *Aesthetic Science: Connecting Minds, Brain And Experience*, Arthur P. Shimamura, Stephen E. Palmer, (Ed.). pp. 250–275.
- Suhaimi, Nazmi Sofian, James Mountstephens, and Jason Teo. (2020). EEG-based emotion recognition: A state-of-the-art review of current trends and opportunities, *Computational Intelligence and Neuroscience*, 2020.
- Ullrich, Martin, and Sebastian Trump. (2023). Sonic Collaborations between Humans, Non-human Animals and Artificial Intelligences: Contemporary and Future Aesthetics in more-than-human worlds *Organised Sound* Volume 28, No. 1.
- Yusa, I., Yu Yu, and Tetiana Sovhyra. (2022). Reflections on the Use of Artificial Intelligence: Works Of Art. *Journal of Aesthetics, Design, and Art Management* Volume 2, No. 2.
- Zhang, Y., Tang, F., Dong, W., Le, T. N. H., Xu, C., & Lee, T. Y. (2023). Portrait Map Art Generation by Asymmetric Image-to-Image translation *Leonardo* Volume 56, No. 1.

# BlindSpot: An AI-Powered Intelligent Mirror Assisting With Facial Hygiene Analysis for Blind People

Toby Lu

Mulgrave School, West Vancouver, British Columbia, Canada

## ABSTRACT

Appearances are important for our everyday lives: it is how we present ourselves. Yet, for people who are blind or visually impaired, it can be challenging to notice or take care of their facial hygiene without sighted assistance. Furthermore, existing AI-powered visual assistive technologies such as Seeing AI do not provide such support. We present BlindSpot, an AI-powered application designed as an intelligent mirror for assisting blind or visually impaired individuals with facial hygiene analysis. By using a reference photo that is uploaded by the user of their typical appearance, the app will compare it with present photos in order to notice and notify the users of any irregularities in their appearance. To understand the preliminary effectiveness of BlindSpot, we conducted a technical evaluation using online and AI-generated photos, with varied lighting, focus, image composition, and skin color, as well as diverse types of facial irregularities such as faulty makeup, facial and oral stains and color changes. Our results show that BlindSpot is able to detect irregularities in various situations. This work represents an initial exploration of AI-powered assistive application for a novel domain of facial hygiene analysis for blind or visually impaired people.

**Keywords:** Accessibility, Assistive technology, Blind, LLM



## INTRODUCTION

Appearances are important for our everyday lives: it is how we present ourselves. Yet, for people who are blind or visually impaired, it can be challenging to notice or take care of their facial hygiene without sighted assistance. Furthermore, existing AI-powered visual assistive technologies such as Seeing AI do not provide such support. We first performed an online search to identify the needs of blind people and people who are visually impaired by checking their appearances. We found that checking appearances is a pertinent need shared by many people who are blind or visually impaired. For example, in the autoethnography by a blind girl (Holly, 2019), the author explains that growing up with visual impairments is physically and mentally strenuous on teenagers. In Li et al. (2022), the authors find that people who are visually impaired “normally keep [their] appearance the same as before the loss of vision. This makes people focus less on [their] face and [their] visual impairments.” And that “people with visual impairments are attentive to their appearance,” by “applying makeup”, so that they “can control [their]



appearance again.” Pradhan and Daniels (2021) found that people with visual impairments care about their appearance as “they know people around them can see them,” they also “rely on sighted or partially sighted people to act as their mirrors.”

**Table 1.** An overview of how BlindSpot works. The user may choose to upload a reference photo as the baseline (left photo). When the user has a need to check their appearances, they can upload a user taken photo (right photo). BlindSpot will give them feedback as shown.

User reference photo	User taken photo	AI feedback
		The photo you've provided appears to be a stylized and polished look, typically seen in makeup artistry for special events or professional photography. Everything seems to be intentionally crafted, and there doesn't appear to be anything that needs cleaning up in terms of makeup application or hair styling from what I can discern. Each element from the hair to the makeup is designed to create a glamorous and cohesive appearance.

[1] <https://www.businessinsider.com/bridal-wedding-makeup-before-and-after-2019-3>

Such prior work highlights the need to develop novel solutions and mechanisms to help blind people check on their appearances more independently. The advances in artificial intelligence, especially with regard to AI's capabilities on analyzing images, brings tremendous potential to augment the vision of people who are blind or visually impaired. In this project, I propose the design of BlindSpot, which is an AI-powered intelligent mirror to help people who are blind and visually impaired check their appearances through uploading images to the platform.

Our prior need-finding search revealed diverse needs by blind people on checking appearances, including checking food stains around mouth area, checking makeup quality, and checking cleanliness of their faces. To use BlindSpot, we imagine users can first upload a reference photo which displays their regular appearance, and then upload another photo that they wanted to have the intelligent mirror check for them. In cases where users do not have a reference photo, they can directly upload a photo of interest and ask BlindSpot to check their appearance for them.

To evaluate the design concept of BlindSpot, we first constructed a repository composed of people's headshot images with different appearance issues ranging from food stuck on teeth, to smudge of makeup. We then prompted GPT-4 to evaluate the headshot images and give users feedback for them to adjust their appearance. We found that in most of the scenarios, GPT-4 is able to generate accurate assessment and give helpful feedback. We report scenarios where GPT-4 makes mistakes and makes suggestions on how to further improve the design.

## RELATED WORK

In this section, we will synthesize prior work that has developed assistive technologies for blind or visually impaired people (BVI). We will first introduce the challenges BVI people experience in their daily routines as shown by prior work. Prior research papers have described the challenges people that are visually impaired experienced in daily lives. For example, Huber et al. (2022) did a multi-year study to explore how visually impaired people can be better supported in their indoor navigation. Moreover, Yairi et al. (2011) proposed a novel graphic representation method with touch and sound as the universal designed touch-screen interface for visually impaired people to understand graphical information. In another study Yairi et al. (2009) developed a method to help visually impaired people intuitively recognize maps using audio and touch panels. This method is universal-designed to help both visually impaired and non-impaired people to explore maps together.

Prior work has also described BVI people's challenges with regard to independent living and everyday activities. For example, prior study by Li et al. (2022) identified a variety of daily living tasks people with upper-body motor impairments encounter. Moreover, Fusseneger and Spiel (2022) used an auto-ethnographic approach to understand the function and relevance of assistive technologies in everyday life. They found that assistive technologies may pose a risk to users, especially when there is a technology failure.

## NEED FINDING

To understand BVI people's needs with regard to checking appearances in their daily routines, we first performed a need finding search through online resources to identify the circumstances BVI people would like to check their appearances. We initiated the search through prompting ChatGPT specifically for links to blog posts written by BVI people. Here is a summary of findings.

Some blogs posted by people with visual impairment described their personal experience as "emotional and physically strenuous" Holly (2019). Some other articles have shared quotes from blind people describing their experiences Mohapatra (2018), for example, "I normally keep my appearance the same as before the loss of vision. This makes people focus less on my face and my visual impairments." "Nonetheless, people with visual impairments are attentive to their appearance in the same proportion as sighted peers, especially in spaces that are guided by social norms around makeup." "When I first lost my eyesight, I was quite sad that I couldn't look in the mirror. Applying makeup is a way that I can control my appearance again."

Specific to checking appearances, prior work has shown that "Blind individuals pay a lot of attention to their appearance because they know people around them can see them" Pradhan and Daniels (2021). It's also disclosed in this study that "blind people rely on sighted or partially sighted people to act as their mirrors."

It is shown from prior blog posts and studies that blind people do care about their appearances, e.g., "She obviously cared about her physical

appearance, as I noticed from her expertly applied make-up and glorious hair” (Quora, 2018). “Blind people are concerned with things about as much as the sighted people around them are, such as their readers, their companions, their coworkers, their friends and families” (Quora, 2018). Posts from Reddit also show that blind people care about their appearances, e.g., “Yes, because they want to fit in like anyone else.” “They still want to be presentable and put-together” (Reddit, 2019).

A YouTube (YouTube, n.d.) video shows that a blind girl enjoys doing makeups for herself and become youtubers guiding other BVI people to do makeups, e.g., “HEY! I’m Molly, a typical sushi and makeup loving millennial girl who just so happens to be blind! I was diagnosed with Retinitis Pigmentosa at just 4 years old and began public speaking at age 5. Even though I can’t see, I know that there are bright spots in everything we face. Let’s find them together!”

As a summary, blind people care about their appearance. Their concerns may be based on factors like personal pride, comfort, cultural influences, and social acceptance. Blind individuals often use their other senses, such as touch and sound, to assess and enhance their appearance. Like anyone else, attitudes toward appearance vary among individuals (HousingLin, n.d.; Tindall, 2019).

## **USER STORY**

Jack will be attending an office party and he wants to check his appearance before going to the party. Since he had just eaten dinner and woken up from a nap, he wanted to check whether there was food stuck on his teeth or whether his hair was presentable. He opened BlindSpot and took a selfie of his whole face and another selfie of him smiling, showing his teeth. Using the voiceover from his phone, he asked BlindSpot to check his teeth and appearance. BlindSpot gave him feedback that the selfie was incomplete, and the right side of his face is cut off in the picture. Jack then retook a photo and asked BlindSpot to check appearance for him. This time, BlindSpot told him that there’s spinach on his right lateral incisor, so Jack cleaned it up. Jack then felt more confident to attend the party.

## **EXPERIMENTS**




I performed preliminary experiments to demonstrate the feasibility of the approach.

First of all, I compared providing a reference photo versus not providing a reference photo for ChatGPT to give feedback based on an image. I found that providing a reference photo made the AI give much more comprehensive feedback to the user, as shown in Table 1.

Next, I’ll present some preliminary results on the mistakes AI may make in this scenario. In my experiments, I found that ChatGPT did a reasonable job detecting anomalies on the person’s face. For example, in one photo where the person has smudged makeup, ChatGPT was able to successfully determine the smudge of lipstick on the person’s lower lip on their right side. However, they are unable to determine the smudged eyeliner, this may be due to the

ambiguity of the aesthetic the AI thinks the person is going for. Moreover, when the person has noticeable dirt on their face, ChatGPT is able to capture them. The results are shown in Table 2.

**Table 2.** Some example outputs from ChatGPT given a user-provided photo. The ChatGPT generated feedback is shown in the middle column, and some of my interpretation is shown in the right column.

Photo	ChatGPT Feedback	Notes
 [1]	In the photo you've provided, you appear to have a smudge of lipstick on your lower lip extending slightly outside the natural lip line on your right side, it's a minor detail, but touching up that area could create a cleaner look. Other than that, your appearance seems quite purposeful and expressive, suggesting a potential creative or artistic context for the photo.	ChatGPT was able to successfully determine the smudge of lipstick on the person's lower lip on their right side. However, it does not detect the smudged eyeliner, this may be due to the ambiguity of the aesthetic the AI thinks the person is going for; where the smudged lipstick is definite.
 [2]	In the photo you provided, there are noticeable smudges or spots of dirt scattered across your face, primarily on your cheeks, nose, and forehead. Here are the details regarding your appearance based on your left and right sides ....	ChatGPT was able to successfully determine smudges or spots of dirt scattered across the person's face, as well as the different sides the dirt was on. It is important to note that the dirt spots are very noticeable.
 [3]	Upon reviewing the image you've provided, it appears that there is a small, dark green piece of what could be food or another material on the tooth that is second from the left on the top row of the teeth (from the viewer's perspective). This is on your right side. It's subtle, but if this were for a professional setting or a photo you intend to share, you might want to clean it off.	ChatGPT at first, was unable to detect anything odd with the person's teeth, however, when asked specifically about the teeth area, they were able to notice the piece of green material on the person's teeth. ChatGPT was also able to correctly identify the area where the material was.

[1]<https://opmakeup.com/microblading-training/risks-involved-in-microblading/>

[2]<https://www.gettyimages.ca/detail/photo/woman-face-covered-with-dirt-and-leaf-at-home-royalty-free-image/1291939946?adppopup=true>

[3]<https://www.gettyimages.dk/detail/photo/businessman-with-food-in-his-teeth-royalty-free-image/57421505?adppopup=true>

I also noticed that when giving specific prompts, ChatGPT is more likely to detect things that are aligned with the users' expectations. For example, when using a generic prompt "Do you notice anything in this photo of me? Is there anything wrong and/or noticeable with my appearance? Is there anything I

need to clean up? Also make sure you tell me these things regarding MY lefts and rights.”, ChatGPT sometimes can’t detect food stuck on someone’s teeth, however, when the prompt is specifically designed to ask about teeth, e.g., “is there anything on my teeth?”, ChatGPT can successfully detect. More output examples are shown in Table 3.

**Table 3.** Some example outputs from ChatGPT given a user-provided photo. The second row includes the specific prompts we used that prompted the model to pay attention to a specific area, and the third row shows the feedback generated by AI.



[1]

Specific prompt: Do you notice anything in this photo of me? Pay close attention to the area around my mouth or teeth. Is there anything I need to clean up? Also make sure you tell me these things in regard to MY lefts and rights.

Feedback: It seems like there’s a bit of lipstick on your teeth on your right side, which is quite a common little mishap when applying makeup. A quick check in the mirror should help you clean it up. Everything else looks great! Your smile is lovely and vibrant.



[2]

Specific prompt: Do you notice anything in this photo of me? Pay close attention to my makeup and its noticeability. Is there anything I need to clean up? Also make sure you tell me these things in regard to MY lefts and rights.

Feedback: Your makeup looks striking and meticulously applied. The eye makeup is particularly dramatic with a bold winged eyeliner that accentuates your eyes. Your eyeshadow is blended smoothly, transitioning from a lighter shade on the inner corners to a darker tone towards the outer edges...



[3]

Specific prompt: Do you notice anything in this photo of me? Pay close attention to the photo’s lighting and composition. Is there anything I need to clean up? Also make sure you tell me these things in regard to MY lefts and rights.

Feedback: In the photo you’ve provided, the backlighting is quite strong, causing the foreground, including yourself, to be in shadow. This is known as being backlit, where the light source is directly facing the camera, making the details of the subject harder to discern. To enhance the photo, you might want to adjust the lighting so that your face and body are more illuminated...

[1]<https://www.gettyimages.ae/detail/photo/smiling-woman-with-food-particle-stuck-between-her-royalty-free-image/88178261>

[2]<https://www.businessinsider.com/bridal-wedding-makeup-before-and-after-2019-3>

[3]<https://rosesandscarsphotography.com/lighting-your-selfie-101-indoor-lighting-overcast-day/>

With a series of prompt engineering, I identified that when the prompt is specific to an area on the person's face, the model is more likely to make detections correctly. For example, "Do you notice anything in this photo of me? Pay close attention to the area around my mouth or teeth. Is there anything I need to clean up? Also make sure you tell me these things in regard to MY lefts and rights.", or "Do you notice anything in this photo of me? Pay close attention to my makeup and its noticeability. Is there anything I need to clean up? Also make sure you tell me these things in regard to MY lefts and rights".

## DISCUSSION

Although the AI is able to display a wide range of features and abilities, there are still some limitations. BlindSpot tends to be unable to detect food stuck on teeth and scars. Firstly, it is unable to consistently detect whether there is food on a user's teeth. This inability needs to be fixed as it is crucial for all the features to be consistently working and able to detect. Another limitation would be the AI's inability to detect scars. This is important as sometimes, when presented with a picture of someone with scars on their face, the AI would detect it as makeup, or a birthmark, giving false feedback which can limit a user's experience.

There are also some additional features that can be added to BlindSpot in order to improve its design and user support. One feature could be uploading a video instead of a picture, as in videos, there are quite a few more details that can be missed when using only one picture. This can also improve the ease of use of the users, as they are not required to take "perfectly" framed photos, instead they can just take a video of their whole face. Another feature that can improve the ease of use could be real time feedback, where while taking pictures, the app can notify the users whether they are out of frame or in good composition.

## REFERENCES

- Fukushima, Y., Uematsu, H., Mitsuhashi, R., Suzuki, H., & Yairi, I. E. (2011, October). Sensing human movement of mobility and visually impaired people. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility* (pp. 279–280).
- Fussenegger, F., & Spiel, K. (2022, October). Depending on Independence An Autoethnographic Account of Daily Use of Assistive Technologies. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–6).
- Holly (2019). A message for young people with a vision impairment. Retrieved from: <https://lifeofablindgirl.com/2019/03/06/a-message-for-young-people-with-a-vision-impairment/>.
- HousingLin. Personal Stories. Retrieved from: <https://www.housinglin.org.uk/Topics/browse/sight-loss-home-the-built-environment/personal-stories/>.
- Huber, S., Alieva, A., & Lutz, A. (2022, October). Vibrotactile Navigation for Visually Impaired People. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 1–4).

- Kim Tindall. (2019, April). The Challenges in Going Blind and Learning to Live Again. Retrieved from: <https://nfb.org/sites/default/files/images/nfb/publications/bm/bm14/bm1404/bm140414.htm>.
- Li, F. M., Liu, M. X., Zhang, Y., & Carrington, P. (2022, October). Freedom to choose: Understanding input modality preferences of people with upper-body motor impairments for activities of daily living. In Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility (pp. 1–16).
- Li, F. M., Spektor, F., Xia, M., Huh, M., Cederberg, P., Gong, Y.,... & Carrington, P. (2022, April). “It Feels Like Taking a Gamble”: Exploring Perceptions, Practices, and Challenges of Using Makeup and Cosmetics for People with Visual Impairments. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1–15).
- Nilisha Mohapatra. (2015, March). 13 Things You May Not Know About Blind People. Retrieved from: <https://www.lifehack.org/articles/communication/13-things-may-know-blind-people.html>.
- Pradhan, A., & Daniels, G. (2021). Inclusive beauty: How buying and using cosmetics can be made more accessible for the visually impaired (VI) and blind consumer. *Cosmetics and Toiletries*, 136(4), DM4-DM15.
- Quora. (2018). How much do blind people care about their physical appearance or fashion? Retrieved from: <https://www.quora.com/How-much-do-blind-people-care-about-their-physical-appearance-or-fashion>.
- Reddit. (2019). Do blind people care about looking attractive? Retrieved from: [https://www.reddit.com/r/TooAfraidToAsk/comments/9pd5te/do\\_blind\\_people\\_care\\_about\\_looking\\_attractive/](https://www.reddit.com/r/TooAfraidToAsk/comments/9pd5te/do_blind_people_care_about_looking_attractive/).
- Yairi, I. E., Azuma, Y., & Takano, M. (2009, October). The one octave scale interface for graphical representation for visually impaired people. In Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility (pp. 255–256).
- Yairi, I. E., Naoe, K., Iwasawa, Y., & Fukushima, Y. (2011, October). Do multi-touch screens help visually impaired people to recognize graphics. In The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility (pp. 237–238).
- YouTube. Molly Burke. Retrieved from: <https://www.youtube.com/@MollyBurkeOfficial>.

# Design Process for Augmented Reality (AR) Experiences From the Perspectives of UX and Game Designers

Jingoog Kim and Lina Lee

Loyola University New Orleans, New Orleans, LA 70118, USA

## ABSTRACT

This paper presents a cognitive study comparing the approaches of User Experience (UX) designers and game designers in Augmented Reality (AR) experience design. The AR Haunted College is a scenario-based sequential AR experience that transforms the Monroe Hall building at Loyola University New Orleans into a chilling Halloween spectacle. The design of AR Haunted College involved contributions from both UX designers and game designers. We characterize the distinct approaches of UX and game designers by analysing the design process of the AR Haunted College developed in an agile project management framework. Our results show common and distinct HCI and game design principles between UX and game designers' perspectives. Building on the findings of the analysis, we propose a set of design principles as a basis for an AR design framework and associated heuristics.

**Keywords:** Augmented reality, AR design principle, AR experience design, Design framework, Design heuristics

## INTRODUCTION

Augmented Reality (AR) technology has experienced rapid growth and integration across various industries, providing innovative and immersive experiences for users. AR applications extend from gaming and entertainment to real-world problem-solving, offering new dimensions for engagement and interaction. Within the realm of AR development, professionals with diverse backgrounds collaborate to create impactful experiences, including game designers and User Experience (UX) designers. Game designers bring their expertise in crafting captivating narratives, interactive mechanics, and engaging gameplay elements to the table. Their focus lies in creating experiences that captivate and entertain users, often with a strong emphasis on storytelling and user engagement. On the other hand, interaction designers are dedicated to ensuring that AR experiences are user-friendly, intuitive, and deliver on the needs and expectations of the end-users. They are concerned with aspects such as accessibility, usability, and overall user satisfaction. This research seeks to uncover and analyze the distinct approaches of game designers and UX designers in developing AR experiences across various domains, fostering innovation and improvement in the field of Augmented Reality.



We present an AR design project, AR Haunted College, that provides a set of sequential AR experiences in a scenario-based Halloween theme. The design of AR Haunted College involved contributions from both UX designers and game designers. We analyze the design process of the AR Haunted College in each perspective of UX designers and game designers to identify their unique design approaches.

This research addresses the following research question:

- *What are the distinct approaches between UX design and game design that influence the design and user interactions in designing augmented reality (AR) experiences?*

This research addresses the lack of studies on design principles for the next generation of AR experiences by proposing a new approach that investigates two different disciplines related to AR experience design.

## **DESIGN PRINCIPLES FOR AUGMENTED REALITY**

While there is some research on augmented reality design guidelines, most of it is narrow, specific to particular problems, and does not provide generalized principles for AR experiences (Dunleavy, 2014; Dünser et al., 2007; Krüger and Bodemer, 2022; Liang, 2018, 2016).

Dünser et al. (2007) investigated how to apply HCI design principles in an AR design. They introduced eight design principles and heuristics with some examples: affordance, reducing cognitive overhead, low physical effort, learnability, user satisfaction, flexibility in use, responsiveness and feedback, and error tolerance. Dunleavy (2014) introduced design principles for augmented reality learning: Enable and then challenge, drive by gamified story, see the unseen. These design principles emphasize the use of affordance and reducing cognitive overhead in AR learning settings. Kourouthanassis et al. (2015) proposed five design principles for the development of mobile augmented reality applications.: Use the context for providing content, deliver relevant-to-the-task content, Inform about content privacy, Provide feedback about the infrastructure's behavior, Support procedural and semantic memory. Liang (2016; 2018) proposed six AR design principles for older adults (Diminished Augmentation, Modality-rich Augmentation, Instantaneous Augmentation, Accurate Augmentation, Augmented Augmentation and Transparent Augmentation) based on five key features of AR design (i.e., Changeability, Synchronicity, Partial one to one, Hidden Reality and Registration). Krüger and Bodemer (2022) applied two basic multimedia design principles in AR learning environments: the spatial contiguity principle with visual learning material, and the coherence principle with audiovisual learning material, leveraging AR-specific contextuality potentials.

The existing research on AR design principles mentioned above relies heavily on human-computer interaction (HCI) design principles and usability principles (Dunleavy, 2014; Dünser et al., 2007; Liang, 2016; 2018), or focuses on the technical characteristics of AR technology (Krüger and Bodemer, 2022; Schmalstieg and Hollerer, 2016), rather than on the holistic

user experience of AR design. As recent AR experiences have become much more complex, with multiple purposes and multimodal interaction, existing HCI design principles are limited in their ability to lead to successful AR experiences.

Recent research on AR design (Alqahtani and Kavakli-Thorne, 2020; Li et al., 2017; Yu et al., 2022; Zuo et al., 2022) has incorporated game design aspects to support educational purposes, demonstrating the potential of game-based learning with AR to facilitate learning in various contexts. The CybAR Game (Alqahtani and Kavakli-Thorne, 2020) is an AR mobile application that teaches cybersecurity concepts and demonstrates the consequences of actual cybersecurity attacks through feedback. To develop CybAR Game, the researchers conducted a focus group to identify key components that needed to be addressed in the design. The key components are roles, goals and objectives, outcome and feedback, rewards, and interaction. These key components include some parts of both HCI design principles and game design principles. Li et al. (2017) reviewed previous studies of AR games for learning. They proposed five considerations for designing AR learning games from their study: learner groups, learning objectives, AR features, game mechanics, and social interactions. Zuo et al. (2022) studied situating learning in AR fantasy and presented design considerations for AR game-based learning. They emphasized the use of narratives and simple interactions to increase learning effects as design considerations.

The existing research on AR game design for learning experiences mentioned above has primarily focused on applying specific game elements to achieve learning objectives, rather than implementing comprehensive game design principles. In this study, we investigate the approach of game designers towards AR design, exploring how game design principles can be integrated with HCI design principles for augmented reality experiences.

## **AR HAUNTED COLLEGE**

To investigate different perspectives on interaction design and game design for AR experiences, we developed a series of AR experiences involving contributions from both UX designers and game designers. The AR Haunted College is a scenario-based sequential AR experience that transforms the Monroe Hall building at Loyola University New Orleans into a chilling Halloween spectacle. Comprising 18 AR experiences that begin at the entrance of the building and culminate on the 5th floor, the AR Haunted College offers users a cohesive narrative to engage with. Some experiences involve jump scares from animated objects, while others require user actions like avoiding zombies or entering specific rooms to complete missions. The AR Haunted College was created using Adobe Aero, and its AR designs include 2D and 3D objects, videos, layered 2D elements, 3D-scanned objects, a blend of video and animated 2D elements, and various sound effects. Interactions within the AR Haunted College involve a dynamic mix of tap interactions and proximity interactions.

Figure 1 shows some of the AR experiences in AR Haunted College. Figure 1a is the first user story. Users face the AR entrance to the haunted house at the entrance of the Monroe Hall building. Ghosts come and go at this entrance, and when the user opens the door, they receive a welcome message and are guided to the next location. Figure 1b shows a user story on the 3rd floor. As the user runs away from a group of small skeletons, they happen to discover a ticket office shrouded in darkness. The skeleton standing in front of the ticket booth looks at the user and asks if they have been invited to Halloween Town or if they have come to conduct official business with Dr. Kim. If the user replies that they have been invited to Halloween Town, the skeleton gives them a ticket. On the other hand, if the user declares that they want to meet Dr. Kim, the skeleton immediately provides Dr. Kim's room number. Figure 1c shows a user story on the 4th floor. Upon arriving at the exhibition hall, the user finds four students, 3d scanning models, transformed into plaster statues. Approaching Lila cautiously, she explains they are cursed and pleads for help. Then, the user approaches Shay's statue, who reveals her sketchbook is cursed and requests its destruction. Finally, approaching Jada's statue, who informs them about a sketchbook on the 5th floor that can lift the curse. The user is instructed to seek clues in the posters on the wall to break the curse. Figure 1d shows a user story on the 5th floor. The user opens locker number 117. Inside, among the thick dust, lies an old sketchbook. As the user picks it up, it comes to life, blinking and writhing as if trying to escape the user's grasp. The user quickly flicks a lighter and the sketchbook bursts into flames, disappearing in a puff of ash.



**Figure 1:** AR Experiences within the AR haunted college.

## RESEARCH METHOD

The design of AR Haunted College involved contributions from both UX designers and game designers. Five interactive design major students enrolled in the Agile Development Process course developed an AR design project using an agile project management framework, an iterative approach to managing software development projects that focuses on continuous releases and incorporates customer feedback with every iteration. Students completed a sprint, a fixed period of two weeks, during which they designed and completed a certain portion of the project. The design project included 4 sprints, and each sprint included a daily meeting called the daily scrum, design work, development work, design testing, a sprint review meeting, and a sprint retrospective meeting.

Additionally, six students enrolled in the Game Design course actively engaged in AR design projects, participating in 4 focus group discussions, one focus group after each sprint, to identify issues and potential improvements from a game design perspective. The evaluation process comprised two sessions: testing and focus groups. The protocol, including the informed consent document, has been reviewed and approved by our IRB, and we obtained informed consent from all students to conduct the study.

To examine the UX designers' perspective on designing the AR experience, we recorded all daily scrum meetings (16 meetings), sprint review meetings (4 meetings), and sprint retrospective meetings (4 meetings). The verbal data reflects how UX designers approach design, highlighting the design aspects they focus on during the design process and self-evaluation, outlining encountered design challenges, and detailing interactions and user experiences they aim to improve and strengthen. For investigating the viewpoint of game designers in shaping the AR experience, we recorded all focus group sessions (4 sessions). The verbal data illustrates how game designers approach the design of AR experiences, the design aspects they focus on in AR design, the game design principles they apply to AR design, and how they use game elements in non-game AR design. All verbal data collected from both the Agile Development Process course and the Game Design course was transcribed for analysis.

## THEMATIC ANALYSIS OF COLLECTED DATA

We characterize the distinct approaches of UX and game designers by identifying emerging themes from the verbal data. Through our analysis, we identified various themes related to design principles, the design process, project management, technical consideration, team collaboration, and documentation. After completing the initial round of coding and grouping with the identified codes, we extracted only the coded data relevant to both HCI design principles (Masmuzidin and Aziz, 2019; Nielsen, 2005, 1994) and game design principles (Brathwaite and Schreiber, 2009), and then proceeded to interpret the data. We categorized the emerging themes into perspectives from UX designers and game designers, and extracted relevant design principles based on these themes.

## UX Designers' Perspective

Five themes emerged from the verbal data of the Agile Development Process course, from the perspective of UX designers:

- **Graphics and Visual Experience:** Students frequently discussed the quality of graphics, noting specific improvements in visual effects and overall immersion. Additionally, suggestions were made to further enhance visual elements by incorporating dynamic effects such as smoke, blood, and moving objects. However, concerns were raised regarding text readability, prompting discussions on the necessity for better contrast and background colors to improve clarity. Students primarily focused on the quality of graphics, dynamic graphic effects, and visual information readability, which are related to HCI design principles of *Aesthetic and Minimalist Design*.
- **Interactivity and Engagement:** Discussions regarding interactive elements and user experience underscored the importance of incorporating engaging elements to prolong user engagement. Suggestions included adding game-like features such as puzzles, challenges, and interactive characters to enhance the overall experience. Furthermore, students emphasized the need for intentional and thematic connections between different elements to maintain coherence throughout the user journey. This relates to HCI design principles of *Engagement*, *Consistency*, and *Informative Feedback*, emphasizing the relationship between interaction and the game design principle of *Theme* (i.e., a background story that provides context for what the players need to accomplish).
- **Audio and Sound Effects:** The analysis revealed a recurring theme surrounding the role of audio and sound effects in shaping the user experience. Students highlighted the significance of sound effects and background music in setting the mood and enhancing immersion. Recommendations focused on improving audio cues, including timing adjustments and ensuring alignment with narrative and visual elements to foster a cohesive experience. The relevant HCI design principles are *Affordance*, *Signifier*, and *Visibility of System Status*. This is also related to the game design principles of *Theme*.
- **Spatial Design and Exploration:** Spatial design has emerged as a crucial aspect of the user experience, particularly in the context of augmented reality (AR) experiences. Students acknowledged the importance of spatial design in creating immersive experiences, suggesting strategies to optimize physical space for exploration and improve the placement of AR elements. Discussions also centered on reducing clutter and optimizing cues to prevent overwhelming users during interaction. Since the AR interface type utilizes the Exploring type of interaction, maximizing physical and digital exploration is crucial in AR user experience. The relevant HCI design principles are *Affordance*, *Signifier*, and *Exploration of Virtual World*. This is also related to the game design principles of *Game Space* (i.e., the entire area of the game).
- **Narrative and Storytelling:** Lastly, students emphasized the pivotal role of narrative and storytelling in driving user engagement and maintaining

interest. Discussions included refining narrative structures, integrating riddles and puzzles, and incorporating physical artifacts to complement digital narratives and encourage deeper exploration. The relevant HCI design principle is *Engagement*. This is also related to the game design principles of *Game Dynamics* (i.e., patterns of play that arise from mechanics once they are set in motion by players) and *Theme*.

### Game Designers' Perspective

Six themes emerged from the verbal data of the Game Design course, from the perspective of Game designers:

- **Interactivity and Engagement:** Students expressed enthusiasm for the interactive elements of the AR experience, such as the storyline, quests, and treasure hunt-like clues. Suggestions were made to amplify interactivity by integrating more interactive elements and puzzles, which could deepen user engagement and immersion. This result closely aligns with the perspective of UX designers.
- **Graphics and Realism:** Feedback emphasized the importance of high-quality graphics and realism in virtual elements. Students discussed the realistic appearance of virtual characters, such as zombies, and their seamless integration into the physical environment. Suggestions were made to enhance graphics quality and improve the seamless integration of virtual and physical elements, thereby further enhancing the immersive nature of the experience. This result suggests that the emphasis on graphics quality aligns with the perspective of UX designers. However, game designers prioritize the relationship between graphics style and the physical environment or context, placing additional emphasis on consistency. The relevant HCI design principles in this context are *Aesthetic and Minimalist Design*, and *Consistency*.
- **Clarity of Instructions and Balance:** Mixed feedback was received regarding the clarity of instructions. While some students found the instructions clear and helpful, others felt confused or uncertain about what to do next in certain parts of the experience. Suggestions were made to strike a balance between providing clear instructions and allowing for exploration and discovery, ensuring users can navigate the experience smoothly while still feeling a sense of agency and autonomy. In HCI design principles, it is important to provide clear instructions, which aligns with the design principle of *Help and Documentation*. However, in game design, ambiguity of *Objectives* (i.e., goals or outcomes that players can work towards) and *Success Criteria* (i.e., ways of knowing when the objectives are met) can enhance the modeling of certain scenarios. Thus, achieving a balance in the clarity of instructions is considered crucial in this theme.
- **Difficulty, Mechanics, and Feedback:** Students provided insights on the difficulty level of the experience and proposed adjustments, such as offering clues or varying difficulty levels for tasks. Suggestions included refining mechanics like object scanning and providing more intuitive feedback for player actions. Emphasis was placed on the importance

of delivering meaningful feedback and rewards for player actions to enhance the overall experience and foster continued engagement. The relevant HCI design principle is *Engagement*, and the relevant game design principles are *Level Design* (i.e., crafting the layout, challenges, and player experience of individual stages), *Mechanics* (i.e., rules of the game), and *Rewards* (i.e., incentives that reward success).

- **Story, Characters, and Atmosphere:** Students discussed the presence of a captivating story and characters within the AR experience. Suggestions were made to further develop characters and incorporate additional storytelling elements to enhance immersion. Feedback stressed the importance of crafting a spooky atmosphere within the physical environment to complement the virtual elements, including recommendations for incorporating details like cobwebs and bats, which could heighten the overall sense of immersion and suspense. This is related to the game design principles of *Theme* and *Game Space*, including environment design, level design, and character design.
- **Game State and Progression:** Students recommended providing visual cues or indicators to help users track their progress and understand their current objectives. Feedback included suggestions for implementing an inventory system and item collection mechanics to enrich gameplay and progression, granting users a sense of achievement and advancement as they traverse through the experience. The relevant HCI design principle is *Affordance* and *Signifier*, and the relevant game design principles are *Game State* (i.e., a collection of all relevant virtual information that may change during play), *Mechanics* (e.g., *Progression of Play*), and *Rewards*.

## DISCUSSION

Through the analysis, we found common and distinct HCI and game design principles between UX and game designers' perspectives, as shown in Table 1. While UX designers primarily focused on HCI design principles such as Shneiderman's Eight Golden Rules (Masmuzidin and Aziz, 2019) and Nielsen's Heuristics (Masmuzidin and Aziz, 2019; Nielsen, 2005, 1994), they also considered some game design principles (Brathwaite and Schreiber, 2009). Conversely, game designers primarily centered on game design principles (Brathwaite and Schreiber, 2009) but also took into account certain HCI design principles (Masmuzidin and Aziz, 2019; Nielsen, 2005, 1994).

Based on the results of our analysis, we propose a set of design principles that encompass both HCI and game design principles. These principles serve as the foundation for an AR design framework and associated heuristics. The 9 HCI principles include Aesthetic and Minimalist Design, Engagement, Consistency, Informative Feedback, Affordance, Signifier, Visibility of System Status, Exploration of Virtual World, and Help and Documentation. Additionally, the 9 game principles consist of Theme, Game Space, Game Dynamics, Objectives, Success Criteria, Level Design, Mechanics, Rewards, and Game State.

**Table 1.** Comparison of design principles in UX and game designers perspective.

	HCI Design Principles	Game Design Principles
UX Designer	<ul style="list-style-type: none"> <li>• Aesthetic and Minimalist Design</li> <li>• Engagement</li> <li>• Consistency</li> <li>• Informative Feedback</li> <li>• Affordance</li> <li>• Signifier</li> <li>• Visibility of System Status</li> <li>• Exploration of Virtual World</li> </ul>	<ul style="list-style-type: none"> <li>• Theme</li> <li>• Game Space</li> <li>• Game Dynamics</li> </ul>
Game Designer	<ul style="list-style-type: none"> <li>• Aesthetic and Minimalist Design</li> <li>• Engagement</li> <li>• Consistency</li> <li>• Affordance</li> <li>• Signifier</li> <li>• Help and Documentation</li> </ul>	<ul style="list-style-type: none"> <li>• Theme</li> <li>• Game Space</li> <li>• Objectives</li> <li>• Success Criteria</li> <li>• Level Design</li> <li>• Mechanics</li> <li>• Rewards</li> <li>• Game State</li> </ul>

While our study was conducted based on a specific design project, the identified design principles still have limitations in terms of generalizability. Additionally, the game design principles need to be redefined and refined to apply to general AR design. For instance, the principle of Game State should be redefined as Experience State, encompassing both the Visibility of System Status and Experience State concepts. Another example is Game Space, which needs to be redefined as AR Experience Space, encompassing both the physical space and the spatial range of digital elements. However, they can serve as valuable design guidelines or evaluation heuristics that can be applied to various AR experience designs and evaluations. Moving forward, we plan to develop an AR design framework by adjusting and refining the design principles we identified to suit general AR design. Subsequently, we will apply this framework to AR design in various contexts, leveraging the insights gained from the results of this study.

## CONCLUSION

This paper describes a cognitive study that investigates the design process for Augmented Reality (AR) experiences from the perspectives of User Experience (UX) designers and game designers. To identify distinct design approaches of UX designers and game designers in designing AR experiences, we analyzed the design process of a scenario-based sequential AR design project, AR Haunted College, which involved contributions from both UX designers and game designers. Our analysis reveals both common and distinct HCI and game design principles between the perspectives of UX and game designers. As a result, we propose a set of design principles, consisting of 9 HCI design principles and 9 game design principles, to serve as the foundation



for an AR design framework. The main contributions of this paper are providing a foundation for developing a theoretical framework for designing AR experiences and uncovering the distinct approaches of game designers and UX designers in AR experience design.

## REFERENCES

- Alqahtani, H., Kavakli-Thorne, M., 2020. Design and evaluation of an augmented reality game for cybersecurity awareness (CybAR). *Information* 11, 121.
- Brathwaite, B., Schreiber, I., 2009. Challenges for game designers. Course Technology/Cengage Learning Boston, Massachusetts.
- Dunleavy, M., 2014. Design principles for augmented reality learning. *TechTrends* 58, 28–34.
- Dünser, A., Grasset, R., Seichter, H., Billinghamurst, M., 2007. Applying HCI principles to AR systems design.
- Kourouthanassis, P. E., Boletsis, C., Lekakos, G., 2015. Demystifying the design of mobile augmented reality applications. *Multimedia Tools and Applications* 74, 1045–1066.
- Krüger, J. M., Bodemer, D., 2022. Application and Investigation of Multimedia Design Principles in Augmented Reality Learning Environments. *Information* 2022, 13, 74.
- Li, J., Van der Spek, E. D., Feijs, L., Wang, F., Hu, J., 2017. Augmented reality games for learning: A literature review, in: *Distributed, Ambient and Pervasive Interactions: 5th International Conference, DAPI 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings* 5. Springer, pp. 612–626.
- Liang, S., 2018. Establishing design principles for augmented reality for older adults. Sheffield Hallam University (United Kingdom).
- Liang, S., 2016. Design principles of augmented reality focusing on the ageing population, in: *Proceedings of the 30th International BCS Human Computer Interaction Conference*. BCS Learning & Development.
- Masmuzidin, M. Z., Aziz, N. A., 2019. The adaptation of Shneiderman's golden rules and nielsen's heuristics on motivational augmented reality technology design for young children, in: *2019 IEEE 9th International Conference on System Engineering and Technology (ICSET)*. IEEE, pp. 62–67.
- Nielsen, J., 2005. Ten usability heuristics.
- Nielsen, J., 1994. Enhancing the explanatory power of usability heuristics, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Celebrating Interdependence - CHI '94*. Presented at the the SIGCHI conference, ACM Press, Boston, Massachusetts, United States, pp. 152–158. <https://doi.org/10.1145/191666.191729>
- Schmalstieg, D., Hollerer, T., 2016. *Augmented reality: principles and practice*. Addison-Wesley Professional.
- Zuo, T., Jiang, J., Spek, E. V. der, Birk, M., Hu, J., 2022. Situating learning in AR fantasy, design considerations for AR game-based learning for children. *Electronics* 11, 2331.

# Verification of an Alternative Wheelchair Control in a Virtual Environment

Jendrik Bulk and Benjamin Tannert

Hochschule Bremen – University of Applied Science, Bremen, 28199, Germany

## ABSTRACT

This work presents a training simulator that is being developed in collaboration with electric wheelchair users and medical professionals. The main function of the software is to serve as a development platform for alternative control systems that enable people with paralysis to control a wheelchair themselves. The project developed two prototype control systems using technologies such as eye tracking and a brain-computer interface.

**Keywords:** Wheelchair, VR, BCI, Eye tracking, Voice commands, Alternative controls, Virtual joystick

## INTRODUCTION

The focus of the research presented here is an alternative electric wheelchair control. The control is verified and improved in a virtual environment (VE) in close collaboration with the target group. People with a spinal cord injury in the neck area (locked-in syndrome) are unable to control a wheelchair in the traditional way. Simple motor tasks, such as operating a joystick, are hardly or not at all possible for them.

The first alternative control developed for the project consists of a combination of several sensors: brain-computer interface (BCI), eye tracking (ET), voice control (VC) and obstacle detection (OD). The second is an eye-controlled virtual joystick. Before a real wheelchair is constructed using one of these or another technique, the controls mentioned are implemented in software and iteratively developed to take the specific needs of the target group into account. This software is meant to be a simulation where the controls can be tested and developed and later, it serves as a training environment.

In order to obtain meaningful test results, the VE contains a realistic residential area. This represents accessible houses, streets and sidewalks. In this way, the control can be tested in various everyday situations. The software operator has the option to activate obstacles, pedestrians or cars in the VE.

After the software is completed, the test environment serves as a training simulator. It should be used before a wheelchair user learns how to control the real wheelchair. The development takes place on a participatory basis

with future users of the software. The focus is equally on wheelchair users and the software operators.

## CHOOSING A CONTROL TECHNOLOGY

In order to find a suitable control technology that takes into account the special requirements of the target group, various technologies are compared in this project and their usability is examined. The BCI procedures, eye-tracking and voice control in particular require a closer look.

There is already various research into BCI controls for wheelchairs. Different BCI measurement methods are used in this research: Steady-state Visual Evoked Potential (SSVEP) (Bastos, 2011; Chen *et al.*, 2022), Motor Imagery (MI) (Bastos, 2011; Palumbo *et al.*, 2021) and Face Expression (FE) (Rabhi, Mrabet and Fnaiech, 2018). What they all have in common is that they have the potential to control electric wheelchairs. With the help of these controls, people with a complete cross-section can gain the opportunity to control their wheelchair themselves.

The P300 technique is another BCI method that offers higher accuracy, less fatigue, and more commands than SSVEP (Puanhvuan *et al.*, 2017). The method by Puanhvuan *et al.* (2017) uses the P300 method when obstacle avoidance and other safety features are activated. This led to an accuracy of 83.42% with the developed control. The latency was around 250 to 500 milliseconds. This accuracy and latency could be insufficient for safe movement on the road. In addition, an extra monitor must be mounted on the wheelchair, which the user has to look at to control it, which also applies to SSVEP. This could make it difficult to observe the surroundings.

The MI-BCI attempts a different approach, but has some drawbacks that make it unsuitable for this project. The process needs a lot of training and user experience to work as a controller. Users need to have the ability to do certain mental activities, such as picturing moving an arm, that match the wheelchair's control commands. Not all users can acquire the necessary control over such a controller. Even experienced users who get along well with the procedure have highly fluctuating and therefore inaccurate success rates when executing mental commands. The resulting delays could pose a high safety risk in road traffic. In addition, MI-BCIs put a lot of cognitive strain on users. This can lead to difficulties when users need to manage multiple tasks at the same time, such as navigating a complex environment (Palumbo *et al.*, 2021).

BCI also allows facial expressions to be recorded. Facial expressions are natural and intuitive to implement. They can be implemented without much explanation from users and require no effort. Facial expressions also offer the advantage that they can be quickly recognized and interpreted. This could lead to responsive control like steering or stopping the wheelchair (Lin and Jiang, 2016).

The same applies to alternative controls based on eye tracking methods (Bai *et al.*, 2016; Cojocararu *et al.*, 2019; Luo *et al.*, 2021; Wanluk *et al.*, 2016; Wästlund *et al.*, 2015). These also show that wheelchair users can carry out the desired control commands even though they suffer from a

spinal cord injury. The most common method is to point the wheelchair in the appropriate viewing direction. This control proves to be responsive and reliable, but also requires a mechanism that prevents unintentional steering maneuvers when users simply want to look around.

Similar results were achieved with voice command based controls (Abdulghani *et al.*, 2020; Cao *et al.*, 2021; Jayakody *et al.*, 2019; Karim *et al.*, 2022).

Voice control is probably suitable for tasks that are not urgent, such as adjusting the backrests or leg supports (Abdulghani *et al.*, 2020; Shinde, 2023). Activating or deactivating various functions could also be implemented via voice. However, time-critical commands such as stopping or sudden steering could take longer than necessary in the event of danger. It is important to remember that not only speaking itself takes time, but also recognizing and processing the spoken words (Dearsley, 2024). In noisy environments, speech recognition can also be more difficult (Dua, Akanksha and Dua, 2023). These possible disadvantages could be minimized by sensible positioning of the microphone and modern ambient noise filters.

In the work of Taher, Amor and Jallouli (2016) BCI, eye-tracking and voice control techniques are combined. It uses an algorithm to determine the most likely control command from the user. The approaches of the work of Taher *et al.* form the basis for one of the developed alternative wheelchair controls. The second one will be based on eye-tracking and voice commands only. Both are described in the next chapter.

## SOFTWARE DESCRIPTION

The goal of the software is to provide users, medical staff like occupational therapist, of different experience levels with an intuitive and logically structured user interface. Calibration of the BCI hardware occurs within the application, eliminating the need for external software. This calibration is carried out step by step and with illustrated instructions. Users have the option to assign available BCI commands and eye movements to individual control commands. The BCI technology used is based on facial tracking, including eyebrow movements, eye blinks, smiles and frowns. Calibration data and control command assignments can be stored in user profiles along with voice commands.

The primary control method combines BCI, eye tracking and voice control. Using voice commands, you can switch between driving and steering controls and adjust the speed. The software recognizes the user's line of sight in order to control the wheelchair accordingly. To avoid unintentional steering, the user signals the intention to steer with a selected BCI command.

The secondary control focuses also on eye tracking and voice commands, but this control uses eye tracking for a different purpose. Due to the assumption that BCI might be prone to interference in everyday use, a second alternative control method was developed. This consists of a virtual joystick that is displayed as a two-dimensional circle. When the user fixates on this circle, it turns red and follows eye movements within a limited area, simulating the operation of a physical joystick. The voice commands were

used for setting the maximum speed, toggle the joystick and a rearview camera on and off. It is also possible to steer, start and stop the wheelchair with voice commands.

There are three virtual environments for testing the control. The first environment allows you to learn controls without obstacles on a straight, endless surface. The second environment depicts part of a small town, including houses, streets and sidewalks where obstacles, pedestrians and vehicles can be added. Some homes are walk-in, allowing testing of controls in confined spaces. In the third environment, users can playfully improve their control skills by completing a narrow route with variable obstacles.



**Figure 1:** Training environments of the software. From left: infinite world environment, neighbourhood, game environment (Own pictures, 2024).

## HARDWARE

The selection of hardware was based on the criteria of user-friendliness, robustness and cost-effectiveness. As a result, the brain-computer interface (BCI) device Emotiv EpocX and the VIVE Focus 3 VR headset with integrated eye sensors were selected. In case of users who suffer from motion sickness, the VR headset would be replaced with a conventional 27 inch-monitor in combination with a Tobii Eye Tracker 5. The software was designed to automatically switch depending on the connected hardware.



**Figure 2:** From left: Emotiv EpocX, VIVE Focus 3, Tobii EyeTracker 5 (adapted from Emotive, VIVE and Tobii, 2024).

The BCI headset used was characterized by a stable fit and ease of use that could even be used by technical laypeople. Also, the Emotive software and the calibration process of the headset is pretty straight forward. The headset proved to be reliable under controlled laboratory conditions. However, movements and vibrations led to a significant deterioration in the quality of the measurement data, which questioned its applicability in real conditions, such as on the road.

## TESTS PROCEDURE

The first assessment of the control systems for wheelchair users took place at the Center for Paraplegics in Hamburg, Germany. Three wheelchair users, two occupational therapists and the senior doctor took part.

The concept behind using a VR headset aimed to deeply involve the user in the simulation and achieve a high level of immersion and realism. Although virtual reality (VR) has now found widespread consumer acceptance, it was surprisingly found that the majority of the test subjects had no prior experience with VR. This led to initial difficulties in operating the controls, which is why a gradual introduction from simpler to more complex control methods were chosen, starting without a VR headset.

To test the developed control systems, the test subjects positioned themselves in their wheelchairs in front the monitor about one meter away. The eye tracker was attached to the bottom of the height-adjustable monitor, which was aimed at the seated person. This test arrangement was used for both types of control.

During the first test phase, the test subjects were introduced to the operation in a virtual infinite world (Fig. 1, left). Initially, control was carried out using line of sight without the use of BCI technology. This method proved to be quick to learn for all participants; Within one to two minutes, users were able to navigate intuitively. Facial recognition using BCI, on the other hand, proved to be more complex to acquire. Individual facial expressions varied in their ability to initiate control intentions. After a maximum practice time of five minutes in open space, the users switched to the ‘neighbourhood’ environment (Fig. 1, centre).

The second control method, a digital joystick, also underwent a testing phase. In contrast to the first control, the joystick can not only be used to steer. The eye control allows you to move the wheelchair back and forth and control the speed, just like a physical joystick. It took the test subjects about five minutes to feel safe in the infinite world. After that, they also went to the neighbourhood environment. It went out, that controlling the joystick and simultaneously keeping track of the environment and obstacles is a hard task. The test persons trained about ten minutes before they felt comfortable with the joystick.



**Figure 3:** Using virtual joystick in neighbourhood environment (screenshot from Simulator software, 2024).

## RESULTS AND DISCUSSION

The test persons rated the BCI-supported eye tracking control as not suitable for everyday use. Many found controlling the wheelchair challenging as the facial expressions required often led to confusion. In everyday life, facial expressions are often used unconsciously, which could lead to unintentional movements with BCI-supported control. Additionally, the unreliability of BCI technology made precise navigation in the virtual environment difficult.

This highlights the need to develop such systems to be robust to unintentional signals and to enable clear distinction between conscious control commands and random facial movements. It is crucial that the technology is both precise and fault-tolerant to ensure safe and effective use in everyday life. A key argument against this technology was the reluctance of all test subjects to use the BCI headset in public. They expressed that they would feel uncomfortable wearing the headset visibly because it would visually emphasize their disability. Consequently, further development of this alternative control technology will be discontinued.

The second control method, the digital joystick, also underwent a testing phase. Having the immediate control above the speed and driving direction, gave the test persons a safer feeling. The resemblance to a real joystick made the virtual one easy to use. Although the control proved workable, it was shown to require more intensive training, especially in an environment closer to reality. It became clear that a more comprehensive evaluation and fine-tuning of the control only makes sense when the simulation of the virtual wheelchair more precisely replicates the behaviour of a real wheelchair.

All test subjects still agreed that the virtual joystick as a different way of controlling is a viable method that should be further developed.

## OUTLOOK

In order to increase the effectiveness of the control, a detailed analysis of user interactions with the digital joystick must be carried out. This would make it possible to evaluate the reaction times, the accuracy of the movements and the intuitive aspects of the controls. In addition, adjustments are being made to the user interface to provide better visual feedback, helping the user learn the controls faster and use them more efficiently.

In a next step, the physical properties of a real wheelchair, such as inertia and acceleration behaviour, are integrated into the simulation to create a more authentic driving experience. These further steps will continue to be carried out iteratively together with wheelchair users.

In the next phase of research, the digital joystick will be developed on the Magic Leap 2 augmented reality (AR) glasses, which will later also be used in real wheelchairs. The joystick initially serves as a control element for the virtual wheelchair in the simulated test environment.

As part of the further development, it is planned to display the wheelchair's user-specific user interface (UI) on the AR display. In addition, the possibility should be created to operate this UI using eye control and voice commands.

After both the test subjects and the medical professionals have rated the control as user-friendly and safe, it is integrated into a real wheelchair. In

collaboration with the test subjects, the control is further optimized in an iterative process.

## CONCLUSION

The alternative wheelchair control methods presented in this study, particularly the digital joystick, show promising approaches to offering more autonomy to people with severe physical limitations. Iterative development and testing in virtual environments enable user-centred optimization that addresses the specific needs of the target group. Although BCI-based control is not being pursued due to accuracy issues and public perception of the technology, the digital joystick offers a viable alternative that will be further developed and implemented in real-world wheelchairs. Future research will focus on fine-tuning the controls and integrating them into physical wheelchairs to ensure safe and effective use in everyday life.

## REFERENCES

- Abdulghani, M. M. *et al.* (2020) 'Wheelchair neuro fuzzy control and tracking system based on voice recognition,' *Sensors*, 20(10), p. 2872. <https://doi.org/10.3390/s20102872>
- Bastos, T. *et al.* (2011) 'Robotic wheelchair commanded by SSVEP, motor imagery and word generation' *Annu Int Conf IEEE Eng Med Biol Soc.* <https://doi.org/10.1109/iembs.2011.6091177>.
- Bai, D. *et al.* (2016) 'Design of an eye movement-controlled wheelchair using Kalman filter algorithm' *2016 IEEE International Conference on Information and Automation (ICIA)*. <https://doi.org/10.1109/icinfa.2016.7832085>
- Cao, W. *et al.* (2021) 'Voice controlled wheelchair integration rehabilitation training and posture transformation for people with lower limb motor dysfunction,' *Technology and Health Care*, 29(3), pp. 609–614. <https://doi.org/10.3233/thc-202386>
- Chen, W. *et al.* (2022) 'An electric wheelchair manipulating system using SSVEP-Based BCI system,' *Biosensors*, 12(10), p. 772. <https://doi.org/10.3390/bios12100772>
- Cojocar, D. *et al.* (2019) 'Using an Eye Gaze New Combined Approach to Control a Wheelchair Movement, 23rd International Conference on System Theory' *Control and Computing*. <https://doi.org/10.1109/icstcc.2019.8886158>
- Dearsley, J. (2024) 'Let's Talk: Voicebots, Latency, and Artificially Intelligent Conversation'. <https://blog.vapi.ai/lets-talk/>
- Dua, M., Akanksha and Dua, S. (2023) 'Noise robust automatic speech recognition: review and analysis,' *International Journal of Speech Technology*, 26(2), pp. 475–519. <https://doi.org/10.1007/s10772-023-10033-0>
- Jayakody, A. *et al.* (2019) 'Smart Wheelchair to Facilitate Disabled Individuals,' *2019 International Conference on Advancements in Computing (ICAC)*. <https://doi.org/10.1109/icac49085.2019.9103409>
- Karim, A. *et al.* (2022) 'Raspberry Pi Based Voice Controlled Smart Wheelchair' *2022 International Conference on Emerging Trends in Smart Technologies (ICETST)*. <https://doi.org/10.1109/icetst55735.2022.9922929>
- Lin, J. and Jiang, Z.-Y. (2016) 'An EEG-Based BCI system to facial action recognition,' *Wireless Personal Communications*, 94(3), pp. 1579–1593. <https://doi.org/10.1007/s11277-016-3700-3>



- Luo, W. *et al.* (2021) 'A Human-Computer control system based on intelligent recognition of eye movements and its application in wheelchair driving,' *Multimodal Technologies and Interaction*, 5(9), p. 50. <https://doi.org/10.3390/mti5090050>
- Palumbo, A., *et al.* (2021). 'Motor-Imagery EEG-Based BCIs in Wheelchair Movement and Control: A Systematic Literature Review,' in: *Sensors* 21(18), p. 6285. <https://doi.org/10.3390/s21186285>
- Puanhvuan, D. *et al.* (2017), 'Navigation-synchronized multimodal control wheelchair from brain to alternative assistive technologies for persons with disabilities', in: *Cognitive Neurodynamics*, Springer, Wang (Ed.), Volume 11. pp. 117–134, <https://doi.org/10.1007/s11571-017-9424-6>
- Rabhi, Y., Mrabet, M. and Fnaiech, F. (2018) 'A facial expression controlled wheelchair for people with disabilities,' *Computer Methods and Programs in Biomedicine*, 165, pp. 89–105. <https://doi.org/10.1016/j.cmpb.2018.08.013>
- Shinde, S. (2023) 'Smart voice controlled wheelchair for physically disabled people,' *International Journal for Research in Applied Science and Engineering Technology*, 11(11), pp. 536–541. <https://doi.org/10.22214/ijraset.2023.56476>
- Taher, F. B., Amor, N. B. and Jallouli, M. (2016) 'A self configured and hybrid fusion approach for an electric wheelchair control,' *2016 IEEE 8th International Conference on Intelligent Systems (IS)*. <https://doi.org/10.1109/is.2016.7737392>
- Wanluk, N. *et al.* (2016) 'Smart wheelchair based on eye tracking' *2016 9th Biomedical Engineering International Conference (BMEiCON)* [Preprint]. <https://doi.org/10.1109/bmeicon.2016.7859594>
- Wästlund, E. *et al.* (2015) 'Evaluating gaze-driven power wheelchair with navigation support for persons with disabilities,' *Journal of Rehabilitation Research and Development*, 52(7), pp. 815–826. <https://doi.org/10.1682/jrrd.2014.10.0228>

# Collaborative Learning Through XR. A Study of Eye- and Hand-Based XR Interactions to Support Collaborative Learning in the Chemistry Classroom

Wisanakorn Boribun and Frank Heidmann

University of Applied Science Potsdam, Brandenburg, 14469, Germany

## ABSTRACT

In this study, we examine the utilization and efficacy of Extended Reality (XR) technologies in education. The focus is on understanding how XR interactions, particularly those centred around eye and hand movements, can enhance collaborative learning. To achieve this, we developed a specialized XR application tailored for this purpose. By leveraging augmented reality technologies, we created an interactive and immersive learning environment where fundamental chemical concepts can be visually and tangibly represented. The application enables students to interact with and study virtual sugar molecules, facilitating collaborative learning as they learn to distinguish between them. Additionally, the application incorporates eye-tracking and hand-tracking technologies, facilitating natural interaction and collaboration among students. Through comprehensive analysis and multiple application tests, we investigate how these advanced technologies can enhance understanding of basic chemical concepts and promote collaboration among students. Finally, we discuss the impact of such technologies on the overall learning environment and classroom dynamics, highlighting both the advantages and challenges of integrating XR technologies into teaching.

**Keywords:** XR (extended reality), VR (virtual reality), Collaborative learning, Augmented reality, Education, Chemistry classroom, Eye-tracking, Hand-tracking, Eye-based interaction

## INTRODUCTION

The digital transformation is reshaping various sectors of our society, with education undergoing significant changes (Radianti et al., 2020). Amid this evolution, Extended Reality (XR), encompassing Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), emerges as a powerful tool for fostering interaction and collaboration, especially within educational settings (Merchant et al., 2014; Billingham & Duenser, 2012; Dunleavy & Dede, 2013). The integration of XR and VR into education holds the potential to enrich learning experiences by facilitating collaboration and shared comprehension (Radu, 2014; Akçayır & Akçayır, 2017). Collaborative learning, in particular, assumes a pivotal role, enabling learners to collaborate on problem-solving, exchange ideas, and glean insights from each other (Johnson et al., 2014). This collaborative approach not

only enhances understanding and knowledge absorption but also nurtures essential skills such as communication, teamwork, and critical thinking (Ott & Freina, 2015). Despite the immense promise of XR and VR, many applications remain segregated (Dalgarno & Lee, 2009). Shared experiences often rely on screen streaming or verbal explanations, potentially excluding individuals who are not directly engaged (Slater & Wilbur, 1997). Even in scenarios where multiple users utilize VR or XR headsets, shared experiences often remain confined to specific contexts, such as multiplayer VR games (Billinghurst et al., 2015). Addressing these challenges, we have developed a prototype harnessing the eye and hand tracking functionalities of the HoloLens 2. This prototype facilitates collaborative interactions with 3D elements within an XR environment. By tracking participants gazes within the XR scene, users can effortlessly identify the elements they are currently viewing. Consider, for instance, the ability to observe what your partner is currently focused on. This capability could significantly enhance communication and shared understanding, as each participant gains insight into the other's focus. Moreover, it promotes collaboration by enabling participants to concentrate on shared elements or issues, facilitating joint problem-solving (Yuen et al., 2011).

Our current prototype is tailored to explore the intricate molecular structures of sugar molecules within a collaborative XR environment (Santos et al., 2014). Teams of users can collaboratively delve into molecules such as fructose, glucose, galactose, and mannose, gaining deeper insights into their unique properties. To evaluate the functionality and cognitive impact of our XR prototype, we recruited 19 participants for assessment, ensuring they lacked prior knowledge or involvement in chemistry or related fields concerning sugar molecules.

Our work contributes to: (i) the exploration of XR technology's potential to enhance collaborative learning in educational contexts, (ii) the development of a novel prototype leveraging eye and hand-tracking capabilities to foster interactive learning experiences, and (iii) the provision of valuable insights into user experience and cognitive impacts, contributing to the broader discourse on XR in education (Wojciechowski & Cellary, 2013).

## **DEVELOPMENT AND STATUS OF XR IN EDUCATION**

Extended Reality (XR) is increasingly gaining popularity as an educational technology. It is extensively used in educational institutions to facilitate asynchronous collaboration among learners at various locations (Merchant et al., 2014). A contemporary example of XR application in education is "ClassVR," which integrates 360-degree experiences and Virtual Reality (VR) into the learning process. While there are already various XR and VR applications that enable and promote collaboration, there remains a need for further development and research in many areas, especially in education. The goal should be to create and improve XR environments that optimally support collaborative learning. Furthermore, XR is being applied in various other fields such as medicine, real estate, retail, and entertainment. For instance, in medicine, XR allows for the simulation of surgeries to train

medical professionals. In the real estate industry, potential buyers or tenants can take virtual tours of properties. In retail, customers can visualize and test products in an XR environment, and in the entertainment industry, XR enables immersive gaming and experiences.

Although the field of Virtual Reality (VR) is experiencing continuous innovations and advancements, the integration of eye-tracking technologies into many modern VR headsets is not yet fully implemented. This does not necessarily represent a deficiency but rather highlights unexploited potential for enhanced interaction mechanisms. With the use of eye-tracking technologies, users would no longer depend on aiming at an object in the XR environment with a controller to interact with it. Instead, they could select which element to interact with through eye movements. This could significantly increase the speed of interaction and improve the user experience. Therefore, there is an increasing need for further research and development in this area to integrate eye-tracking and other intuitive interaction methods into XR systems.

In our work, we explore new ways of visualization and interaction in XR systems to enhance collaborative learning within the XR environment. Unlike previous studies that focused on specific aspects of XR systems, our goal is to develop a system that provides a comprehensive and immediate overview of the shared learning environment (Billinghurst & Dünser, 2012).

## **COGNITIVE ASPECTS OF EYE-BASED XR INTERACTIONS**

To explore the cognitive aspects of eye-based XR interactions, it is essential to have a solid understanding of eye-tracking technology and its integration into human-computer interfaces. Majaranta and Bulling (2014) provided a comprehensive overview of the development of eye-tracking technology, emphasizing its central role in medical and psychological research before its application in interactive, gaze-based applications. The implementation of eye-tracking in human-computer interaction relies on the eye's ability to convey cognitive processes and intentions (Majaranta & Bulling, 2014). The link between cognitive processes and eye movements is particularly significant for the design of XR systems. Majaranta and Bulling (2014) highlighted how eye tracking can enhance user interaction by leveraging natural eye movements, a principle applicable to XR applications for creating intuitive and efficient user interfaces.

The primary goal of our test prototype was to apply these principles and findings in practice. Specifically, we aimed to utilize natural eye movements to analyse user interactions in the XR environment. For instance, gaze fixation on an object, observed long enough to be processed by the brain's visual system, can serve as a cognitive task triggering interactions in virtual environments. By integrating eye-tracking technology into our prototype, we seek to gain deeper insights into the cognitive aspects of eye-based XR interactions.

## FIXATION DURATION IN HUMAN-COMPUTER INTERACTION

Within the context of eye movement research, fixation is defined as a sustained visual focus on a specific point (Negi and Mitra, 2020). Longer fixations are believed to indicate that a person is processing or interacting with the information at that point. When developing our prototype, we considered findings from a study by Negi and Mitra (2020), who emphasized the importance of fixation duration in human-computer interaction. They found that longer fixation durations often indicate deeper cognitive involvement. Based on this understanding, we adjusted the interactivity of our user interface, designing it to activate only after a fixation duration of 300 ms and allowing a tolerance of up to 500 ms (Negi and Mitra, 2020). This design approach serves to improve the user experience in two important ways: firstly, by ensuring that interactions and information become accessible or visible only when the user is ready to engage with them, and secondly, by helping to prevent accidental selection of objects within the XR layer.

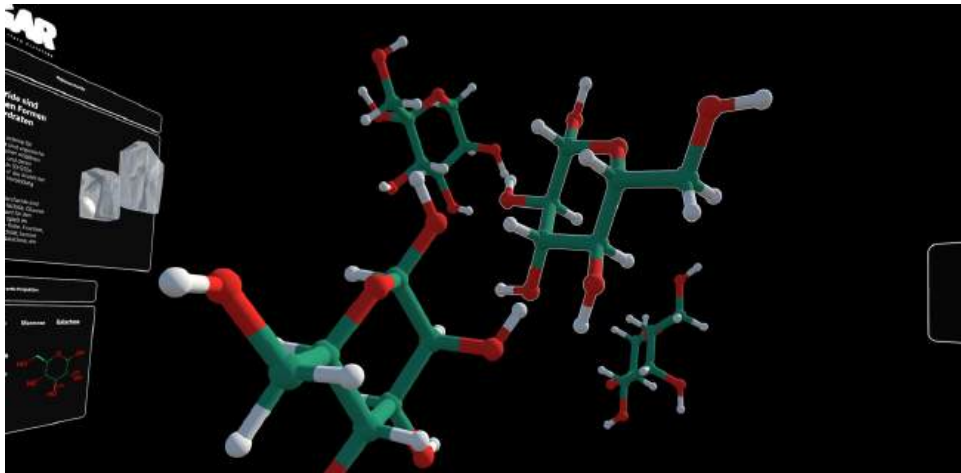
**Table 1.** Overview of fixation durations and their impact on learning processes based on findings by Negi and Mitra (2020).

Fixation Duration	Typical Significance	Potential Impact on Learning
Under 100 ms	Often too short for conscious perception	May indicate distraction or rapid scanning
100–200 ms	Brief fixation, limited information intake	Suggests minimal engagement or familiarity
200–300 ms	Normal for reading simple texts	Standard fixation duration for text learning
300–500 ms	Longer fixation, deeper cognitive engagement	Indicates intense engagement, likely comprehension
500–800 ms	Very long fixation, potential difficulty or interest	Could signify confusion or deep interest
Over 800 ms	Extremely long fixation uncommon in normal reading	May suggest cognitive overload or intense focus on a complex issue

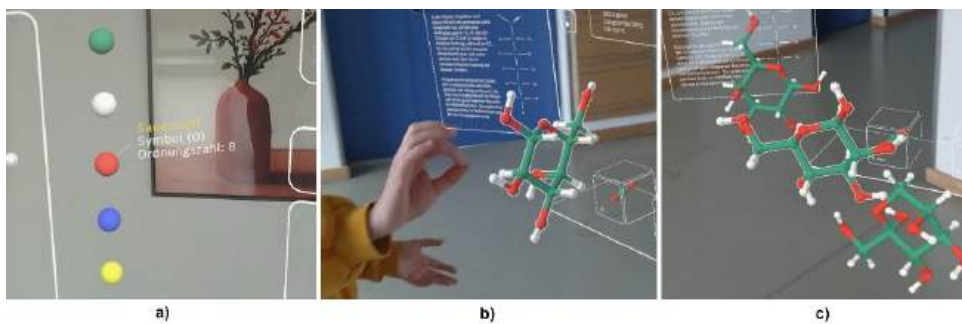
## XR PROTOTYPE (XVGAR)

The prototype is implemented using Unity 2021.3.5f1 and the Mixed Reality Toolkit (MRTK), designed specifically for standalone use with the HoloLens 2 mixed reality glasses, without needing a wired connection to a computer. A crucial feature of our prototype is gaze interaction, facilitated by the HoloLens 2's eye-tracking capability. This technology captures the user's gaze direction in real-time, accurately pinpointing their focus point. This is achieved through raycasting, where an invisible virtual beam, originating from the user's eyes, determines their gaze direction and collects collision data with 3D objects in the XR environment. When the raycast intersects with a 3D object, such as interfaces in the XR environment or a model of a sugar molecule, that object becomes the user's current focus point.

When using eye-tracking technologies, it is common to capture the gaze direction of each eye separately. However, since we naturally focus on a specific point with both eyes, determining a midpoint between the two gaze directions is necessary. This process, called interpolation, calculates a single central focus point from the gaze directions of both eyes. This method is particularly crucial in three-dimensional applications like the HoloLens 2, enabling precise interaction with elements in three-dimensional space. Additionally, eye-tracking systems account for small deviations, known as “vergence,” to ensure accurate calculations and determine the precise midpoint between the two gaze directions.



**Figure 1:** An exemplification of the XVGAR prototype. An eye-based interaction (fixation) is displayed, enabling users in the XR-Environment to identify the molecule currently being observed - glucose, fructose, mannose, and galactose.



**Figure 2:** a) The graphic illustrates a tool-tip within the prototype, which becomes visible upon achieving a fixation duration of 300ms. Specifically, the sphere (oxygen atom) detects the user's fixation and subsequently displays pertinent information. b) This scenario incorporates eye-based interaction in conjunction with hand-tracking capabilities. The user employs a pinching gesture to interact with the object identified through the gaze-tracking mechanism. c) This graphic depicts the synchronization of fixation, indicating that all users are collectively focusing on a single molecule within the XR environment, as other sugar molecules are not highlighted.

The synchronization of the captured gaze information is done using the Photon Unity Networking (PUN2) plugin. PUN is a network plugin for Unity that facilitates real-time data exchange between users. In our project, we utilized PUN to synchronize the gaze information captured by the eye-tracking function. This gaze data, packaged into data packets, is transmitted to all connected users. Each packet contains detailed information about the position and orientation of the raycast, as well as the objects struck by the raycast. When a user fixates on a specific object, it becomes visible to the entire user group. After a designated fixation time (300–500ms), a visual signal is generated, indicating that the object is being viewed, enabling other users to precisely identify the object the first user is focused on.

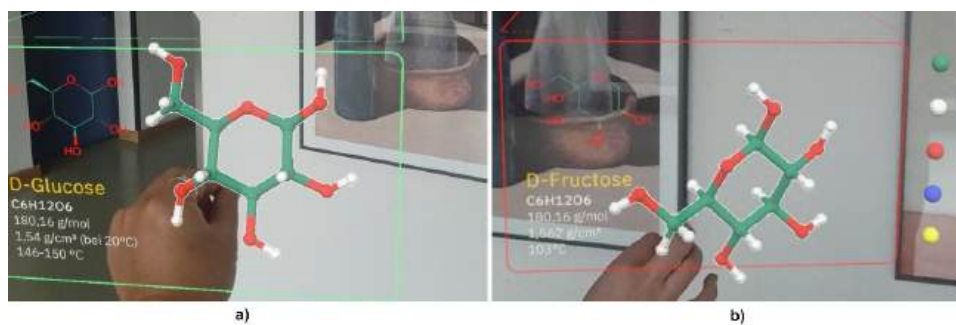
In addition to gaze data synchronization, we have also implemented hand tracking data synchronization. This means that with each shift of an object, data packets are transmitted, ensuring that every change is immediately visible on the displays of all users. This facilitates seamless and intuitive collaboration as users work on the same object in real time, with modifications visualized instantly. Technically, each interaction with the object generates a data packet sent via a network protocol to all connected devices. These packets contain specific information about the position and movement of the manipulated object, interpreted and implemented by other devices to maintain a cohesive representation of the interaction for all users.

In designing the prototype, we placed significant emphasis on interaction with the physical world. The interface is designed to be translucent, allowing users to perceive their physical environment while simultaneously interacting in augmented reality. This ensures that the interface does not impede physical tasks' performance and that users remain connected to their real environment. By integrating these technologies and design principles, we have developed a prototype that serves as an effective tool for collaborative work in augmented reality (Greenwald et al., 2017). These characteristics render the prototype ideal for our learning scenario and the execution of our study.

## EVALUATION

In our initial formative study, we evaluated the effectiveness of the XR system, particularly focusing on its ability to simplify complex chemical concepts and promote collaborative learning. The study comprised two phases. In the first phase, a tutor explained the properties of existing sugar molecules, the meaning of different projection types, and how to distinguish and interpret them. In the second phase, participants had the opportunity to examine sugar molecules in the XR environment and collaboratively decide which sugar molecule should be inserted into each differentiation system.

We invited a group of nineteen participants, not directly involved in the field of chemistry, to test our XR prototype in small groups. Surprisingly, we could almost forego an explanation of the controls, as participants perceived eye control as an intuitive method for interaction in the XR environment. This improved the naturalness of the communication process and allowed users to quickly familiarize themselves with control and communication at the XR level. Using eye and hand control felt like a natural interaction for most.



**Figure 3:** a) The XR environment features a synchronized system for distinguishing sugar molecules. In this illustration, the user places the correct molecule in the corresponding field. The box changes its color to green to visualize the correct placement of the molecule. b) This graphic displays the distinction system when an incorrect molecule is returned to the wrong field, causing the box to turn red, visible to all users.

Interestingly, our findings suggested that our XR system enabled participants to distinguish between different sugar molecules more quickly. This could be attributed to the interactive and playful way molecules were differentiated together, possibly due to the increased focus enabled by direct gaze control. Besides improving communication, our system also appeared to enhance individual learning and information processing abilities, indicating its potential to accelerate the learning process and enhance the overall user experience. Essentially, this study provided valuable insights into the potential of integrating eye-based interactions into XR systems for educational purposes. The results underscore the potential of XR technology to transform educational practices by creating engaging, intuitive, and collaborative learning environments.

## CONCLUSION

In this work, we have introduced a comprehensive XR system based on eye and hand tracking interactions. We presented our investigations into the use of eye-tracking technologies in XR environments and described the development of our prototype employing these technologies. Moreover, we explored the cognitive aspects of eye-based XR interactions, highlighting the interplay of cognitive processes and eye movements. The possibility of utilizing eye-tracking technology with XR to synchronize fixations opens new opportunities for various fields, not just in educational contexts. Currently, we are testing our prototype in several studies, particularly aiming to enhance the user experience in collaborative XR environments. We are reflecting on the use and development of our technologies and presenting the insights gained. Our work aims to contribute to the continuous improvement of these tools and the resulting research in XR technologies. However, we believe that XR technologies should not replace traditional teaching methods and books, but rather complement them. They should serve to enhance students' collaboration and interaction, making learning more effective and engaging.



## ACKNOWLEDGMENT

The authors would like to thank all participants and chemistry teachers for their participation in the evaluation. Their educational insights were invaluable to us.

## REFERENCES

- Akçayır, M. & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11. <https://doi.org/10.1016/j.edurev.2016.11.002>
- Billinghurst, M., Clark, A. & Lee, G. (2015). A Survey of Augmented Reality. *Foundations And Trends in Human-computer Interaction (Print)*, 8(2–3), 73–272. <https://doi.org/10.1561/11000000049>
- Billinghurst, M. & Duenser, A. (2012). Augmented Reality in the Classroom. *Computer*, 45(7), 56–63. <https://doi.org/10.1109/mc.2012.111>
- Dalgarno, B. & Lee, M. J. W. (2009). What are the learning affordances of 3-D virtual environments? *British Journal Of Educational Technology*, 41(1), 10–32. <https://doi.org/10.1111/j.1467-8535.2009.01038.x>
- Dunleavy, M. & Dede, C. (2013). Augmented Reality Teaching and Learning. In Springer eBooks (S. 735–745). [https://doi.org/10.1007/978-1-4614-3185-5\\_59](https://doi.org/10.1007/978-1-4614-3185-5_59)
- Greenwald, S., Corning, W. & Maes, P. (2017). Multi-User Framework for Collaboration and Co-Creation in Virtual Reality. MIT Media Lab. Abgerufen von <https://dspace.mit.edu/handle/1721.1/108440>
- Johnson, D. W. & Johnson, R. T. (2014). Cooperative learning in 21st century. *Journal Article*, 30(3), 841–851. Abgerufen von <https://doi.org/10.6018/analesps.30.3.201241>
- Majaranta, P. & Bulling, A. (2014). Eye Tracking and Eye-Based Human–Computer Interaction. In *Human-computer interaction series* (S. 39–65). [https://doi.org/10.1007/978-1-4471-6392-3\\_3](https://doi.org/10.1007/978-1-4471-6392-3_3)
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W. & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70, 29–40. <https://doi.org/10.1016/j.compedu.2013.07.033>
- Negi, S. & Mitra, R. (2020). Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal Of Eye Movement Research*, 13(6). <https://doi.org/10.16910/jemr.13.6.1>
- Ott, M. & Freina, L. (2015). A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. *eLearning And Software For Education*. <https://doi.org/10.12753/2066-026x-15-020>
- Radianti, J., Majchrzak, T. A., Fromm, J. & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers And Education (Print)*, 147, 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal And Ubiquitous Computing*, 18(6), 1533–1543. <https://doi.org/10.1007/s00779-013-0747-y>
- Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J. & Kato, H. (2014). Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation. *IEEE Transactions On Learning Technologies*, 7(1), 38–56. <https://doi.org/10.1109/tlt.2013.37>

- Slater, M. & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence* (Cambridge, Mass.), 6(6), 603–616. <https://doi.org/10.1162/pres.1997.6.6.603>
- Wojciechowski, R. & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers And Education* (Print), 68, 570–585. <https://doi.org/10.1016/j.compedu.2013.02.014>
- Yuen, S. C., Yaoyuneyong, G. & Johnson, E. (2011). Augmented Reality: An Overview and Five Directions for AR in Education. *Journal Of Educational Technology Development And Exchange* (Print), 4(1). <https://doi.org/10.18785/jetde.0401.10>

---

# Virtual Ergonomics - Ergotyping in Virtual Environments

**Stefan Pfeffer, Marc Rößler, Simone Maag, Lisa-Marie Langwaldt, Lara Schunggart, Markus Strigel, Abigail Senger, and Moritz Ochtrop**

Furtwangen University, Tuttlingen, 78532, Germany

## ABSTRACT

Ergonomic assessment of manual work processes is important to prevent workplace injuries. Virtual reality simulations can be used to carry out an evaluation of work equipment and workplaces very early on. In combination with motion tracking analyses, data on posture during task performance and product use can then be collected. However, not all work situations can be equally represented in a virtual simulation. In particular, the virtual analysis of load handling poses a challenge in simulation, as body posture changes under the influence of external load weights. The aim is to increase immersion to bring the body movements in the virtual simulation closer to those in the real simulation with weights. For building up VR simulations with different aspects of visual, auditory and haptic immersion a scheme called immersion cube is presented. In order to be able to simulate load handling in VR, the immersion cube is used to investigate how much haptic immersion is needed to obtain sufficiently good data for the body movements measured in a VR setting. The first study showed that the deviation between real and virtual executions depends heavily on the task (lifting from the ground, move while standing, lifting over the shoulder). In some tasks, virtual and real simulation are very close to one another for certain body movements and could therefore in principle be used for ergonomic assessment. On the other hand there are still movements that vary between these two forms of execution and therefore show a need for increasing the immersion.

**Keywords:** Ergonomics, Virtual reality, Immersion cube, Weight perception

## INTRODUCTION

Musculoskeletal disorders (MSDs) resulting from inappropriate ergonomic design of workplaces represent a main factor for days of incapacity for work and therefore loss of productivity of companies and economic burden of health systems all over the world (Bevan, 2015; Briggs et al., 2018). In particular, the ageing society and the need to continue working in old age make it necessary to pay even more attention to ergonomic workplace design in order to ensure the ability to work and to be able to employ older people and workers with reduced working capacity, taking into account their individual capabilities and limitations. This view goes hand in hand with the WHO's guiding principles for the decade of healthy aging.

In Germany, most of the illnesses caused by mechanical impacts at work (reported suspected cases) can be attributed to “load handling” - with an increasing trend in numbers year by year (BAuA, 2023). For this reason, the German regulation on the implementation of EU directives in occupational health and safety, the so-called load handling regulation - also points out that physical suitability of the employees must be taken into account.

Load handling is defined as the handling of loads equal or greater than 3 kg (Serafin et al., 2020). Furthermore a distinction is made between the following forms of load handling:

- Lifting, lowering or transferring is the movement of a load from one position to a lower, same or higher position by muscular force,
- Holding is the fixing of a load in a certain position by muscular force
- as a predominantly static process,
- Carrying is the predominantly horizontal transportation of a load that does not reach the ground using human strength and by carrying it on the body.

In order to design a workplace ergonomically with regard to load handling, the nature of the load to be carried (weight, shape, size, etc.) and the working environment (space requirements, floor conditions, lighting, etc.) must be taken into account, as well as the employee’s work task. This includes, for example, body postures and movements and the relation of the load object to the person. Unfavorable stresses resulting from body postures and movements under load caused by the workplace design include, for example, all work requiring extreme trunk flexion, lifting with simultaneous trunk rotation and lateral flexion of the upper body or hollow back posture (trunk extension) when carrying loads.

## **VIRTUAL ERGONOMICS AND ERGOTYPING**

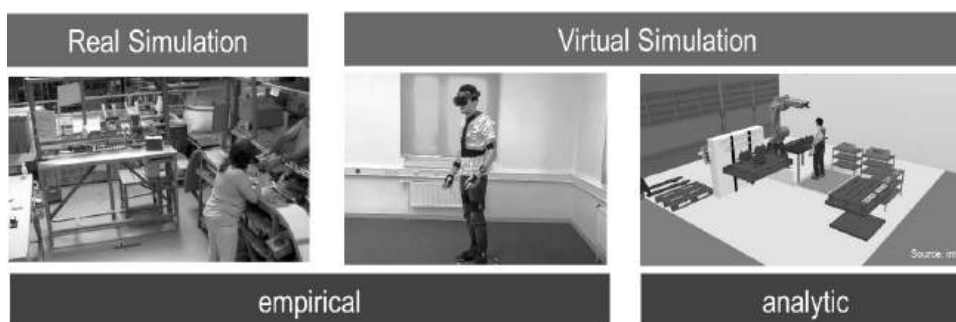
Virtual ergonomics refers to the use of computer-aided methods and tools for the ergonomic design of products, work systems and processes (Bullinger-Hoffmann & Mühlstedt, 2010). The term “ergotyping” can be traced back to Kamusella and Schmauder (2010) and, as an artificial word made up of the terms ergonomics and prototyping, describes a discipline in which ergonomic assessments can be implemented as early as possible in product or process development by using methods and digital tools to analyze, evaluate and design ergonomic aspects. The earlier an assessment of ergonomic conditions can take place, the more cost-effective and faster product and process development will be.

As a typical tool of virtual ergonomics, digital human models are used to create prospective simulations of a work system. For example, with the digital human model of the ema Work Designer (emaWD) software, scenes and behavior modeled in 3D can be assessed analytically with regard to ergonomic criteria (see Figure 1, right). In this context, analytical means that calculations, for example of posture, are not based on directly empirical collected data. In contrast, real simulations of work processes can be set up prototypically with simple materials (cardboard engineering). Here, for

example, body postures can be objectively determined using motion tracking systems. The use of motion tracking systems for the evaluation of workplaces has already been elaborated in several studies (e.g. Rybníkář et al., 2023; Caputo et al., 2018). The advantage of this method is that inter-individual differences in execution can also be taken into account. Without wanting to replace these tools, but rather as a supplement, virtual-empirical simulation is positioned here as a method between these two poles (see Fig. 1, middle). Here, workplaces are modeled virtually and empirical data is obtained through the use of virtual reality (VR) technologies in conjunction with motion tracking. The combination of VR and motion tracking has also been used in studies, although these were predominantly scenarios without load handling (e.g. Kačerová et al., 2022; Simonetto et al., 2022). Virtual-empirical simulation offers the following advantages in addition to the other two methods:

- individual movements of the workers can be taken into account, whereby the virtual workplace conditions can be adapted directly,
- there is no risk from a real load, as the load objects are only handled virtually. This means that people with reduced working capacity and older people can carry out the simulation without injury,
- the virtual environment can be scaled so that a person can experience the simulated workplace in different anthropometric body sizes - both as a very small person with challenges in accessibility and limited upward reach and as a very tall person with the challenges of space restrictions and limited downward reach.

Further advantages of the virtual-empirical approach are also mentioned by Kačerová et al. (2022). Compared to real simulation, considerable amounts of material can be saved, simulations of workplaces in potentially hazardous environments are possible and many alternative solutions can be tested in a very short time.



**Figure 1:** Possibilities for early simulation and ergonomic testing of workstations.

For workplaces where movements without loads  $\geq 3\text{kg}$  take place, body postures and movements, reachability, posture and dimensional analyses can already be assessed using the virtual-empirical method. However, as soon as movements with an external load  $\geq 3\text{kg}$  take place, such as when

lifting, holding and carrying objects, the posture changes in comparison to load-free movements. Adaptive changes in posture take place in order to compensate for the additional load and ensure the stability and efficiency of the movement. The posture data collected in virtual-empirical studies then could lose validity and the assessment of ergonomic conditions could lead to misjudgments. Kačerová et al. (2022) already found differences between real and virtual simulations with regard to the duration of postures taken.

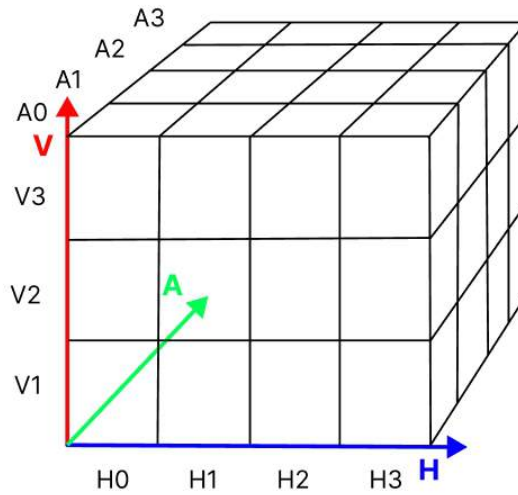
In order to be able to use the virtual-empirical method for the early assessment of load handling, the aim must be to make the body movements in the virtual simulation as real as possible and to discuss a compromise between data quality, technical effort and benefits for early assessment. In order to investigate this compromise, a scheme called “immersion cube” was developed to systematically increase physical immersion in subsequent studies in the hope of increasing mental immersion and the experience of presence to such an extent that sufficiently good posture data can be generated. Weight perception is particularly important for the simulation of load handling in VR. There are already concepts for this from research. Lim et al. (2021) provide a comprehensive overview.

### **IMMERSION CUBE**

Since the appearance of the first standalone head-mounted displays, VR technology has become broadly attractive for use in industry and research. It is known from VR training that an increase in immersion and the associated more realistic simulation of a situation or process can lead to better learning outcomes (e.g. Månsson, 2018). The better the immersion and presence experience, the more valid the results that can be drawn from it in simulations. It is known from studies that immersion can be increased through a multimodal presentation (Dinh et al., 1999; Gallace et al., 2012).

Figure 2 presents a trimodal scheme for visual, auditory and haptic perception. As an extension of Milgram and Kishino’s diagram (Milgram & Kishino, 1994), the immersion cube lists auditory and haptic immersion in addition to visual immersion and represents these three modalities in the form of a cube. Each modality is divided into levels in terms of the expected degree of immersion, with the levels representing the degree of virtuality or reality of content from 0=purely virtual to 3=real.

In the visual dimension, a distinction is made between virtual objects in virtual environments (V1), virtual objects in real environments (V2) and real objects in real environments (V3). In V2, a further distinction can be made between virtual objects in indirectly (V2.1) or directly perceived real environments (V2.2). Whereas V1 is called “Virtual Reality” by Milgram and Kishino (1994) and V2.2 is called “Augmented Reality”, there’s a new form of visual immersion coming from the technological possibilities of passthrough modes, where the real world is only indirectly perceived through a real-time video (V2.2). Another sub-form of level V2.2 results from open-built headsets such as the Quest Pro, where the virtual objects are not superimposed on the real environment, but the virtual and real environments are experienced simultaneously adjacent to each other.



**Figure 2:** Immersion Cube; V=Visual; H=Haptic; A=Auditory.

The auditory dimension would be possible without sound (A0), with virtual object sound and virtual surrounding sound (A1), real object sound and virtual surrounding sound (A2) as well as real object and surrounding sound (A3).

Both visual and auditory perception are distant senses. Haptics, on the other hand, is a near sense. In the real world, we perceive objects haptically in the area of exteroception through direct contact. These real haptics can also be made experienceable in virtual settings through the use and tracking of so-called haptic proxy elements. Following Lim et al. (2021), this form of haptic immersion is called “direct haptic” (H3). In addition, there are many studies and technical concepts for integrating haptic perception into VR via substitute cues. For the haptic submodality of force feedback, for example, this could be haptic gloves (Caeiro-Rodríguez et al., 2021). The force feedback of the gloves makes it possible for virtual objects that are gripped to be experienced indirectly haptically. Another form of indirect haptics would be the substitution of one specific submodality of haptic perception with another. An example of this is the replacement of proprioceptive perception (e.g. weight perception) with exteroceptive perception (e.g. vibrotactile presentation of information via a vest). Lim et al. (2021) call this “indirect haptic” (H2). Another way of transferring haptic information from the real world to the virtual world is to transmit the information via a perceptual channel other than the haptic one (H1). Substitution by the visual channel has been investigated many times. Here, for example, a pseudo-haptic effect is generated by changing the control-display ratio (Samad et al., 2019). Finally, level H0 describes a simulation in which no haptic information is transmitted. This is basically only possible using hand tracking, but the use of controllers (without haptic feedback) should also be included here.

In the combination of the three dimensions, simulations can now be described in terms of their immersion modalities. For example, a classic

VR simulation with hand tracking could be described as V1/H0/A1. With controllers that can reproduce vibrotactile haptic feedback, this would become a V1/H2/A1 simulation. The simulation described by Milgram and Kishino (1994) as “augmented virtuality” would be equivalent to e.g. a V1/H3/A1 and “augmented reality”, with headsets such as a Hololens 2 or Magic Leap 2, would be equivalent to a V2.2/H0/A1. The mixed reality made possible in many new headsets using passthrough mode and with hand tracking would be a V2.1/H0/A1.

## EXPLORATORY STUDY

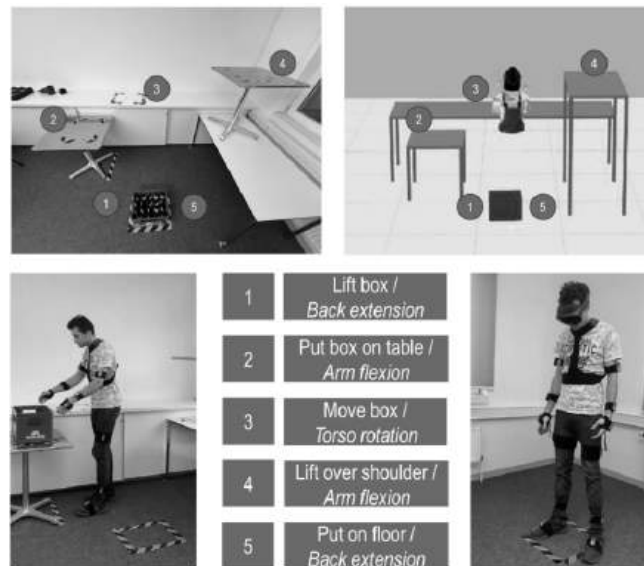
A standard procedure was developed to investigate the different effects of haptic immersion options on body movement during load handling (see Figure 3). This includes the typical, problematic activities of lifting (1), in which the back extension is considered, setting down an object with flexed arms (2), transferring with back rotation (3), lifting above shoulder height with flexed arms (4) and setting down on the floor with back extension (5). In the first exploratory study, this procedure was carried out in the specified order 1–5 in four executions - in each case with three real weights (light, medium, heavy) - according to the Immersion Cube - in a V2.2/H0/A0 simulation. The three real executions were balanced for all test subjects in a test plan. The executions in VR always took place in the fourth position. The virtual scenario was set up in the ema Work Designer software (see Figure 3, above right) from imk Industrial Intelligence GmbH, Chemnitz, Germany. The VR headset used was the Oculus Quest Pro with controllers. This allowed the participants to see the virtual visual simulation in the central field of view and the real environment in the peripheral field of view. The controllers only served as a means of interaction for gripping the box and did not transmit any haptic feedback (H0). An acoustic simulation was not implemented (A0).

A differently filled bottle crate was selected as load object, which handles offers a typical gripping situation, in certain cases also a problematic gripping situation. An empty bottle crate with a weight of 1.7kg was selected as the light weight for both male and female participants. A distinction was made between men and women for the medium and heavy crates. The medium crate weighed 13.9 kg for men and 7.5 kg for women. The heavy crate weighed 19.1 kg for men and 10.4 kg for women. The reason for the differentiation is that both genders should be subjected to a similar load. The loads were estimated using the “Leitmerkmalmethode” (Serafin et al., 2020) and selected comparably for both genders. The heavy load condition was therefore selected so that it is in a rather low risk range with a moderately increased load. This ensured that the risk of injury to the participants is kept to a minimum, but that the heavy crate still requires effort to complete the task.

Body movements were measured using inertial sensors in all executions. The T-sens sensors and the CAPTIV software from TEA, France were used. In total 6 participants took part in the test. Of these, 3 participants were female and 3 participants were male. The average age of the participants was 23.8 years. The average arm length of the participants was 72.2 cm, with the



shortest arm length being 67.5 cm and the longest arm length being 79 cm. The average height of the participants was 176.5 cm. Both taller and shorter participants were represented, with the shortest participant being 168 cm tall and the tallest 186 cm. In order to limit individual forms of execution somewhat, the participants were instructed to observe the following: the box should be lifted from the floor with legs extended (step 1) and put down to the floor with legs extended (step 5). In step 3 the rotation should be made from the upper body and after each step the crate should be released and an upright posture assumed. Before the data collection took place in the virtual simulation, the participants had the opportunity to familiarize themselves with the VR situation and test the interaction with the controllers.



**Figure 3:** Standard procedure for the real (left) and virtual (right) simulation of load handling.

## RESULTS

The results of the first study are presented below. The data was analyzed with regard to initial trends. Data analysis was ultimately possible for four of the six test participants, as the motion tracking data for two test participants could not be used for all executions due to test artefacts. Intra-individual comparisons were carried out for the remaining four participants. For this purpose, the measured movement data was first divided into categories commonly used for body posture analysis in ergonomics. As long as the body movement is in the “green” angle range, there is no risk potential. In the angle ranges “orange”, there is a possible risk and measures to eliminate the hazard are recommended. In the “red” range, there is a high risk and measures to eliminate the hazard are urgently required. Table 1 shows the angle ranges for the six movements evaluated in this study.

**Table 1.** Angle ranges for the six movements evaluated in this study.

Movement	Orange area	Red area
Lower back (rotation)	-20° to -10°	>20°
	10° to 20°	<-20°
Back (flexion/extension)	20° to 60°	< 0°
		> 60°
Shoulder right (flexion/extension)	20° to 60°	< 0°
		> 60°
Shoulder right (abduction/adduction)	-20° to -60°	< -60°
		> 0°
Shoulder left (flexion/extension)	20° to 60°	< 0°
		> 60°
Shoulder left (abduction/adduction)	-20° to -60°	< -60°
		> 0°

As an example, Table 2 shows the result of the comparison of the real executions for the three load cases with the virtual execution in step 3 “move box”. In this type of evaluation, it was only counted whether a test participant was in the orange and/or red area during the exercise. The numbers therefore represent the number of test participants who experienced a medium or higher level of strain during this activity. In terms of ergotyping, it would be these activities that could be considered for optimization.

**Table 2.** Number of test participants for the two areas orange and red in the four executions.

Movement	Light		Medium		Heavy		VR	
	Orange	Red	Orange	Red	Orange	Red	Orange	Red
Lower back (rotation)	4	3	4	3	4	2	4	4
Back (flexion/extension)	4	4	3	4	4	4	4	4
Shoulder right (flexion/extension)	4	4	4	4	4	4	4	4
Shoulder right (abduction/adduction)	4	4	4	4	4	4	4	4
Shoulder left (flexion/extension)	4	4	4	4	4	4	4	4
Shoulder left (abduction/adduction)	4	4	4	4	4	4	4	4

If Table 2 is viewed in terms of signal detection theory, the simulation in VR would produce the same result for the “move” (step 3) activity as the real execution for certain movements. Even for different load weights. The geometric conditions of the standardized workstation would therefore have the same effect on the considered body angles of e.g. the shoulder in the real and virtual situation. The rotation of the lower back on the other hand would be overestimated. The VR simulation would produce more false alarms regarding this movement. In our small study, it was even the case that the heavier the weight, the smaller the range of lower back rotation was. Lower back rotation in VR and with light weight in the real execution were wider and therefore more often found in the risky areas. The results for

flexion and extension of the back were again comparable, although one test subject only had movements in the red area and there was no orange area.

## CONCLUSION

The results of this first small study give us hope that it is possible to assess load handling using motion tracking in a virtual simulation. In our case of the defined standard procedure, risky body movements occurred more frequently in the virtual simulation than in the real situations. The aim is not to measure body movements exactly. The method would be successful for early assessment in terms of ergotyping if it can be used to identify neuralgic points of a work system. Nevertheless, further studies are to follow in which the haptic immersion in the sense of the Immersion Cube is to be increased in order to bring the movements in the virtual simulation even closer to those in reality. The interesting trade-off here will be to be able to create a sufficiently accurate simulation with as little effort as possible. In terms of quality, when comparing the movements in reality with the movements in the virtual simulation, we were able to observe that a simulation like the one we carried out without haptic immersion would not have been sufficiently accurate for some activities. For example, the virtual crates were held overhead for longer in the transition from step 4 to 5 and even made a body turn in this position, whereas the real weights were brought directly close to the body at hip height when they left the table. This is also partly consistent with the findings of Kačerová et al. (2022) who observed that subjects spent longer in shoulder flexions  $>60^\circ$  in VR execution than in real execution.

For this reason, haptic immersion is to be successively increased in further studies. Initially in the sense of H1 by means of pseudo-haptic feedback as control display ratio. Subsequently, indirect haptic feedback will be used to investigate whether this can increase the awareness of a load and bring the movements in VR closer to reality. The plan is to use vibrotactile wristbands (H2) here. In a further step, a model for direct haptics will be developed (H3), which will act as a load proxy and enable adaptive weight adjustment.

In addition, deficiencies from the first study are to be rectified in the follow-up studies. Carrying loads over 5m will also be included as a further activity. Since the controllers caused artifacts in the motion tracking due to their condition, the following studies will be set up using hand tracking, which will also ensure a more realistic gripping situation. Initial attempts to work with haptic gloves were not promising at this point. A further test for the method will be to gradually dissolve the standard procedure and move on to free forms of execution in order not to exclude individual differences in execution.

With regard to the equipment used so far, we can say that we have had good experiences with the Quest Pro stand-alone headset for this type of study. Due to the design of the headset, the test participants were able to see the real environment and their bodies in the peripheral field of view, which meant that the movements were more realistic than in the first test runs of our study with a completely closed headset (Quest 3).

## REFERENCES

- Bevan, S., (2015). Economic impact of musculoskeletal disorders (MSDs) on work in Europe. *Best Practice & Research Clinical Rheumatology*. 29.

- Briggs, A. M., Woolf, A. D., Dreinhöfer, K., Homb, N., Hoy, D. G., Kopansky-Giles, D., Åkesson, K. and March, L., 2018. Reducing the global burden of musculoskeletal conditions. *Bulletin of the World Health Organization*, 96(5), pp. 366–368.
- Bullinger-Hoffmann, A. C. and Mühlstedt, J., 2017. *Homo Sapiens Digitalis - Virtuelle Ergonomie und digitale Menschmodelle*. Berlin Heidelberg New York: Springer-Verlag.
- Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, 2023. *Sicherheit und Gesundheit bei der Arbeit - Berichtsjahr 2022. Unfallverhütungsbericht Arbeit*. 1. Auflage. Dortmund: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. ISBN: 978-3-88261-761-0
- Caeiro-Rodríguez, M., Otero-González, I., Mikic-Fonte, F. A. and Llamas-Nistal, M., 2021. A Systematic Review of Commercial Smart Gloves: Current Status and Applications. *Sensors*, 21(8), p. 2667.
- Caputo, F., Greco, A., D'Amato, E., Notaro, I. and Spada, S., 2018. On the use of Virtual Reality for a human-centered workplace design. *Procedia Structural Integrity*, 8, pp. 297–308.
- Dinh, H. Q., Walker, N., Hodges, L. F., Song, C. and Kobayashi, A., 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In: *Proceedings of IEEE Virtual Reality*. pp. 222–228.
- Gallace, A., Ngo, M. K., Sulaitis, J. and Spence, C., 2012. Multisensory presence in virtual reality: Possibilities & limitations. In: *Multiple sensorial media advances and applications: New developments in MulSeMedia*. IGI Global, pp. 1–38.
- Kačerová, I., Kubr, J., Hořejší, P. and Kleinová, J., 2022. Ergonomic Design of a Workplace Using Virtual Reality and a Motion Capture Suit. *Applied Sciences*, 12(4), p. 2150.
- Kamusella, C. und Schmauder, M., 2010. *Ergotyping im rechnerunterstützten Entwicklungs- und Gestaltungsprozess*. München: GRIN Verlag.
- Lim, W. N., Yap, K. M., Lee, Y., Wee, C. and Yen, C., 2021. A Systematic Review of Weight Perception in Virtual Reality: Challenges and Road Ahead. *IEEE Access*
- Månsson J., 2018. *Using a Virtual Fire Extinguisher as a Tool for Safety Training*. Ph. D. Master Thesis. Lund: Faculty of Engineering, Lund University.
- Milgram, P. and Kishino, F., 1994. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12), pp. 1321–1329.
- Rybníkář, F., Kačerová, I., Hořejší, P. and Šimon, M., 2023. Ergonomics Evaluation Using Motion Capture Technology—Literature Review. *Applied Sciences*, 13(1), p. 162.
- Samad, M., Gatti, E., Hermes, A., Benko, H. and Parise, C., 2019. Pseudo-Haptic Weight: Changing the Perceived Weight of Virtual Objects By Manipulating Control-Display Ratio. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland, 4–9 May 2019. New York: ACM, p. 13.
- Serafin, P., Klußmann, A., Liebers, F., Schust, M., Brandstädt, F., Hartmann, B., and Gebhardt, H., 2020. Die neue Leitmerkmalmethode Manuelles Heben, Halten und Tragen von Lasten (LMM-HHT): Darstellung der Methode und Ergebnisse der Methodentestung, GfA Frühjahrskongress, Berlin.
- Simonetto, M., Arena, S. and Peron, M., 2022. A methodological framework to integrate motion capture system and virtual reality for assembly system 4.0 workplace design. *Safety Science*, 146.

# Combining System Dynamics and Agent-Based Simulation to Evaluate and Visualise Sustainable Airport Operations

Florian Rudolph<sup>1</sup>, Martin Jung<sup>2</sup>, and Axel B. Classen<sup>2</sup>

<sup>1</sup>German Aerospace Center, Lilienthalplatz 7, 38108 Braunschweig, Germany

<sup>2</sup>German Aerospace Center, Linder Höhe, 51147 Cologne, Germany

## ABSTRACT

Facing the backdrop of the climate crisis, we are currently witnessing an intense transformation process in aviation. Aim of this process is a climate-friendly air transport system. In addition to aircraft manufacturers and airlines, airports also must contribute to this transition by improving their operations. A key objective of airport management is therefore to foster climate-neutral aviation and energy-efficient airport operations. European airports are committed to achieving these goals by 2050. An important contribution to these objectives is to enable airport operators to draw informed operational decisions while balancing traffic impacts with economic and environmental aspects. Therefore, we present a concept for combining different simulation techniques resulting in a comprehensive, integrated hybrid simulation model and visualization tool. For this purpose, we combine two simulation techniques, namely an agent-based network simulation which implements the Airport Collaborative Decision-Making Concept (A-CDM) for joint decision-making at airports via state charts as discrete events with a flow simulation based on system dynamics.

**Keywords:** A-CDM, Climate-neutral aviation, Hybrid simulation, Holistic airport management, System dynamics, Agent based, Ecologic, Economic, Key performance indicators, Situational awareness

## INTRODUCTION

Airports provide a complex infrastructure and many different players are active at the airport. We therefore view an airport as a holistic system with various operational areas and stakeholders in which an overarching and coordinated management allows for targeted prioritisation, e.g. of sustainability parameters. Consequently, there are many places where energy is used and CO<sub>2</sub> is emitted.

In air traffic, standstill consumes resources and energy. At airports, this statement is not a contradiction in terms, as an aircraft waiting on the ground for the next process requires energy to keep engines, electronics and air conditioning running. The quicker an aircraft is able to take off again, the less energy it consumes at the airport. Therefore, a smooth flight is an energy-efficient flight. The formula sounds simple at first, but the forces behind it are complex process and coordination chains between all the system partners

involved in flight management. This applies to the day-to-day interaction between airport operators, airlines, ground handling companies and air traffic control.

### Airport Collaborative Decision Making

Airport Collaborative Decision Making (A-CDM) (EUROCONTROL, 2017) is an operating principle that can help reduce energy consumption at the airport. The aim of A-CDM is to increase operational efficiency of airports by making aircraft turnaround processes faster and more predictable. This is achieved by the stakeholders at the airport (airport operators, airlines, ground handlers and air traffic control) and the network manager working together transparently and cooperatively in operations and exchanging relevant, accurate and timely information. A-CDM focuses on a series of selected milestones along a flight (arrival, landing, taxi in, turnaround, taxi out and take-off) where the partners involved change. This approach to the turnaround process tracks the progress of a flight through a continuous sequence of different events, called milestones. Rules are set for updating downstream information and the accuracy of estimates. Different Airport-CDM partners may be responsible for different milestones, with the aim of integrating all milestones into one common seamless process for each flight. The main objective of the milestone approach is to further improve the common situational awareness of all partners during all turnaround phases of each flight. An overview of the milestones relevant to the turnaround process is provided in Figure 1.

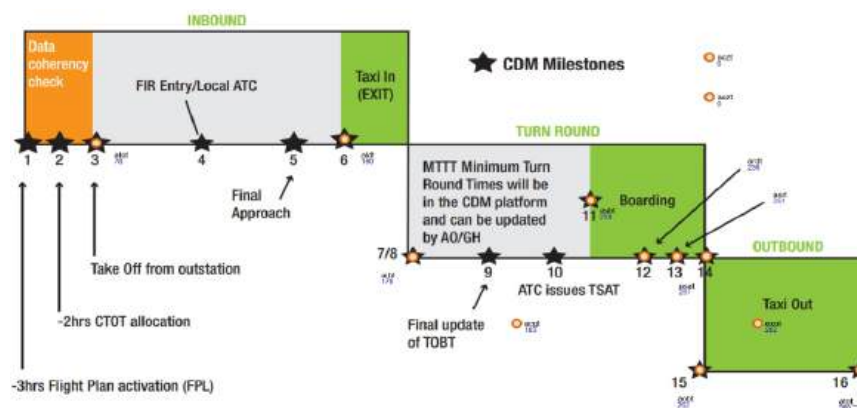


Figure 1: A-CDM-milestones.

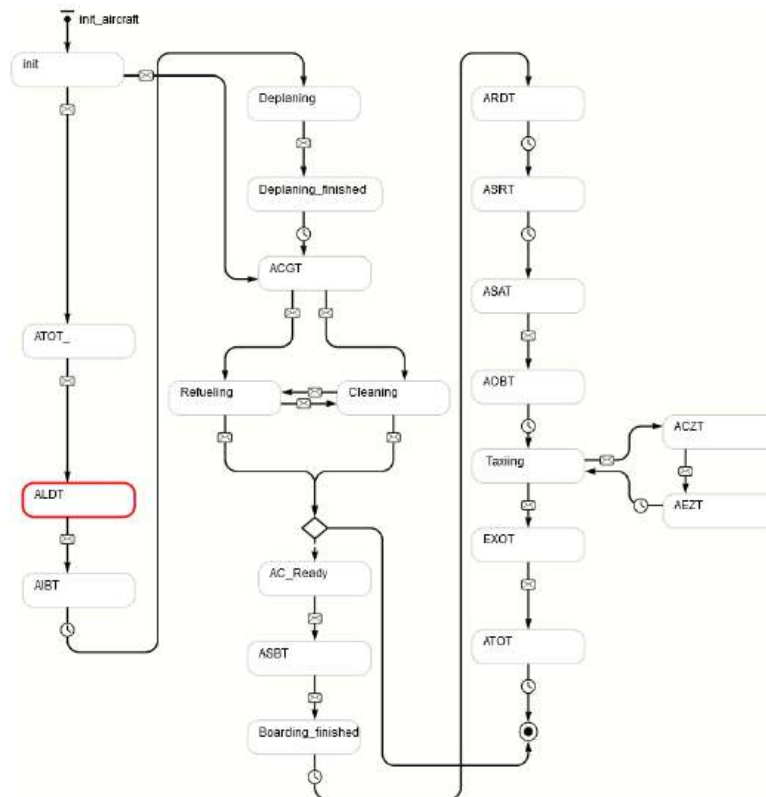
### Agent-Based Simulation of A-CDM Milestones

Our approach combines two different simulation paradigms into a coupled simulation environment. This chapter describes the agent-based part of the simulation. The System Dynamics-based part is then described in the next chapter.

The task of the agent-based simulation is to map the A-CDM milestones on the basis of predicted scenarios and to evaluate them in terms of joint

decision-making by calculating key performance indicators (KPIs) such as punctuality, connectivity or throughput from the milestones (Schier et al., 2016). This agent-based simulation is complemented by a system dynamics simulation to map the energy consumption in the scenarios. The aim is to develop new indicators for evaluating energy efficiency.

The A-CDM simulation provides a continuously changing operational forecast of A-CDM milestones and maps them with modelling methods from software engineering, in this case as a state chart diagram as part of the Unified Modelling Language (UML) (Fowler, 2003). This is a general-purpose visual modelling language and is intended to provide a standard way to visualize the design of a system. A state chart diagram shows the states of a state machine that are permitted at runtime and specifies events that trigger its state transitions. A state chart diagram thus describes a hypothetical machine that is in exactly one state of a finite set of states at any given time. This automaton consists of states and transitions (Harel, 1986). In our case, each milestone is represented by a state in the diagram.



**Figure 2:** A-CDM milestones as state charts.

This kind of Modelling allows us to specifically map the dependencies between milestones. The processes in and around the aircraft during turnaround are implemented using a network model. This means that the topology of the airport is not considered. The various dependencies of

turnaround and terminal processes are implemented via queues and process times. The relevant parameters to describe the airport are stored in an external data structure. They are read in at the start of the simulation. Those parameters include the flight plan, runway properties, aircraft stands, number of available ground handler teams, etc.. Based on the flight plan, the aircraft are created as simulation agents within the simulation upon approach or on the apron and move through the queue and process network until they leave the milestone simulation with their departure. Acting as a simulation agent, each aircraft has certain initialization parameters and individual control of the state chart diagram model for the aircraft's A-CDM milestones (Figure 2).

The individual states of the aircraft are controlled via messages. If the aircraft reaches a certain process in the network simulation, a message is sent to the simulation agent. This triggers the transition to a new state. For example, if the aircraft has completed the landing process, the achievement of this milestone is attributed to the aircraft as an ALDT (Actual Landing Time) message. When this message is received, the aircraft is set to the corresponding status and the timestamp is saved and stored in an external data structure.

A diagram is generated for each aircraft to be handled (Figure 2) and the current process status is displayed. In this case, ALDT (Actual Landing Time) is the current status (marked by a red border). The status changes are triggered by a network simulation via messages. When an aircraft reaches a certain node in the network, it receives the message for the status change. If a milestone is reached as a status, this is stored in the object memory and can be made available to the management system from there.

The network model provides the geographical path-time interrelations as well as the queue processing at the corresponding neuralgic points where only sequential processing of aircraft is possible. In our simulation, the processing of the route network is separated from the queueing logic. The logic and interaction of the agents (in this case aircraft) is implemented with blocks of a modelling language, while the distances are read and recorded directly from the topology of the airport using a geo-information system. This combination of state charts and network simulation thus provides a digital twin for determining the A-CDM milestones. These milestones are generated for each simulated flight, broadcasted and stored during runtime and are immediately available to the management system. The following times are obtained as an example data set for a flight at the day of operations:

**Table 1.** Output of A-CDM milestones for each flight (example).

simtime	Id	Registration	Type	callsign_a	callsign_d	sobt
435	497	LZEHD	both	AFR211	AFR212	06:54
sibt	atot	aldt	aibt	acgt	asbt	ardt
05:56	04:12	05:42	05:56	06:01	06:20	06:39
asat	aobt	aczt	aezt	atot	delay	dur_tax
06:57	06:58	00:00	00:00	07:13	4	15



Our simulation environment and input data is based on an exemplary day of operations, the topology, and capacity utilization of a medium sized European international airport in the Mediterranean region. In total the air-traffic scenario runs from 0:00 to 23:59 and comprises 904 flights (457 arrivals and 447 departures with approx. 110.000 Passengers). We aggregated the scenario inputs and parameters in an Excel table from where it is dynamically fed into the simulation.

### Energy Flow Simulation in System-Dynamics

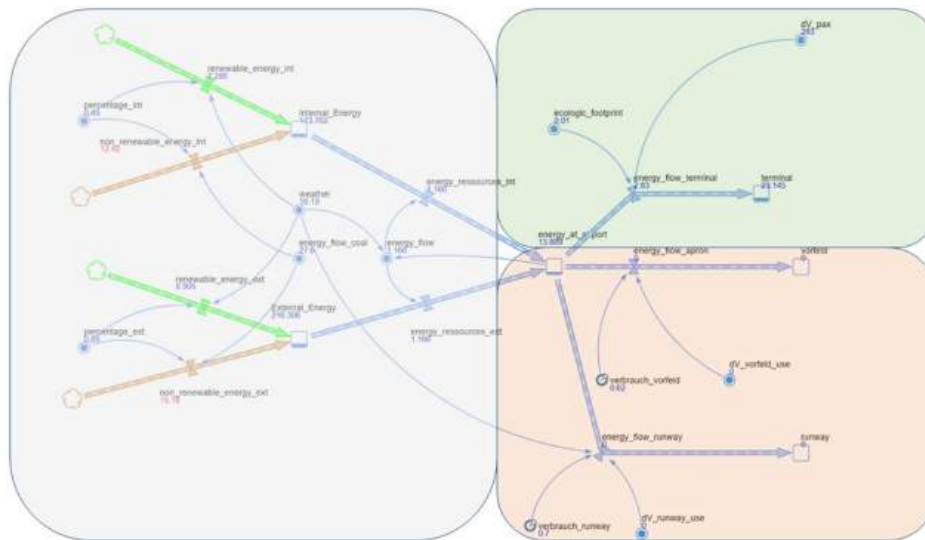
To analyse the energy flow at an airport, we use the paradigm of system dynamics simulation. System Dynamics (SD) is a methodology for modelling and simulating complex systems, particularly in the context of energy consumption. SD is a tool to model and understand nonlinear behaviour of complex systems over time. It uses and visualises stocks, flows, internal feedback loops, table functions and time delays to model a system. System Dynamics was developed back in the 1950s by J. W. Forrester (Forrester, 1990). As an application model of economic cybernetics, SD is used today particularly in the fields of economics and business administration to analyze dynamic and complex issues. Examples come from the public and private sectors: production management, strategic planning, analysis and design of business models, business forecasting and scenario analysis.

In our case we use SD to get a holistic view of the airport by considering interconnected components, feedback loops, and dynamic behaviour of all airport parts and stakeholders to visualize causal relationships. For modelling and simulating we use the simulation software Anylogic. Anylogic is a multi-methods simulation software supporting system dynamic, discrete events and agent-based modelling. It is also capable of mixing these simulation methods within one model (Anylogic, 2024).

In this paper we focus on a first simulation prototype to examine the technical feasibility, its possibilities and limitations. From a technical point of view, we focused on combining SD and Agent-Based Modelling (see next chapter). For this, we first modelled the main energy flow at an airport with SD. Later in the project, the aim is to simulate the interaction of the different stakeholders and to model the environmental and (macro) economic impacts, as well as the possibility of different scenarios with SD.

Figure 3 shows a screenshot of the SD model that we are building in Anylogic. The model is divided into three main areas. The cluster on the left represents the energy input, divided into internal and external energy sources, as well as the subdivision into renewable and non-renewable energy and its dependencies, such as weather. One of the consumers on the top right is the sub-simulation of the terminal building. The results of this sub-simulation serve as input for the SD simulation (see description of *dV\_pax* in next section). As a hypothesis, we assume that energy consumption is a function of the number of passengers using the building, in addition to a base load. The arrival distribution of passengers per flight is based on passenger survey data and historical observed patterns that set up the inflow for the Terminal part of the model (Alers et al., 2013). The terminal outflow

is combined with the agent-based simulation. Each time an aircraft is in the status of “boarding completed”, the number of booked passengers are “leaving” (subtracted from) the terminal. To take account of environmental improvements of the terminal building, we include an “Ecologic Footprint Factor” in our simulation model chain. This will enable us to run different scenarios with varying environmental improvements.



**Figure 3:** System dynamics energy flow.

The bottom right of Figure 3 represents the simulation sub-model of the apron area and runway system, which represents the second part of the energy consumers. Our initial hypothesis here is that energy consumption depends on the use by airplanes, again in addition to a base load of energy consumption for the basic provision of airside infrastructure.

### Interconnection of System-Dynamics and Agent-Based Simulation

The combination of System Dynamics and agent-based simulation allows us to understand both the macroscopic flows and the microscopic interactions at the airport. In addition, agent-based simulation can help to assess the impact of fine-grained changes in airport layout or operations by simulating how individual agents - in this case passengers, ground staff and aircraft – react to these changes. On the other hand, System Dynamics can be used to analyse overarching trends and patterns.

The challenge in this modelling lies in the interface between the agent-based network simulation and the energy flow simulation implemented in System Dynamics. Ultimately, it should be possible to read out the resource utilization via the processes that the simulation agents (i.e. airplanes and passengers) pass through in order to determine the required energy consumption. For this purpose, the number of agents that are currently in

defined processes in the network simulation is determined and imported into the SD simulation as a dynamic variable during runtime. As a dynamic variable, this value has a direct impact on the simulated (energy) flow and thus on the stock levels. Therefore, we use the three variables “*dV\_runway\_use*” for utilization of the runway, “*dV\_apron\_use*” for the number of aircraft using the apron and “*dV\_pax*” for the number of passengers using the terminal building.

### Initial Results

As a first approximation, we used the configuration and real data based on an exemplary day of operations of an international European airport in the Mediterranean region for the simulation. In order to be able to further develop an assessment tool to measure the congruence between resource utilization, A-CDM milestones and energy consumption.



**Figure 4:** Energy flow at the airport.

Figure 4 shows a conceptual representation of energy consumption and available energy at an airport over time, expressed in minutes on the x-axis. As in the system dynamics model, consumers are differentiated into runways (grey), apron (light-green) and terminal building. Energy available at the airport is shown in black. It is easy to see how this storage behaves in the opposite direction to consumption. When consumption is high, the storage buffer sinks and when consumption is low, it fills up. So, the consumption is directly dependent on the utilization of each area and the aircraft, vehicles and passengers using them.

Since the application runtime of the scenario under consideration is only a few seconds, this coupled simulation environment can be used as a forecast simulation during regular airport operations. Current consumption values are just being recorded as part of a project, which means that reliable values can only be forecast once the quality of the input data has been established and validated. In the current phase, it is mainly important for us to quantify the relationship between utilization and consumption.

## CONCLUSION

In this paper we investigated the potential of combining System Dynamics and agent-based simulation modelling to provide a toolbox for airport operators. By adjusting model parameters, policies, and external influences, we can evaluate the impact of various interventions on energy consumption.

The combination of these two modelling paradigms in a hybrid simulation enables us to model the deduction of energy consumption in relation to resource utilization. With state charts in the agent-based sub-simulation and the direct visualization of System Dynamics, users receive a traceable overview of airport processes and a holistic view of the system's behavior. This also allows what-if studies to be conducted on how consumption will change under different conditions or future operational procedures or to evaluate technologies.

Thanks to the short application runtime of just a few seconds, our simulation can also be used for operational forecasting. It provides the ability to aggregate individual flight-related information to a level that allows airport stakeholders to assess demand, capacity utilization, and performance parameters, such as punctuality or delays as well as resulting energy flows. In addition, the forecast can be used to implement possible countermeasures (e.g. runway closure for a certain period of time) and assess their impact to resolve potential disruptions. This can help to improve energy-efficiency of airport operations.

## REFERENCES

- Alers, T., Classen, A. B., Kellner, S. (2013) More Input - Generic Data for Microscopic Airport Simulation, AIAA Aviation, Los Angeles.
- AnyLogic – Software tools for simulation modelling & business solutions <https://www.anylogic.de/> (Stand 4.2024).
- EUROCONTROL Airport CDM Team (2017) Airport Collaborative Decision-Making (A-CDM) Implementation Manual.
- Forrester, J. W. (1990) Principles of Systems - System dynamics series. ISBN 9780915299874 <https://books.google.de/books?id=oMASAQAAMAAJ>
- Fowler, M. (2003) UML Distilled: A Brief Guide to the Standard Object Modeling Language ISBN 9780321193681.
- Harel, D (1986) Statecharts: A Visual Formalism for Complex Systems. Science of Computer Programming 8.
- Schier, S., Timmermann, F., Pett, T. (2016). Airport Management in the Box – a human-in-the-loop simulation for ACDM and airport management. Deutscher Luft- und Raumfahrt Kongress 2016, Braunschweig, Germany.

# The Construction of Egocentric and Allocentric Spatial Representations in Visual-Spatial Working Memory in Highly Immersive Virtual Reality (CAVE)

Olga A. Saveleva<sup>1,2</sup>, Boris B. Velichkovsky<sup>2,3</sup>,  
Galina Ya Menshikova<sup>2</sup>, and Grigory S. Bugriy<sup>2</sup>

<sup>1</sup>I.M. Sechenov First Moscow State Medical University, Moscow, Russia

<sup>2</sup>Moscow State University (Lomonosov MSU), Moscow, Russia

<sup>3</sup>Moscow State Linguistic University, Moscow, Russia

## ABSTRACT

In solving spatial tasks, neurocognitive egocentric and allocentric spatial representations storing in the visual-spatial working memory. Egocentric representations encode the visual scenes in self-centered coordinates and allocentric representations – in world coordinates regardless of the observer's position. Previously studies showed a good consistency in spatial processing about real environments compared to virtual reality environments. A presentation method was developed for memorizing and reconstructing 3D scenes using the highest immersive CAVE virtual reality system. A space for task, library of objects and virtual scenes were designed, each containing seven virtual objects located in different 3D positions. Three viewpoints were given for reproduction: «the front» viewpoint (to reproduce the memorized scene from the imaginary egocentric position), «the left» and «the above» viewpoints (to reproduce the scene from the left or above imaginary allocentric positions, respectively). The participant had to reconstructed memorized scene in a natural way by choosing objects from the library and placed it in virtual space in accordance with the given imagine viewpoint. The score of object localizations was estimated separately by three parameters – topology, metrics, and depth. The results showed, that for both types of spatial representations schematic topological properties were preserved better in visual-spatial working memory than the exact metric information (especially for the egocentric representations). Overall, the egocentric representations were more effective in the reconstruction of 3D scenes than allocentric representations. It was also found that when using an allocentric representations, the need to add a height axis (vertical rotation) diminishes the effectiveness of the scene reconstruction from visual-spatial working memory, compared to rotations in the horizontal plane. The results suggest that both egocentric representations and allocentric representations can be formed in visual-spatial working memory, but that egocentric representations are more basic in the solution of spatial tasks using visual-spatial working memory. These results not only have theoretical significance in cognitive psychology, but also have the potential for wide practical application in healthcare, education, developmental and sports psychology, human factor research and related interdisciplinary fields.

**Keywords:** Egocentric spatial representations, Allocentric spatial representations, Spatial memory, Visual-spatial working memory, Topology, Metric, Depth, Immersive virtual reality

## **INTRODUCTION**

### **Spatial Representations (SRs)**

Modern approaches for studying spatial cognition using the concept of spatial representations (SRs), usually understood as cognitive units of spatial cognition (Menshikova et al., 2020). Also, very often the concept of spatial representations using synonymously with the concept of processing strategies, cognitive systems, frames, etc. Two types of SRs are identified in the literature: egocentric representations (ESRs) and allocentric representations (ASRs) (Colombo et al., 2017; Derby, 2021; Klatzky, 1998; Serino et al., 2014). The dissociation of cognitive systems that ensure the formation of egocentric and allocentric spatial representations has been proven in many studies. ESRs encode object locations relative to the observer and ASRs encode object locations by specifying the relative positions between objects. The two types of spatial representations are related in a complex fashion, ASRs being considered more abstract than ESRs (Klatzky, 1998). Both ESRs and ASRs are involved in the solution of many spatial tasks including wayfinding and spatial memory. An important topic in ESR/ASR research is the study of their cognitive and neurocognitive mechanisms. This is the more important as deficits in ESRs and, especially, ASRs have been shown in a variety of conditions including old age and Alzheimer's Disease (Colombo et al., 2017; Serino et al., 2014; Tuena, 2021), autism (Presley, 2021), and Williams syndrome (Broadbent, 2014).

Current research has recently suggested the involvement of hippocampus and surrounding areas in the formation of SRs (Danjo, 2020; Fidalgo, Martin, 2016). This may be well related to the findings that hippocampus is heavily involved in spatial processing in both animals and humans (Bird, 2008). ESR/ASR formation has also been shown to be associated with the parietal lobes which are related to attention management and multimodal sensory integration (Iachini et al., 2009). This indicates that a complex network of anatomical structures and cognitive processes supports the establishment of ESR/ASR while solving complex spatial tasks (Danjo, 2020). These may include (spatial) memory processes as indicated by the involvement of the hippocampus. In this research, we would like to extend the research on the neurocognitive underpinnings of ESR/ASR by the study of the role of working memory in their formation.

This study is interdisciplinary. The results are very important in healthcare (especially in children and involutionary age), education, medical rehabilitation of cognitive and neurobiological disorders, sports psychology, as well as the study of human factors and the development of artificial intelligence systems used in cognitive ergonomics. Well known systematic review point out in conclusion that the study of cognitive deficits based on cognitive concepts of ESR/ASR might be crucial to make accurate diagnosis and rehabilitation (Tuena, 2021).

### **Spatial Representations (SRs) of Virtual Reality (VR) Environments**

In the modern digital world, it is especially important to study the psychological processes of processing spatial information not only in the

real environment, but also in a virtual reality. Highly immersive technologies allow to conduct cognitive experiments in unique conditions that are difficult to create using standard laboratory methods, for example, to study the construction of ESR/ASR at the earliest stages of their formation.

Among the methods for studying spatial representations, the method using virtual reality (VR) technologies is increasingly being used. Despite subtle differences, real-world and virtual versions showed good overlap for the assessment of spatial memory even in clinical subjects, not just healthy human subjects (Tuena, 2021).

On the other hand, in the digitalizing modern world it is very important to study the processing of spatial representations in virtual reality. We used the CAVE virtual reality system, which provides the highest immersive conditions for conducting a cognitive experiment in working memory methodology.

### **Visual-Spatial Working Memory (VSWM) and SRs**

Working memory (WM) is a hypothetical cognitive system involved in the temporary storage and processing of information (Velichkovsky, 2017; Baddeley, 2012). Within WM, a visual-spatial component (VSWM) has been identified, among others. VSWM is used for processing of spatial and (independently) of visual information and is heavily involved in solving spatial tasks (McAfoose, Baune, 2009). It can be assumed that VSWM should be involved in processing of SRs. Indeed, recent research has shown a correlation between individual differences in VSWM and the effectiveness in SRs encoding (Ishikawa, 2023; Wen et al., 2013). Below, we report on a study about how information about 3D scenes is preserved and processed in VSWM over short-periods of time when a 3D scene reconstruction task is performed involving an ESR (no need for a perspective change, reconstructing the scene “from the front”) or an ASR (with a need for perspective change, reconstruction the scene “from the left” or “from above”).

## **MATERIALS AND METHODS**

### **Participants**

Thirty nine volunteers (21F, 18M, age range  $22\pm 3$  years) took part in the experiment. All subjects had normal or corrected to normal vision and had no vestibular dysfunctions or injuries.

### **Equipment**

We used the CAVE system Barco Ispace 4.0, which consisted of four large flat screens (each  $2.5\times 2.5$  m in size) combined to form a cube. Active eyewear CrystalEyes 3 Stereographics was used to create stereo effects. The manipulations of virtual objects were carried out using a Flystick 2. The A.R.T. DTrack 2 tracking system was used to record the participant’s positions with update rate of 8 Hz (see Figure 1). The virtual scenes were developed with the use of VirTools 4.0 software.



**Figure 1:** CAVE virtual reality system and the typical scene view (in the upper right corner).

### Stimuli

Six virtual scenes were designed, each containing 7 virtual objects located in different 3D positions (see Figure 1). All objects were organized in a virtual volume of  $(20^\circ \times 20^\circ \times 7^\circ)$ , the angular size of each object was about  $2^\circ \times 2^\circ$ . The average brightness of the objects varied slightly within 15 – 20 Cd/m<sup>2</sup>. The participants were recommended to keep constant position (2.3 m) relative to the virtual scene to control angular sizes of the virtual objects. The objects were presented against a background which represented a dark 3D space, in which a lot of small white balls ( $0.5^\circ \times 0.5^\circ \times 0.5^\circ$ ) were uniformly distributed. Their density in the space was about 6 cpd.

### Procedure

Each of six virtual scenes was presented 3 times for 25 seconds. The order of scenes presentation was quasi-random. The participants were asked to remember the objects and their location and then to reproduce the memorized scene. The algorithm of reproduction was as follows. Immediately after the scene presentation an arrow was shown for 3 seconds. Its orientation showed from which point of view the participant had to imagine and reproduce a 3D scene that he had just seen. Three viewpoints were given for reproduction: «the front» viewpoint (to reproduce the memorized scene from the imaginary egocentric position), «the left» and «the above» viewpoints (to reproduce the scene from the left or above imaginary allocentric positions, respectively). Then, the library consisting of 21 familiar and unfamiliar objects was shown. The participant had to choose objects from it and placed them in virtual space in accordance with the given point of view. The object locations in the virtual space were recorded. Each scene was reproduced using each of three points of view. A series of training exercises had been carried out before the experiment to allow participants to get acquainted with manipulations of virtual objects. The average time required to reproduce a scene was about 20–30 sec. The training exercises and the main experiment took 20–25 minutes.



## Data Processing

The score of scenes reproduction was estimated as the sample-average of the number of correctly reproduced virtual objects, separately for each mental viewpoint position. The score of object localizations was estimated separately by three parameters—topology, metrics, and depth.

## RESULTS

The accuracy of spatial memory was assessed by computing the probability of objects being correctly identified during 3D scene reconstruction (averaged over probes, subjects, and experimental conditions, see Table 1). The scores were very high meaning that, generally, around 6–7 visual objects were successfully held in subjects' VSWM. This is well in line with the magical number of 7 plus/minus 2 as identified by Miller (1965). However, this is distinctively over the 4 elements limit of VSWM identified by (Luck & Vogel, 1997) which may be related by an absence of an addition attention-demanding task as employed in the Luck & Vogel study. Overall, we think the using 7 objects in our experiment didn't exceed typical limitations of VSWM and we obtained results pertaining to visual scene processing in VSWM without the need for recruiting additional storage mechanisms as was demonstrated for verbal WM (Oberauer, 2002). Importantly, the accuracy of object identification didn't differ significantly between experimental conditions (all  $p > 0.1$ ).

**Table 1.** The accuracy of SRs from all imaginary viewpoints.

Variable	Front Viewpoint (ESRs)	Left Viewpoint (ASRs)	Above Viewpoint (ASRs)	F-test	F-value	p-value
The accuracy of objects identification	0,94(0,09)	0,90(0,11)	0,93(0,09)	0,12	2,14	0,58
The accuracy of topology	0,88(0,10)	0,81(0,15)	0,67(0,24)	0,01	--	--
The accuracy of metric	0,48(0,23)	0,41(0,20)	0,39(0,23)	0,19	1,71	0,36
The accuracy of depth	0,54(0,15)	0,39(0,18)	0,31(0,14)	0,01	--	--

The analysis of how the topological properties of the remembered visual scenes were reconstructed revealed some differences between experimental conditions. Topological properties were not preserved if the relative location in any pairs of objects was wrongly reproduced. The number of wrongly reproduced relative locations was averaged over probes, subjects, and experimental conditions. These averages were submitted to pairwise t-tests. The accuracy of topology (probabilities of preserving the topology of the remembered visual scene) are presented in Table 1. It can be seen that topology is well preserved for ESR (about 90% correct in the front condition), but not in both ASR conditions (81% correct in the left condition and 67% correct in the above condition). T-test revealed significant differences between all experimental conditions ( $\llcorner\text{front}\llcorner-\llcorner\text{left}\llcorner$ ,  $t(38) = 3,88$ ,

$p < 0,01$ ; «left»–«above»,  $t(38) = 3,15$ ,  $p < 0,01$ , «front»–«above»,  $t(38) = 5,33$ ;  $p < 0,01$ ). These results suggest a gradual decrease in the accuracy of topological information in ASRs as they require more mental rotation.

The accuracy of metric properties in VSWM was also analyzed for different types of SR. To this end, we computed the averaged probabilities of there being no metric errors in the visual scene reconstruction. A metric error was coded as the deviation of an objects position to more than 20% from its original position. The averages for the probability of metrically correct reconstructions are given in Table 1 and were also submitted to a series of pairwise t-test. It can be seen from Table 1 that metric accuracy in VSWM is much lower than topological accuracy and that the results don't differ much between SRs (if anything, there is a tendency for the ESR to be more slightly effective in this respect than both ASR). T-tests support this conclusion and reveal no significant differences between experimental conditions (all  $ps > 0.1$ ). However, the standard deviations reported in Table 1 for metric accuracies are relatively high (around 0.2, which mean the individual accuracies vary on average as much as from 20% to 60%) which suggest relatively high variability of metric accuracy across probes and subjects. The last point suggests a possible influence of individual differences in the spatial abilities.

Finally, we assessed ESR/ASR differences in preserving depth relations. To this end, we computed the relative number of depth errors. Depth errors were identified as false ordering of the reconstructed objects along the depth axis (Z-axis in the front condition, X-axis in the left condition, and Y-axis in the above condition). The average probabilities of depth reconstruction accuracy are presented in Table 1. It can be seen that the depth-related accuracy is generally relatively low (around 50%-40%). It is highest in the ESR front condition (54%) with a gradual decrease of the both ASR conditions (left – 39%, above – 32%). Pairwise t-test comparisons revealed that all experimental conditions differ significantly with respect to depth-related accuracy (“front”–“left”,  $t(38)=3,587$ ,  $p = 0,01$ ; “left”–“above”,  $t(38)=2,32$ ,  $p = 0,03$ ; “front”–“above”,  $t(38)=6,02$ ,  $p = 0,01$ ). As maintaining depth relationships are related to the maintenance of metric information, these data suggest that VSWM is not very good in keeping exact spatial information about the perceived visual scenes. They also suggest that the depth information is best reconstructed based on an ESRs.

Differences in spatial abilities are consistently reported for men and women (Chen et al., 2020; Conrad, Hull, 1964; Hedges, Nowell, 1995), but see, for example, (Bartlett, Camba, 2023; Self et al., 1992) for a critical review of gender differences in spatial abilities. Given that we had a similar number of man and women in our study (21F, 18M), we also performed a brief check of possible gender differences in our data. No gender differences were found for the recognition accuracy. This conclusion is supported by a t-test comparison ( $p > 0.1$ ). Also, no differences were found neither for topological accuracy, nor for metric accuracy, nor for depth-related accuracy (all  $ps > 0.1$ ). These results strongly suggest that the formation of SRs in VSWM is a basic cognitive process not dependent on the specifics of the male or female

brain (e.g., hemispheric asymmetries). This is not surprising given that the ecological and evolutionary requirement in representing basic relationships within the spatial environment are most likely shared for man and women. Our data shows that there are no gender differences in the basic processes of ESRs and ASRs build-up in VSWM.

## CONCLUSION

Considering our data, several general conclusions can be drawn. First, it seems that the characteristics of ESRs and ASRs created in VSWM are very similar to that typically reported for SRs in long-term memory (LTM). In fact, we have seen that while topological relationships within the retained visual scene are very well preserved, the more exact metric and depth information is lost to a large extent. That is, our data show that SRs in VSWM tend to be schematic – a trend often observed for spatial LTM but also for verbal LTM (e.g., Bartlett's schema theory, see Wagoner, 2013). This raises the question of how VSWM and spatial LTM interact when building ESRs and ASRs. In accordance with classic cognitive memory models (Atkinson, Shiffrin, 1968), we would advocate the position that SRs (at least, ESRs) are initially built within the VSWM and are later transferred to spatial LTM. Clarifying this relationship would surely somewhat change the tradition of preferably studying ESRs/ASRs as representations within LTM. Of course, the relationship between VSWM and spatial LTM may be very complex with respect to ASRs as ASRs may be abstracted from many instances of ESRs stored in spatial LTM (Röhrich et al. 2014, for a recent model on the ESR/ASR-related interplay of VSWM and spatial LTM).

Second, generally we have seen a higher effectiveness of ESRs in the reconstruction of spatial information in VSWM. This could be expected as ESRs are by definition more sensory-related and seem to provide a more exact replica of the perceived visual scene. This finding is in line with the idea that VSWM storage is more in terms of sensory features than LTM, where storage is semantically organized (Conrad, Hull, 1964). It is interesting that we also see the presence of much more abstract ASRs in VSWM which means VSWM storage may be multimodal and abstract (Baddeley, 2012). Generally, our data on the higher efficiency of ESRs over ASRs in VSWM under the condition of limited stimuli presentation and response suggest that (1) ASRs are built-up on the basis of ESRs in the VSWM and the (2) the build-up of ASRs in VSWM is a time-consuming process. Thus, manipulating the time intervals within spatial VSWM task would be a strong experimental manipulation for the study of ASRs abstraction in VSWM in future research.

Third, within the ASRs we have consistently found that spatial information is better preserved in the “left” condition than in the “above” condition. This is a controversial result since we actually had expected that the “above” condition would be most effective. This could be expected as (modern) humans seems to be very well adjusted to form 2D “maps” for the bird's perspective out of 3D environmental information. However, exact the opposite was found. Speaking in ecological terms, abstracting ASRs staying “on the ground” is easier than adding a height direction and abstracting

an ASRs “from above” (as humans naturally are not used to flying). This distinction can be cast in terms differences within 3D mental rotation as mental rotation within a horizontal plane may be easier than mental rotation within a vertical plane (although, generally, 3D mental rotation is more effective than 2D mental rotation (Bartlett, Camba, 2023; Neubauer et al., 2010; Paraskeva et al., 2010)). We think these ideas should be elaborated on more thoroughly in future research.

Fourth and last, we can envisage a differential psychology perspective in the ESR/ASR research. First, a gender perspective may be advanced. Our data shows that there are no gender differences in the basic processes of ESRs and ASRs build-up in VSWM. Still, we had a relatively small sample, so this null result may be due to the low power of our research. On the whole, we think the lack of gender differences can be expected as the processes we study are of fundamental importance for effective spatial interactions for all humans regardless of gender. However, we would hypothesize that if some gender differences will be found in this area of spatial cognition, these would be located more within the ASR domain. This is because ESRs are so basic and immediately related to sensory experience that it would hard to expect fundamental differences between genders within these basic spatial mechanisms. Also, while there were no gender differences, we found indications of substantial individual differences in the reconstruction of exact metric information. This can be a basic and very specific inter-individual variation. Future research should show how it is related to the formation of ESRs and ASRs in VSWM and how it is related individual differences in spatial abilities.

This study is, to the best of our knowledge, the first study of ESR/ASR formation in VSWM using ecologically valid 3D visual scenes presentation in an immersive virtual reality. We found that the formation of ESRs and ASRs is possible in VSWM as a prerequisite for storing a SR in the spatial LTM. We also found that schematic topological information is better preserved than exact metric information already at the level of ESRs, and that there the ASRs built from horizontal rotations are more exact than ASRs built from vertical rotations. We also found evidence that there may be individual differences in the ability to form ESRs/ASRs in VSWM. Overall, our results suggest that VSWM may be involved in the formation of ESRs and ASRs as an important aspect of human spatial cognition.

## **ACKNOWLEDGMENT**

The authors would like to acknowledge Russian Scientific Fund. This study was supported by RSE, grant #22-18-00689.

## **REFERENCES**

- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89–195). Academic press. 89–195.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual review of psychology*, 63, 1–29.

- Bartlett, K. A., & Camba, J. D. (2023). Gender differences in spatial ability: A critical review. *Educational Psychology Review*, 35(1), 8.
- Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature reviews neuroscience*, 9(3), 182–194.
- Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Egocentric and allocentric navigation strategies in Williams syndrome and typical development. *Developmental science*, 17(6), 920–934.
- Chen, W., Liu, B., Li, X., Wang, P., & Wang, B. (2020). Sex differences in spatial memory. *Neuroscience*, 443, 140–147.
- Colombo, D., Serino, S., Tuena, C., Pedroli, E., Dakanalis, A., Cipresso, P., & Riva, G. (2017). Egocentric and allocentric spatial reference frames in aging: A systematic review. *Neuroscience & Biobehavioral Reviews*, 80, 605–621.
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British journal of psychology*, 55(4), 429–432.
- Danjo, T. (2020). Allocentric representations of space in the hippocampus. *Neuroscience research*, 153, 1–7.
- Derbie, A. Y., Chau, B. K., Wong, C. H., Chen, L. D., Ting, K. H., Lam, B. Y.,... & Smith, Y. (2021). Common and distinct neural trends of allocentric and egocentric spatial coding: An ALE meta-analysis. *European Journal of Neuroscience*, 53(11), 3672–3687.
- Fidalgo, C., & Martin, C. B. (2016). The hippocampus contributes to allocentric spatial memory through coherent scene representations. *Journal of Neuroscience*, 36(9), 2555–2557.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269(5220), 41–45.
- Iachini, T., Ruggiero, G., Conson, M., & Trojano, L. (2009). Lateralization of egocentric and allocentric spatial processing after parietal brain lesions. *Brain and cognition*, 69(3), 514–520.
- Ishikawa, T. (2023). Individual differences and skill training in cognitive mapping: How and why people differ. *Topics in Cognitive Science*, 15(1), 163–186.
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition: An interdisciplinary approach to representing and processing spatial knowledge* (pp. 1–17). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281.
- McAfoose, J., & Baune, B. T. (2009). Exploring visual–spatial working memory: A critical review of concepts and models. *Neuropsychology review*, 19, 130–142.
- Menshikova, G. Ya., Saveleva, O. A., Velichkovskiy, B. B. & Bugriy, G. S. (2020). Formirovaniyeegotsentricheskikhiallotsentricheskikhprostranstvennykhrepresentatsiy v rabocheympamyati. *Voprosypsikhologii*. (6). 131–140.
- Neubauer, A. C., Bergner, S., & Schatz, M. (2010). Two-vs.three-dimensional presentation of mental rotation tasks: Sex differences and effects of training on performance and brain activation. *Intelligence*, 38(5), 529–539.
- Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 411.
- Paraskeva, F., Mysirlaki, S., & Papagianni, A. (2010). Multiplayer online games as educational tools: Facing new challenges in learning. *Computers & Education*, 54(2), 498–505.

- Presley, A. (2021). The extent to which autistic traits are predictive of impairments in allocentric spatial navigation.
- Röhrich, W. G., Hardiess, G., & Mallot, H. A. (2014). View-based organization and interplay of spatial working and long-term memories. *PloS one*, 9(11), e112793.
- Self, C. M., Gopal, S., Golledge, R. G., & Fenstermaker, S. (1992). Gender-related differences in spatial abilities. *Progress in Human Geography*, 16(3), 315–342.
- Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: a systematic review. *Ageing research reviews*, 16, 32–44.
- Tuena, C., Mancuso, V., Stramba-Badiale, C., Pedrolì, E., Stramba-Badiale, M., Riva, G., & Repetto, C. (2021). Egocentric and allocentric spatial memory in mild cognitive impairment with real-world and virtual navigation tasks: A systematic review. *Journal of Alzheimer's Disease*, 79(1), 95–116.
- Wagoner, B. (2013). Bartlett's concept of schema in reconstruction. *Theory & Psychology*, 23(5), 553–575.
- Wen, W., Ishikawa, T., & Sato, T. (2013). Individual differences in the encoding processes of egocentric and allocentric survey knowledge. *Cognitive science*, 37(1), 176–192.

# GlasgowSim - Glasgow Coma Scale Simulation

**Dominique Correia de Oliveira<sup>1</sup>, Yassin Aziz Rekik<sup>2</sup>,  
and Stéphane Malandain<sup>2</sup>**

<sup>1</sup>Haute Ecole de Santé de Vaud, Lausanne, VD 1000, Suisse

<sup>2</sup>Haute école du Paysage, d'ingénierie et d'architecture, Genève, GE 1200, Suisse

## ABSTRACT

The training of healthcare professionals is a challenging and prolonged process, requiring a profound grasp of theoretical concepts as well as technical, non-technical skills. Commonly, during the early stages of medical education instructional techniques involve static and unrealistic learning materials, based on the old philosophies. However, contemporary trends favor innovative approaches leveraging new technologies and advanced simulators. The integration of serious games and VR simulation can contribute to the training and enhancement of the skills of healthcare professionals, thereby improving the quality of care, increasing patient safety, and reducing costs of training programs. In this project, we have developed a simulation-based serious game designed for students and professionals to hone their proficiency in evaluating patients' level of consciousness according to the Glasgow scale.

**Keywords:** Health professional education, Serious games, Virtual reality, Patient safety, Glasgow coma scale

## INTRODUCTION

The Glasgow Coma Scale (GCS) is a neurological instrument that measures the “severity” and extent of impaired consciousness. The GCS has become the most used tool in the world to document alterations in the level of consciousness caused by brain damage (Ingram, 1994). In combination with other neurological examinations, the scale is used to estimate the vital prognosis of patients with a severe brain injury. Because of its ease of use for all health professionals in all care settings, the scale has become an essential tool in all training programs. The neurological evaluation requires frequent simulation-based education to improve the cognitive, psychomotor and communication skills of the health students. However, current simulation approaches are resource-intensive and not routinely offered in all healthcare schools. Also, alternative approaches are needed to improve working memory, decision-making skills, and teamwork performance. Serious games may be effective and more accessible alternatives if they use active, experiential, and problem-based learning. These tools are highly likely to ignite student motivation and facilitate their acquisition of knowledge within intricate learning scenarios. A multidisciplinary team including experts from healthcare, education and engineering ensure a

coherent interaction between the game, content, and pedagogical features. The GlasgowSim Project aims to create a serious game designed to enhance the training of aspiring healthcare professionals in Neurological Evaluation by encouraging knowledge acquisition and practical skill development.

## CONTEXT

### Teaching Neurological Assessment at the Haute Ecole de Santé du Canton de Vaud (HESAV)

Students undergo comprehensive training through a series of practical workshops, which alternate between role-playing and the analysis of practical cases. HESAV employs simulation as a teaching method to enhance the precision of clinical observation and stress the significance of promptly identifying clinical alerts in neurological assessment—a vital skill across all healthcare disciplines. The knowledge and skills acquired during these simulation workshops represent a broadening of the clinical skills profile of Bachelor nurses in Switzerland. These skills, applied in clinical assessment, help to improve the quality of care (Lindpaintner et al., 2009) and at the same time enhance the position of the nursing profession in everyday clinical practice. As a result, students acquire in-depth knowledge and skills in taking a medical history and in basic neurological examination techniques.

Given the paramount importance of assessing consciousness levels clinically, we teach the use of the Glasgow Coma Scale (GCS). Accurately determining the severity of the altered state of consciousness is not only to establish a vital and functional prognosis, but also to monitor more effectively the patients with brain injury and to optimize the use of available resources.

### The Glasgow Coma Scale (GCS)

When Teasdale and Jennett published the Glasgow Coma Scale (GCS) in 1974, they promoted a major change in the clinical assessment of patients with an altered state of consciousness, by taking particular care in constructing their score. The scale was designed to be used reliably by all healthcare professionals. The authors aimed to establish a means of based on simple items with clear, unambiguous, and easily translatable definitions (Ledoux et al., 2008).




The scale is based on the separate clinical assessment of three distinct aspects of the behavioral response to stimuli during altered states of consciousness (Teasdale and Jennett, 1976): i) the motor response (M, five then six levels); ii) the verbal response (V, five levels); and iii) the eye opening (E, four levels).

The opening of the eyes reflects the integrity of the nerve connections within the central portion of the gray matter and cranial nerves III, V and VII in the brainstem. The verbal response makes it possible to estimate the degree of damage to Wernicke's area (language comprehension center of the temporal lobe) and Broca's area (speech production center of the frontal lobe). The motor response indicates the capacity of the sensorimotor fibers



to capture and interpret sensory information coming from the body or the environment and react to it in an appropriate manner.

By summing the evaluations of the different responses to obtain an overall score, with a scale ranging from three to 15 points, making it possible to establish a correlation between the severity of altered states of consciousness and patient outcomes. This allows healthcare professionals to align the severity score with appropriate clinical interventions (see Figure 1).

Behaviour	Response
 Eye Opening Response	4. Spontaneously 3. To speech 2. To pain 1. No response
 Verbal Response	5. Oriented to time, person and place 4. Confused 3. Inappropriate words 2. Incomprehensible sounds 1. No response
 Motor Response	6. Obeys command 5. Moves to localised pain 4. Flex to withdraw from pain 3. Abnormal flexion 2. Abnormal extension 1. No response

**Figure 1:** The “GCS” – Glasgow Coma Scale (Adobe stock).

The relationship between GCS score and the outcome constitutes the basis of a common classification of acute head injuries (Jain and Iverson, 2023): Severe, GCS 3 to 8 - Moderate, GCS 9 to 12 - and Mild, GCS 13 to 15.

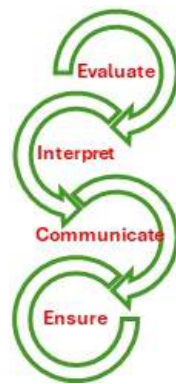
The GCS is recognized for its ease of administration by all healthcare professionals in all care settings. The scores obtained guide professionals in their clinical decisions, to predict the patient’s prognosis and to evaluate the effectiveness of the treatments provided. A decrease in score from one assessment to another may reflect a worsening of the brain injury requiring immediate medical intervention, while an improvement in the score indicates a positive response to treatment. The frequency of application of the Glasgow scale varies depending on the neurological condition of the patient and the time elapsed since its admission, or even since the appearance of the first signs of brain damage. However, there is a lack of rigor in the evaluation of states of altered consciousness because the Glasgow scale is only used in 42% of care units (Santos, 2016). Equally worrying, health professionals claim to have learned to use GS by reproducing the habits of their department without necessarily having understood the principles of application (Reith et al., 2016). Thus, clinical experts agree on the fact that it is necessary to demystify this scale and to implement periodic training that uses active, immersive teaching methods based on real situations (Edwards, 2001).

### **The GlasgowSim Project**

The GlasgowSim project aims to develop a pedagogical innovation that meets the requirements of the curriculum and extends the teaching options for the

Glasgow Assessment Scale. The user is invited to interact with a computer-based device that combines teaching aspects with playful elements derived from video games (Chabert, 2010).

Theoretical and simulation workshops are currently being organized, but straight neurological assessment needs to be made more accessible to health students who will be applying ECS daily as future clinicians. When used for demonstration purposes, these approaches are resource-intensive (mannequins, technicians, teachers, etc...). Moreover, through the various satisfaction evaluations, students express their desire for access to diverse teaching tools to anticipate forthcoming challenges in clinical settings. On the one hand, the teachers in the HESAV nursing care stream have developed clinical vignettes that promote the integration of knowledge and student learning. On the other hand, the teachers in the HEPIA IT stream have developed a virtual environment that allows users to immerse themselves in a real-life neurological assessment situation. The present approach aims to cultivate participants' reflective practice, thereby refining their clinical judgment in assessing and managing patients with neurological disorders. The trainers are convinced that simultaneous learning of *Why* and *How* the Glasgow Coma Scale is administered would optimize the clinical judgment of health professionals in the neurological assessment of brain-injured patients. For this reason, the team project has determined 4 structuring stages in the use of the game (see Figure 2):



**Figure 2:** Four stages of GlasgowSim.

1- EVALUATE: the level of consciousness of a patient with neurological impairment using the Glasgow scale.

2-INTERPRET: the scores obtained following the evaluation of three components of the Glasgow scale.

3-COMMUNICATE: critical information regarding a patient's care in a structured manner.

4-ENSURE: clinical monitoring of a patient with neurological damage.

In the game, the user will have to transmit important clinical information through a standardized tool, the SBAR: Situation - Background - Assessment -Recommendation (see Figure 3).



**Figure 3:** SBAR, transmission structuring tool.

SBAR allows structured and effective communication between healthcare professionals, promoting the transmission of relevant information and facilitating the coordination of patient care.

## GLASGOW SIMULATION DESIGN

### Multidisciplinary Team

Our aim was to create a tool that is appealing, user-friendly, adaptable, and customizable to cater to the requirements of both students and educators, while incorporating best practices. GlasgowSim project concerns three domains: (1) healthcare, for the development of clinical scenarios; (2) engineering, for the software application development and game technology advice, (3) pedagogy, as the science and art of teaching and learning.

The development of the immersive gamified simulation in GlasgowSim project relied on a multidisciplinary team. The team was composed of a health professional educator, an educational scientist, two computer scientists and students from Master of Science HES-SO in Computer Sciences Engineering. The team followed an iterative and collaborative approach to design and develop the serious game.

### Initial Conceptual Choices

GlasgowSim has been designed and developed by adopting six conceptual choices.

**Immersion:** recent literature clearly shows that a high level of user immersion during a simulation increases acceptance, strengthens motivation, and intensifies realism. That is why, from the outset of the project, we opted for an immersive simulation using virtual reality. At the outset, some thought was given to whether an approach based on augmented reality would be more appropriate. However, the intended mode of use did not justify such an approach. The scenario did not involve any interaction with real objects or the visualization of a real setting. The aim was clearly to simulate the entire context and immerse the learner in a highly realistic virtual reality.

**Realism:** Realism was desired for two essential reasons. Firstly, the realism of the setting and environment intensified the immersion of the learners and gave them a sense of 'déjà vu' compared with real hospital diagnoses. The second aspect is the realism of the patient and the diagnosis. Indeed, the aim

of the simulation is to train learners to carry out a diagnosis according to GCS by performing the necessary gestures and procedures. It is therefore imperative that the patient and his diagnosis are as realistic as possible. However, in a Virtual Reality situation, it is not always easy to represent and capture these indicators. Taking the patient's pulse is an example of this difficulty in the absence of a haptic device.

**Pedagogy:** With the aim of providing a comprehensive training tool, we chose to develop a simulation that could be used in various modes. We have selected three main modes of use. The learning mode, in which the learner will be guided and have access to a teaching aid to reinforce their theoretical and practical knowledge. Then there is the practical mode, in which the learner can carry out a diagnosis without support or assistance and be challenged by variations and constraints. Finally, the assertion mode enables the learner's learning to be examined in a controlled, time-bound mode, with the possibility of weighting the assessment criteria.

**Gamification:** The gamification of a simulation makes it attractive and thus intensifies learner motivation and commitment. In addition to the classic gamification techniques (introduction of scores, time limits, rewards, etc), we decided to introduce two innovative aspects. The first is the transformation of training itself into an adventure in which a GSC diagnosis becomes a stage for acquiring clues and resources for progressing in the game. The second aspect is to model the teaching aids as limited game resources that require effort to obtain. This allows the resources to be better valued and therefore better appropriated and memorized.

**Modularity:** the two recurring problems with gamified simulations are their development costs and their rigidity. To reduce these two aspects, we decided to develop our simulation in a modular way that would allow it to be adapted and extended in the future. The idea is to develop the simulation by aggregating various modular and configurable objects. This makes it easier to adapt these objects and add new ones. The typical example of this approach is the patient himself, who has been developed as an autonomous object that can be integrated into various settings and contexts. Other examples are characters, instruments, diagnostic tools, etc.

**Genericity:** the final choice we have adopted is not to freeze the scenarios in the game, but rather to allow teachers to define their own scenarios and case studies. The simulation we are going to develop will be a player of scenarios given as parameters. This functionality will empower educators to generate a multitude of scenarios, encompassing the wide array of situations encountered in real-life settings. Each scenario will be crafted using an interactive, assisted tool, which will produce a comprehensive description sent to the simulator. This description will establish parameters for all components, including patients, settings, instruments, and behaviors.

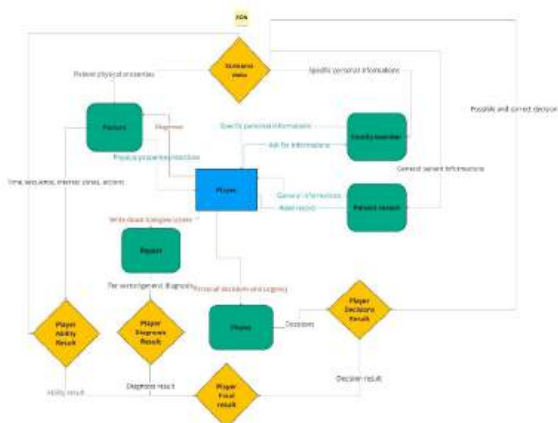
### **GlasgowSim Architecture and Components**

Based on the design decisions mentioned above, we have broken down the GlasgowSim solution into three essential components. The first component is the simulation tool, which allows the learner to interact with the patient, make diagnosis, and deliver assessments and recommendations. This tool creates the complete simulation environment based on a description of the

scenario in JSON format and aggregates the necessary objects with the right configuration. The second component is the interactive patient. We deemed this element to be essential due to the pivotal role of the patient in facilitating learner diagnosis. As will be detailed later, this patient has been developed to enable realistic GSC diagnosis. Finally, the last component is the interactive scenario creation tool. This tool will enable teachers to define all the parameterizable elements of a scenario and to export the scenario in JSON format to the simulation tool.

**Immersive Simulation Tool**

One of the aims of the GlasgowSim project is to allow a multitude of scenarios and settings. For this reason, the simulation environment has been designed by combining two types of modules. Firstly, there are the user interface modules, which are responsible for providing information to the user or receiving data from them. These include the patient module, which represents the virtual character with which the user interacts. We also have the patient record module, which contains relevant medical information, and the diagnosis rendering module, which allows the user to submit their results through virtual tablet, virtual phone, virtual patient record, etc. Then, we have the data processing modules, which retrieve information from the scenario definition file (JSON format) and process it. These modules can compare the user’s results with the scenario data or simply provide information to the user interface modules. Their role is to guarantee the smooth operation of the simulation by utilizing the relevant data. By using this modular approach, we can effectively manage user interactions, provide the necessary information to the learner, and react to the learner actions and assessments within the immersive and entertaining simulation. This modular approach allows also for extensions to be added later with minimum adaptation. The generic scenario model used in GlasgowSim is shown in the Figure 4.



**Figure 4:** Scenario model for Glasgow simulation.

### **Interactive Animated Patient**

For this project, patient modelling is an essential component. It must be as realistic as possible to make the simulation more immersive. The patient must necessarily have a skeleton comprising the whole of his body, including his head, eyes, and mouth, so that we can then create animations adapted to the practices and diagnoses to be performed by the learners with a precise level of detail. The creation of the patient in the GlasgowSim project was based on UMA2, also known as Unity Multipurpose Avatar, which is an open-source project developed for Unity3D. Its aim is to enable the creation and customization of 3D characters in a flexible and modular way. UMA2 provides prefabs for dynamically generating humanoids. From these prefabs, it is then possible to add clothes to the patient, change his morphology and personalize him.

In addition to animations and tactile interaction with the patient (colliders zones), we have also developed voice interaction. This interaction is based on *vosk-model-small-en-0.22*. This is a model in French created exclusively for small devices. The model is based on the following sources: i) an acoustic model containing the sounds of the language; ii) a language model containing word sequences; and iii) a phonetic dictionary to break down words into sounds.

Vosk's small language models allow the vocabulary to be reconfigured at runtime. This allows the language model to recognize only the words given following reconfiguration.

### **Scenario Generator Application**

To be able to create new scenarios and new clinical situations simply and easily, a scenario generator has been developed in the form of a web application connected to the simulator. This tool enables new scenarios to be created by defining the corresponding medical records and all the patient's medical and behavioral parameters. Once created, a scenario is packaged in the form of a JSON file and sent to the simulator, which can then parse it, extract the information, and automatically generate a new simulation.

## **IMPLEMENTATION, VALIDATION, DISCUSSION**

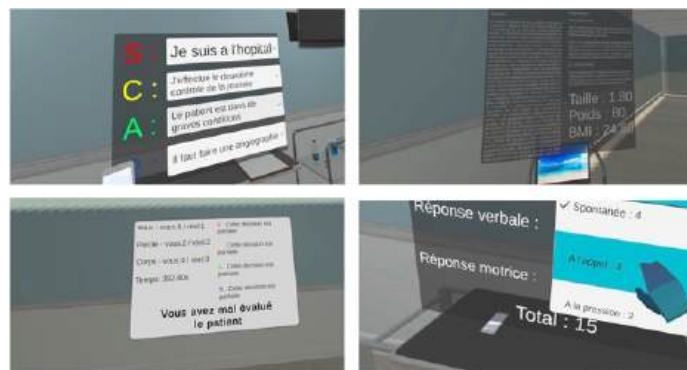
### **State of Progress in the Project**

The GlasgowSim project is still in progress, but we already have a first functional and testable version. In this first version, all the scenarios take place in a realistic hospital (see Figure 5). This is a choice made by the HESAV teaching staff to replicate real-life sessions as closely as possible. However, there are plans for future diversification of environments to include training for diagnoses outside of hospital settings, such as road accidents, domestic accidents, etc.

The current simulation tool covers the three types of GCS diagnoses and allows practice on the SCAR tool recommended by the HESAV partners, as shown in Figure 6.



**Figure 5:** 3D environment used in the current version of GlasgowSim.



**Figure 6:** Some 3D GUI objects used in GlasgowSim simulation tool.

The scenario generation tool is also functional and uses a simple Web interface to create and configure new scenarios, store them in a scenario database and then import them into the simulation tool. So, as far as these two tools are concerned, we consider our project to be well advanced and the pilot users are showing very positive feedback.

In terms of the interactive patient, however, a lot of work remains to be done. The current version of the patient does not seem to achieve the desired degree of realism. Muscle animations, which are responses to stimulation of specific areas by the learner, need to be better adjusted to approximate real movement. Eye movements also need to be improved and equipped with more intelligent behavior to adapt to the context: the learner's location, movements in the vicinity, movement tracking, etc. Finally, the patient's current conversational ability is judged to be very limited. To streamline operations, we initially opted for straightforward models that minimize server-side processing. However, it turns out that this choice is questionable and does not give satisfactory results. Finally, although some gamification aspects have been introduced in the current version, a lot of work remains to be to complete a fully gamified version in the form of a well-structured and balanced adventure-economy game.

## Validation and Discussion

The GlasgowSim solution is currently operational. For the time being, we have only carried out tests with pilot students outside the course. The aim is to assess the degree of satisfaction and usability. These initial tests have produced very positive results, which have encouraged us to plan tests next academic year.

## CONCLUSION

Several studies investigating the application of the GSC have underscored that healthcare professionals lacking experience and ongoing training may inadvertently commit significant errors, especially in cases involving intermediate levels of consciousness. Such errors are particularly critical as they impede accurate detection of neurological changes crucial for patient monitoring (Rowley and Fielding, 1991). It is therefore important to integrate active teaching methods and simulation to diversify the learning scenarios which will allow students to train in environments that closely mimic real-life situations, thereby enhancing their preparedness for clinical practice.

## ACKNOWLEDGMENT

The authors would like to acknowledge the entire teaching team of Neurological Clinical Evaluation, HESAV and the students who participated in the construction of the prototype, namely Mr. Jamal Albadri and Mr. Teo Ferrari.

## REFERENCES

- Chabert, G., Desmarais, C. and Jacques Ibanez Bueno (2010). Des cadres de santé 'jouent' à évaluer. Le 'Serious gaming' pour appréhender l'évaluation ?. *Communication et organisation*, 38(38). doi: <https://doi.org/10.4000/communicationorganisation.1419>.
- Edwards, S. L. (2001). Using the Glasgow Coma Scale: Analysis and limitations. *British Journal of Nursing*, 10(2), pp. 92–101. doi: <https://doi.org/10.12968/bjon.2001.10.2.5391>.
- Ingram, N. (1994). Knowledge and level of consciousness: Application to nursing practice. *Journal of Advanced Nursing*, 20(5), pp. 881–884. doi: <https://doi.org/10.1046/j.1365-2648.1994.20050881.x>.
- Jain, S. and Iverson, L. M. (2023). *Glasgow coma scale*. [online] National Library Of Medicine. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK513298/>.
- Ledoux, D., Piret, S., Boveroux, P., Bruno, M., Vanhauzenhuysse, A., Damas, P., Moonen, G. and Laureys, S. (2008). Les échelles d'évaluation des états de conscience altérée. *Réanimation*, 17(7), pp. 695–701. doi: <https://doi.org/10.1016/j.reaurg.2008.07.004>.
- Lindpaintner, L. S., Bischofberger, I., Brenner, A., Knüppel, S., Scherer, T., Schmid, A., Schäfer, M., Stoll, H. R., Stolz-Baskett, P., Weyermann-Etter, S. and Hengartner-Kopp, B. (2009). Defining Clinical Assessment Standards for Bachelor's-Prepared Nurses in Switzerland. *Journal of Nursing Scholarship*, 41(3), pp. 320–327. doi: <https://doi.org/10.1111/j.1547-5069.2009.01286.x>.



- Reith, F. C. M., Brennan, P. M., Maas, A. I. R. and Teasdale, G. M. (2016). Lack of Standardization in the Use of the Glasgow Coma Scale: Results of International Surveys. *Journal of Neurotrauma*, 33(1), pp. 89–94. doi: <https://doi.org/10.1089/neu.2014.3843>.
- Rowley, G. and Fielding, K. (1991). Reliability and accuracy of the Glasgow Coma Scale with experienced and inexperienced users. *The Lancet*, 337(8740), pp. 535–538. doi: [https://doi.org/10.1016/0140-6736\(91\)91309-i](https://doi.org/10.1016/0140-6736(91)91309-i).
- Santos, W. C., Vancini-Campanharo, C. R., Lopes, M. C. B. T., Okuno, M. F. P. and Batista, R. E. A. (2016). Assessment of nurse's knowledge about Glasgow coma scale at a university hospital. *Einstein (São Paulo)*, 14(2), pp. 213–218. <https://doi.org/10.1590/s1679-45082016ao3618>
- Teasdale, G. and Jennett, B. (1976). Assessment and prognosis of coma after head injury. *Acta Neurochirurgica*, 34(1–4), pp. 45–55. doi: <https://doi.org/10.1007/bf01405862>.

# Hybrid Improvisational Theatre: A Thematic Review of the Production Processes

Daniel Lock<sup>1</sup>, David Gochfeld<sup>2</sup>, and Ben Kirman<sup>1</sup>

<sup>1</sup>University of York, York, UK

<sup>2</sup>The Royal Central School of Speech and Drama, London, UK

## ABSTRACT

The recent uptake in hybrid forms of communication and modern consumer XR products has led to the development of a number of successful hybrid theatre productions. However, while such productions are being produced there has been limited academic work investigating its impact on the development of theatre production from the practitioner's perspective. This paper aims to rectify this by developing an in-depth understanding of the creative and technical challenges posed by hybrid theatre by examining the creation of such production through interviews with the practitioners and audiences. The results of this paper highlight a clear correlation between the audiences and practitioners understanding and expectations of hybrid productions.. However, there is an interesting disparity between the two viewpoints on the difficulty of implementing change, setting an elevated expectation on what is viewed as possible compared to what can be achieved with the constraints of budget, time and resources. We conclude that the use of hybrid technology positively affects the theatre space and provides opportunities for novel, exciting avenues for immersive and interactive productions. Furthermore, a rich understanding of the needs of practitioners and audiences can positively affect the theatrical production development process.

**Keywords:** Hybrid, Improvisational theatre, Thematic analysis, Production development, Virtual audiences, Digital twin, Multi-location interactions, Interactive media, XR, Virtual reality

## INTRODUCTION

Hybrid events, offering both in-person and online access, have become increasingly common, since the COVID-19 pandemic as a way to make both work and entertainment more accessible (Vyas, 2022; Kurt, 2024). This global cultural shift along with the growing uptake in modern consumer XR products (Alsop, 2024) has contributed to the development of a number of well received hybrid theatre productions (Boosted, n.d.; Deakin, n.d.). Such productions generally aim to offer an entertaining experience for both in-person viewers and virtual, globe-spanning, audiences through the use of a range of different, often interactive, elements such as live streams, WebVR and live chats. This paper takes a different direction with a focus on understanding how the needs of both the practitioners

(Lock and Kirman, 2023) and audience may be realised to encourage co-beneficial development of future productions. To investigate these topics the researchers posed the following questions:

What are the developmental challenges of hybrid theatre?

What are the viewership challenges of hybrid theatre?

What affordances can be made to support both perspectives?

By concentrating on these questions; the outcomes of the study aimed to develop a rich understanding of the needs of both the practitioner and audiences. Through consideration of both perspectives; this paper identifies opportunities that may further the development of hybrid theatre productions while minimise compromises to quality, maximise viewer experience and reduce barriers faced by practitioners. For this purpose, researchers from the University of York collaborated with the theatre company FANDCO (Fandco, n.d.) to create a hybrid improvisational theatre production, from ideation to performance, interviewed both the production team and hybrid audience to and analysed the results through thematic analysis. The following paper will highlight nuances in hybrid theatre and related work within the area considering how the needs of both the practitioners and audiences are widely viewed in the immersive theatre space. The paper will then highlight the development process of 'House of Masks' the hybrid production developed in collaboration with FANDCO focusing on introducing how such productions are developed. Finally the paper will present a thematic analysis through data collected during interviews and via a questionnaire that highlight the needs of both practitioners and audiences. The main contribution of this paper is the rich understanding of how the development of hybrid production affect both the practitioners and audiences alike, highlighting how these two perspectives directly impact the other's access to the space, focusing particularly on what can be achieved by a production with the correct tools and solutions available, allowing to captivate its audience through interaction and accessibility.

## **HOUSE OF MASKS**

House of Masks is a hybrid improvisational theatre production that was created for this research project through a collaboration between the researchers and the theatre company FANDCO (Fandco, n.d.). This section will provide a detailed overview of the production's premise along with an outline of how the production work for the in-person and online audience members.

As an improvisational production the composition of this play was somewhat different to how regular plays are developed. A number of key areas of the performance, including the story, stage direction and locations, were not decided upon as part of the planning and development process. Instead, these details are decided upon during the live performance, with the audiences participation. This is common in improvisational theatre (Zaunbrecher, 2011) and is what forms part of the uniqueness of such productions, where no two shows are exactly the same. However, while such details did not exist an overarching story premise that outlined the world the

story was set in was created. This is often the case in large improvisational productions to ensure a high level of cohesion can be achieved. Within the premise of this production this was especially important as some aspects, such as the virtual set and atmosphere would need to be developed beforehand to ensure that their aesthetic matched the tone of the production.



**Figure 1:** Actors performing in the physical and virtual environments.

As a hybrid production the format can be expressed as two halves, the physical and the virtual. For the former the actors performed in a Black Box (Hannah, 2024) theatre with no props or costume. The only asset used was a projector screen that was placed behind the actors to show the virtual scene that the online audience were witnessing and to act as a backdrop adding further context to the production. In the latter, a virtual world created by FANDCO in Unreal Engine 4 (Unreal Engine 4, n.d.) was utilised. This consisted of two environments, a house and a cloud, and three avatars which were scans of the actors. To observe this environment the online audience were given a link to a YouTube live stream (Google, n.d.). To act between these two spaces simultaneously the actors wore Motion Capture suits and their movements were translated onto their virtual avatar providing accurate, one to one movements.

The development of the story as part of the live performance experience required a high level of collaboration between the actors and the audience. For this collaboration the audience were encouraged to provide prompts in the forms of phrases, actions and props that the actors would then build upon to steer and alter the direction of the story. This interaction was delivered via different formats to cater for the different audience viewing experience, the in-person audiences were provided with paper and asked to submit prompts before the show began whereas the online audience were actively asked throughout the production for prompts via an interactive live stream chat, these prompts were then randomly selected. The researchers decided to create a new production for this project for two primary reasons, first to gain firsthand experience of the development process deepening their

understanding of the practitioners perspective enabling them to ask more meaningful questions during interviews. Second, to create an opportunity to collect first hand data from a production team and audience who had recently participated in a hybrid theatre production, be that via the development of said production or as a member of the audience, ensuring synergy in their responses to interview questions.

## **METHODS**

### **Participant Recruitment**

Participants for this study were proactively recruited from both the in-person and online audiences. Both audiences were invited to participate in answering a questionnaire and were asked to participate through an informal invitation before they viewed the production, once the performance had concluded the researchers spoke to the in-person audience members directly and either interviewed them in the theatre or scheduled an interview for a later date. Online audience members were recruited in a similar way with the chat moderator offering a link to sign up to an interview. Both audiences were provided with a link to the questionnaire and asked to complete it. As discussed in the previous section, the choice of interviewing the audience of this production ensured that everyone has a similar definition and understanding of what a hybrid improvisational theatre production was. It would also ensure that the processes followed by the practitioners in creating this production would be critically reflected upon by the audiences enabling a rich discussion of how they could be altered and improved in the future.

### **Data Collection**

When conducting research on the audience two data collection methods were employed for this research project, questionnaires and semi-structured interviews. The questionnaire was created using Google Forms (Google Forms, n.d.) and consisted of a mix of questions that could be answered with either a multiple choice selection or through writing a short paragraph. Multiple choice questions were used to understand the status of the audience member responding to the questionnaire and included questions on their mode of viewership from the production and their overall experience in the area of hybrid improvisational theatre. The questions that required a short paragraph answer were directly related to the participants experience of the show and how they would critically evaluate the overall production.

The interviews were conducted using the video conference software Zoom (Zoom, n.d.). The use of semi structured interviews enabled research to keep on topic while opening up interesting areas discovered during the interview. Likewise, the practitioners were interviewed in the same manner as the audience, without the completion of a questionnaire. This decision was taken as the aim of the questionnaire was to enable the researchers to gain valuable and rich data from a large portion of the audience who might be unavailable to take part in a longer interview. From the data collected

the questionnaire received 13 responses with 6 audience interviews and 7 practitioner interviews, due to the richness of the data in these interview no further interviews were conducted.

### **Data Analysis**

Once the interviews and questionnaires had been completed the data was prepared for analysis. To do this the interview recordings were transcribed and the questionnaire data extracted and formatted. To analyse the data the researchers performed a thematic analysis using the process developed by Braun and Clarke (Braun and Clarke, 2016). Utilising this method the researchers first read the formatted data multiple times to develop their familiarity of the data. This data was then analysed and coded by looking for patterns within the data which meant were in line with the research questions. After coding the data the codes were examined and 4 themes were developed that encompassed homogenous codes. After further coding of the data, the researchers developed 4 themes: Affective Collaboration, Opportunities Afforded by Improvisation, Disbalanced Hybridity and Positive Outlook on Production Issue. These themes highlight the critical feedback received from both the audience and practitioners with a focus on understanding their experiences of the hybrid production and where friction and opportunities for further collaboration between audiences and practitioners lay.

## **RESULTS**

The four themes presented in this section provide insight into the developing relationship between the practitioners, the audience and their joint experience of a hybrid production. The rest of this section will break down the developed themes and provide context to the results.

### **Affective Collaboration**

An element running through the creation and viewing of the production is the development of a collaborative process between the improvisational actors and the audience where prompts are provided seamlessly before the production or through a live stream chat to steer the performance's direction. This process requires a high level of trust between both parties in order to ensure the audience felt immersed in and part of the production. For the majority of the experience this goal was achieved leading to a high level of synergy between both the actors and the audience:

“The actors were talking after our first show and saying how we actually didn't know that a lot of those suggestions were from the audience because they fit so beautifully.”

Furthermore it is clear that this deeper level of communal collaboration provided an positive increased level of immersion throughout the production for the audience in manner that may not be easily achieved in more non-hybrid improvisational production:

“It’s kind of like the audience is going ‘they [actors] are invested’ and they’re going on this with you or with them.”

Likewise, this is mirrored by the actors themselves who, without the use of technology, would have had to rely on the use of prompt given before the performance began, which could lack the context of the current pacing or may even result in short pauses in the production throughout to receive prompts which break the immersion for all:

“Whereas what I think is beautiful about this [giving prompts through a chat] is it doesn’t break your immersion in the narrative and in the world and in the lives of these characters. Ah Yeah. So I think it’s that that seamlessness, that immersion ... it’s not an interruption at all it just contributes to the flow.”

However, this collaborative method was not accessible to all audience members. With a focus on the live stream chat as a way to provide prompts during the performance the in-person audience were only able to observe the story progression. From the practitioner perspective this method is successful at allowing the pace and flow of the production to continue seamlessly:

“In a live space to be able to facilitate, it would interrupt the flow, but by having it online and sort of a separate dedicated space, we’d have [Moderator] there filtering in case there is anything that is unhelpful.”

On the other hand, as this left a gap in the overall viewing experience for the in-person audience, which was not filled through a different interaction method. This led to several members of the in-person audience feeling dissociated with the collaborative process:

“It was very interactive [at the] start obviously when they were like getting suggestions from the audience and bringing them all into it. I think afterwards it felt, it felt more like I was, you know, watching live theatre. I didn’t feel like I was in it, but I definitely was like invested in what was happening.”

Within this, several interviewed audience members suggested a solution that could be implemented with little to no extra resources required. The suggestion was to add a notification to the virtual scene that informed the audience what prompt had been selected; or was currently being used, which would help in person audiences identify the prompts easier and overcome any confusion created by the prompts in the scene:

“An example of in what way they are being guided, what kind of suggestion they’re getting ... it would just be good to know how fully improvised it all is and how much they’re being led by other people’s suggestions.”

### **Opportunities Afforded by Improvisation**

Improvisation is often utilised in productions that blend the line between theatre and technology. This is employed as a method to overcome issues

that may arise from working live with experimental technology, this was also present in this production:

“At one point [actor’s] tracking broke and [their] head was in the floor, and that they took that as a as input for their improv and used that for a moment, until it got fixed. So I think there’s a lot of opportunities there that are really I’m excited to continue exploring it.”

However, *House of Masks* takes it a step further as a completely improvised performance. While this relies on the skill of the actors it also requires the technology to be robust and adapt to different prompts. Within these elements it is possible to argue that a difference in understanding can be found between the two perspectives as the audience wanted to see the two worlds, physical and virtual, more aligned:

“The thing that I found most interesting was, when the actors were able to align with what was happening in the back, behind them. So I’d love to see more of that because it really connected the two, most of the time they were kind of just ... disconnected.”

On the other hand, with the context of the actor’s constraints, it is clear that this is a direction the practitioners also wanted the production progress in, but due to limitations on what was possible with the technology used it could not be realised in the current format:

“I’m ignoring the back screen. Because otherwise I’ll just stare at it And then my character will stare at it And it will go on forever in a kind of infinite loop.”

Ultimately this provides a starting point for investigating how and where improvements can be made to the hybrid improvisation theatre creative process both in terms of the setup, placement and feasibility of the technology and the high skill level of working within the space as an actor:

“The director said he had taken the learning after to put monitors in front so they could interact.”

### **Disbalanced Hybridity**

Part of the process of creating a hybrid production is to ensure that an equal experience can be received by both in-person and online audiences. While this does not necessarily require both viewpoints to observe the production in a near identical way it does set an expectation that both audiences have a similar experience. Within the context of this research the practitioners aimed to deliver this by proving the in-person audience access to the virtual world through a projection which was met with a good level of success:

“Well as I say, for the first, for the first night I felt it was just, we’re just watching the actors because it doesn’t really matter what’s going on the screen. Whereas the second time I saw it, I felt like, oh, they are bringing more into what’s happening on screen so I can watch the screen and feel that’s like, I’m still getting the story, whereas I did it the first time.”



However, it is clear that the issue being faced by the audience is not the separation between the in-person and online viewing experience but the disparity in how the production are catering to the two leading to the audience to favour the in-person option to receive the full experience:

“I think using that technology very much in the theatre space rather than just being something you can stream online is much more impressive and you get more it and it’s much more of a, of an experience. You come out and I’ve seen a show where there’s amazing aspects that are happening all around me. Whereas if you’re watching online, you don’t really know what’s going on.”

This does not suggest that the production provided unequal viewing experiences for the audience but rather that the process or ensuring the production itself is truly hybrid for all involved may require further immersive elements to be considered and included, as the practitioners have highlighted:

“As an actor you still feel the presence of the audience and it affect the performance, and I think [you have to] sort of make that explicit and then look at how we take that out of the theatre building, because obviously it’s easier to have that connection if you’re in the same room with somebody.”

### **Positive Outlook on Production Issues**

The creation of anything new will come with issues and challenges that need to be addressed. This is no different with hybrid improvisational theatre where a degree of randomness through live prompts adds to the complexity of creating a robust experience. Factoring for such sets a high demand and causes strain on the resources, such as time, budget and expertise, available:

“From a technical perspective, obviously the overhead is significantly higher. We have to get in, set everything up much more in advance. Rehearsals are an interesting challenge where we have to kind of bear in mind where the tech may or may not be and then respond to that in the week when we get in [to the theatre].”

This can sometimes cause overestimations on what is possible within the space compared to the production itself as there is not often the context for what resources are available from the audience perspective:

“I would’ve loved to see, to have more happen. Like if they, you went to different rooms in a house or, I don’t know, there was more there somehow.”

However, in the context of the House of Masks this was rarely seen. When discussing the issues in audience interviews, a deep level of understanding was exhibited:

“It felt pioneering and so you were, you would forgive a lot of stuff, you know, which I think is a great way to do it because you know, you feel, as an audience member, like you’re seeing something really new.”

This may have been caused by the experimental nature of the production leading to a heightened interest from theatre and technology experts attending as audience members, nevertheless it suggests how audiences are developing an increasingly in-depth understating of what it can take to produce such technology lead productions:

“It was very different from a, a normal improv show, but then they were doing something different. I think it will evolve. Over time as they become more used to it and we [the audience] become more used to it because it’s technological, that’s the thing.”

## **DISCUSSION**

The results rationalised above provide an in-depth and rich investigation into the practitioner and audience perspectives on the creation, performing and viewing of a hybrid improvisation theatre production. As an improvisational production the collaborative process of developing the story live, with prompts from the audience, is a highly important area of the production that required well developed improvisation methods to ensure it is implemented successfully. With this in mind there is a clear level of expectation from the audience on how this may be delivered, Whilst this was not identical across the two audience viewpoints, it did provide a well-developed and seamless way to send and receive prompts. As suggested by the audience, the collaboration could be further improved by providing the audience feedback in the form of a notification on what prompt was currently being used. It is also clear that the hybridity of the production could also be altered. Although the production worked well overall the experience of both the in-person and online audience was quite different. This is somewhat to be expected as, just like hybrid meeting using a mix of in-room and video conferencing software, there are parts of each viewing option that makes the experience unique. Although this is important to ensure that both viewpoints have an enjoyable and immersive experience, some areas such as giving prompts during the production, could be implemented for both audiences in a non-invasive and non-distracting way. With the use and combination of numerous, some experimental, technologies there is always a potential for technical issues to appear during the creative process. Although the majority of such issues are often resolved during the development of a production, it is somewhat inevitable that challenges may arise, more so in the case of improv productions due to the high level of randomness introduced by audience interaction and collaboration. In such instances, the strengths of improvisation can be seen and, in a similar way to VR theatre production (Lock, and Kirman, 2023) the use of improv enables the actors to overcome or explain away the problem by acknowledging it and proving context.

## **CONCLUSION**

In conclusion, hybrid improvisational theatre is a novel form of theatre that combines the recent advances in immersive and interactive technology with the accessibility of hybrid and the timelessness of theatre. Through the analysis of data from practitioner and audience interviews there is a clear

path forward to continue the development and improvement of the creative hybrid improvisation theatre process.

## FUTURE WORK

Throughout this paper it is clear that the area of hybrid improvisational theatre as both a mode of performance and an area of research is quite novel and has numerous areas that could be improved with a focus on its individual parts. With this in mind it is clear that the areas of technical implementation, prompt submission and hybrid integration are candidates for future studies that aim to develop a focused, in-depth understanding of how these areas could be enhanced, utilised and progressively built upon to improve the experience for both the practitioners and the audience.

## ACKNOWLEDGMENT

The authors would like to thank FANDCO and Mary Stewart-David for the development and performing of the House of Mask production, and StoryFutures for funding the research and development of the production.

## REFERENCES

- Alsop, T. (2024) *VR headset unit sales worldwide 2024*, Statista. Available at: <https://www.statista.com/statistics/677096/vr-headsets-worldwide/>.
- Boosted. (no date). *Crowdfunding Arts in New Zealand* (no date) Boosted. Available at: <https://boosted.org.nz/projects/theatre-live-online>
- Clarke, V. and Braun, V. (2016) ‘Thematic analysis’, *The Journal of Positive Psychology*, 12(3), pp. 297–298. doi: 10.1080/17439760.2016.1262613.
- Deakin, F. (no date) *The galaxicle implosions*, Fred Fandco. Available at: <https://freddeak.in/project/the-galaxicle-implosions/>.
- Fandco (no date) *FANDCO site*. Available at: <https://www.fandco.xyz/>
- Get started with live streaming - computer - youtube help* (no date) Google. Available at: <https://support.google.com/youtube/answer/2474026?hl=en-GB&co=GENIE.Platform%3DDesktop>.
- Google Forms (no date) *Google forms: Online form creator | google workspace*. Available at: <https://www.google.com/forms/about/>
- Hannah, D. (2024) ‘(I M) m a T E R (I a L I T Y) and the black-box theatre as an “empty space” of re-production’, *idea journal*, 4(1), pp. 23–33. doi: 10.37113/ideaj.vi0.229.
- Kurt, I. (2024) ‘Covid-19 and sociovirtualization: Exploring new ways to socialize’, *London Journal of Social Sciences*, (7), pp. 23–36. doi: 10.31039/ljss.2024.7.186.
- Lock, D. and Kirman, B. (2023) ‘Identifying the developmental challenges of creating virtual reality theatre’, *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* [Preprint]. doi: 10.1145/3573381.3603359.
- Unreal Engine 4. (no date). *The most powerful real-time 3D creation tool - unreal engine* (no date) Unreal Engine 4. Available at: <https://www.unrealengine.com/en-US>.
- Vyas, L. (2022) “‘New normal’ at work in a post-COVID world: Work–life balance and labor markets’, *Policy and Society*, 41(1), pp. 155–167. doi: 10.1093/polsoc/puab011.
- Zaunbrecher, N. J. (2011) ‘The elements of improvisation: Structural tools for spontaneous theatre’, *Theatre Topics*, 21(1), pp. 49–60. doi: 10.1353/tt.2011.0015.
- Zoom (no date) *One platform to connect*, Zoom. Available at: <https://www.zoom.us/>

# Formative Usability Assessment of a Rehabilitative Hand Exoskeleton – Directions for User-Friendly Physical Interfaces

Sedef Süner-Pla-Cerdà<sup>1</sup>, Batuhan Şahin<sup>1</sup>,  
and Kutluk Bilge Arıkan<sup>1,2</sup>

<sup>1</sup>Department of Industrial Design, TED University, Ankara 06420, Türkiye

<sup>2</sup>Department of Biomedical Engineering, Neurotechnology Center of Excellence (NÖROM), Ankara University, Ankara 06560, Türkiye

## ABSTRACT

This paper presents a formative usability test study involving eight healthy individuals to identify user-centred criteria and directions for improving the usability of a rehabilitative hand exoskeleton system. The formative usability test was applied as a procedure accompanying the motor learning tests for system validation. The test adopted a qualitative approach combining structured observations during exoskeleton use to complete motor control tasks. The observations were followed by semi-structured interviews immediately after use. Qualitative findings from the formative usability tests revealed issues related to use comfort, wearability, simplicity and perceived safety of the proposed exoskeleton system. Based on these findings, practical design recommendations are provided to enhance the donning and doffing of the device, adjustability of finger connections to accommodate anthropometric ranges, material selection and component layout for improved physical comfort. A finger ring system designed to improve physical ergonomics and usability is introduced. The outcomes of this study are expected to contribute to both the usability improvements of the current system and serve as a reference to the research community in general while developing user-friendly physical interfaces for wearable robotics.

**Keywords:** Rehabilitation, Exoskeleton, Assistive technologies, Wearable robotics, User requirements, Usability assessment, User-centred design

## INTRODUCTION

Stroke ranks as the second most common cause of mortality worldwide and the third most common cause of disability. In Europe, stroke affects 1.1 million individuals and results in 440,000 fatalities annually (Béjot et al., 2016). It is believed that current strategies will be ineffective in lowering these rates in the next years. The major factor is the increasingly ageing population. The number of stroke incidents is predicted to approach 1.5 million by 2025, with the number of persons affected by stroke increasing by 27% between 2017 and 2047 (Wafa et al., 2020). Despite recent major pharmacological advances in stroke treatment, there is no feasible, successful, or permanent

medical treatment for stroke. Post-stroke rehabilitation interventions are commonly used in stroke therapy and research is increasingly focusing on the rehabilitation process. The objective of rehabilitation is to improve impaired function and attain the best level of independence feasible within the constraints of chronic stroke-related impairments. Post-stroke physical therapy is a prolonged treatment that requires the involvement of a therapist.

Robot-assisted training is a modern neurorehabilitation approach that has shown efficacy in stroke patients. Robots are commonly utilised by clinicians in stroke rehabilitation because they allow the user to carry out highly repetitive actions precisely. Robotic training with exoskeletons has shown promise in the recovery of motor functions within clinical rehabilitation settings (Prange et al., 2006). Hand exoskeletons, a sub-category of such wearable robotic devices, aim to aid patients in regaining their motor functions. These robotic devices are designed to manipulate the joints of the fingers, primarily for rehabilitation and/or interaction.

Current hand exoskeleton systems pose numerous usability issues due to challenges stemming from the system complications dictated by the complexity of hand kinematics and the diverse tasks it performs (Almanera et al., 2017; Zhu et al., 2022). Size, bulk and weight are among those complications governing most rigid exoskeletons, negatively affecting the devices' comfort, adjustability, portability and wearability (Ferguson et al., 2019). Many systems have low technology readiness levels, posing challenges to acceptability, marketability and home deployment (Martinez-Hernandez et al., 2021).

Engineers encounter difficulties in selecting suitable components and designing control systems to address kinematic complexities, adaptability to various hand sizes, and the need to support different movements and tasks. These exoskeletons require various actuators such as pneumatic, ultrasonic, and DC motors, with considerations for torque, power requirements, and control system design (Kabir et al., 2022). Moreover, post-stroke rehabilitation efforts emphasise functional recovery through motor re-learning, highlighting the importance of effective control mechanisms in robotic rehabilitation (Agarwal & Deshpande, 2015; Levin & Demers, 2020; Marchal-Crespo & Reinkensmeyer, 2009; Sans-Muntadas et al., 2014; Washabaugh et al., 2018).

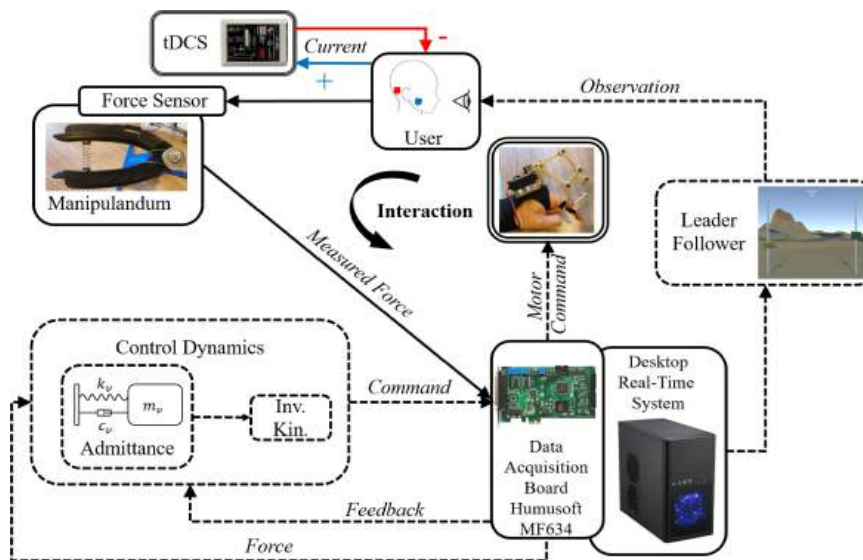
Despite the technical advancements, the commercial availability of hand exoskeletons remains limited, with most systems at prototype levels. User involvement in the development process is crucial to ensure better usability, satisfaction, and acceptance of such technologies. However, understanding primary user requirements and end-user involvement in the design process is limited (Hill et al., 2017; Kobbelgaard et al., 2021). Issues related to physical ergonomics such as comfort, portability, and wearability are recognised concerns (see, for example, Almanera et al., 2017; Ambrosini et al., 2014; Baltrusch et al., 2020; Majidi Fard Vatan et al., 2021; Martinez-Hernandez et al., 2021), but subjective user perceptions on these aspects and other design requirements are often overlooked.

This paper focuses on identifying user-centred criteria and directions for improving the usability of a rehabilitative hand exoskeleton system. The proposed robotic exoskeleton is a two-degree-of-freedom, fully actuated

system. It is designed for the index finger using an optimisation technique that minimises a cost function which is composed of the isotropy measure and the required actuator torque. The rest of the paper introduces the system design and methodology of the formative user evaluation. Recommendations are made to improve the usability of the system design from a user-centred perspective.

## SYSTEM DESIGN

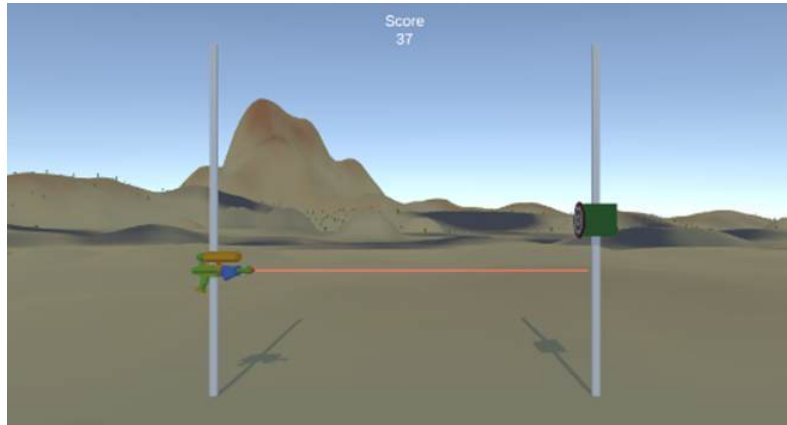
Our project focuses on designing an exoskeleton robot specifically for rehabilitating the pinching action by index finger and thumb, which is particularly challenging to regain after a stroke. The design seeks to facilitate the process of motor relearning by implementing a control system strategy and a specific actuation type in the exoskeleton mechanism. We hypothesise that the patient needs various forms of interaction with the robotic exoskeleton during the whole rehabilitation period to maximise motor relearning. The control system can implement kinematic control, interaction-based control, and force control, which may be applied as either an assistive or resistive technique. The exoskeleton can be fully or under-actuated. The under-actuated form facilitates motor variability, which is essential for motor learning. In this study, the exoskeleton is used in the fully actuated form and the control is the admittance type of interaction control. The overall system architecture is presented in Figure 1.



**Figure 1:** System architecture.

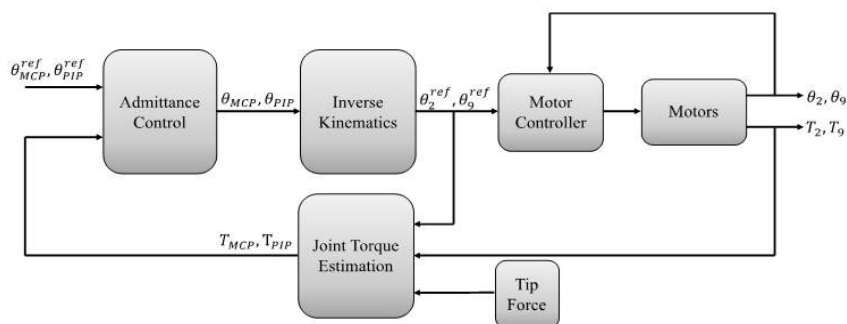
The leader-follower game is the main element of the system. Users produce an upward force in this game by pinching the elastic interface, generating an interaction force. This force drives the avatar of the follower in the upward direction. Likewise, when users let off the interface, a simulated gravitational force causes their avatar to go downward. Users attempt to

strike the vertically oscillating target that follows a predetermined pattern. The pattern is composed of the cumulative sum of three distinct frequency harmonics: 0.07 Hz, 0.2 Hz, and 0.25 Hz. The game interface, created using Unity (Unity Technologies, US), is seen in Figure 2.



**Figure 2:** Leader-follower game in unity: the left avatar is the follower, and the right avatar is the leader.

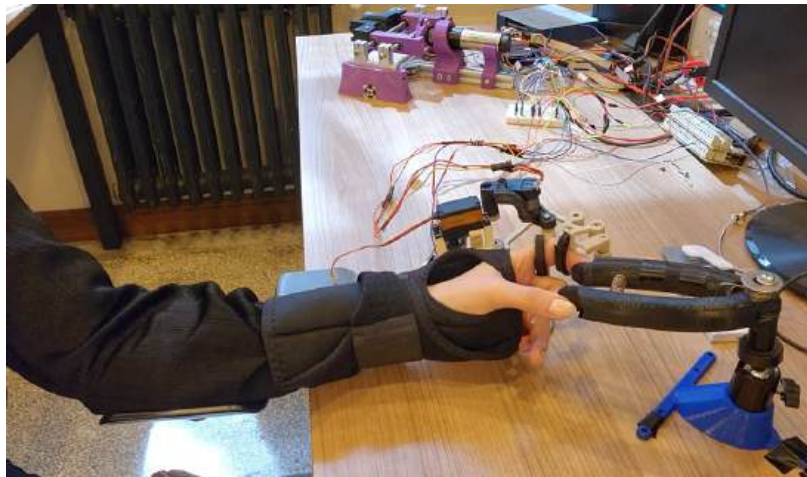
The elastic pinching interface (see Figure 4) connects the user and the game. The device uses a force sensor that quantifies the exerted force by the user's index finger at the tip. The force signal is sampled at 2 kHz and sent to a desktop computer running Simulink® Desktop Real-Time 2021b simulation via a data acquisition device (DAQ) (Humusoft MF634). The exoskeleton controls the metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints of the index finger. The controller uses an admittance-based interaction control. The controller operates to follow the voluntary finger motions of the participant and uses virtual elastic, damping and inertial forces to restore the finger to its normal position. This system may simply be converted into an assistive or resistive strategy. The control system architecture is given in Figure 3.



**Figure 3:** Admittance type control system.

## METHODOLOGY

A preliminary user evaluation study was designed to test the current exoskeleton developed for system validation. We planned a user feedback protocol to test the current version of the system with healthy users and evaluate it in terms of ease of use, comfort and perception, and to plan a design revision based on these evaluations. Eight healthy individuals, all sophomore undergraduate students at TED University in Ankara, Turkey, volunteered to participate in the study. All participants signed an informed consent form before the test.



**Figure 4:** Test setup and video-recording angle.

### Test Protocol

User evaluations included observing the participants using the exoskeleton for approximately 15–20 minutes, followed by a short post-test interview. Each test lasted approximately 30–40 minutes including preparation, use and interview. Use and interview sections were video-recorded. Test setup and video recording angle are shown in Figure 4. The post-interview questions were based on the literature review conducted before the tests, which involved four key user dimensions: comfort in use, wearability, simplicity, and perceived safety. Further observations are made for adjustability. Detailed test procedure is presented in Table 1.

### FINDINGS

Three hours of videos were recorded during the use and interview phases of the eight test sessions. These recordings were watched by one researcher (industrial designer specialised in user research) to identify and document the usability issues observed and stated by the participants. This section presents the identified issues regarding comfort in use, wearability, simplicity, and perceived safety.



**Table 1.** Test procedure.

Test phase	Duration	Activities
Preparation	5–10 minutes	<ol style="list-style-type: none"> <li>1. Participant reads and signs the consent form.</li> <li>2. Participant is introduced to the setup and the test procedure as follows: where and how to sit, introduction and a short demo of the game.</li> <li>3. Participant's arm is placed on the elbow rest and the placement and the controller are adjusted.</li> </ol>
Use and observation	15–20 minutes	<ol style="list-style-type: none"> <li>4. Participant plays the game without the exoskeleton.</li> <li>5. Researcher helps wear the exoskeleton, and the participant plays the same game with it.</li> <li>6. Participant is asked to take off the exoskeleton without help. <i>Observed interactions:</i> Difficulties when donning and doffing; Placement of the fingers and wrist inside the device; Stability of the device and if the participant wants to fix it; If the device affects the hand posture; Difficulties while taking the device off alone</li> </ol>
Post-interview	5–10 minutes	<ol style="list-style-type: none"> <li>7. Open-ended questions are asked of the participant. <i>Interview questions:</i> How did you feel using the exoskeleton, and why? What are your comments on its comfort in use? Have you encountered any difficulties while taking it off, can you show it on the device? What are your comments on its general look and structure? Do you think it is safe to use it, and why? Do you have additional comments or suggestions?</li> <li>8. The participant plays the game one last time without the exoskeleton to complete the performance test, and the procedure is concluded.</li> </ol>

### Comfort in Use and Wearability

The vast majority of the findings point out issues with use comfort, including the method for placing and fastening the finger pieces, weight and balance of the exoskeleton, and the material.

- The current design utilises a standard-size, 3D-printed finger placement form. These finger pieces are fastened to the middle and proximal phalanges via Velcro straps attached to two protrusions on semi-rings. During use, these protrusions are observed to be rubbing against the middle finger, causing physical discomfort and limiting the movement of the user.
- The design of the finger piece does not accommodate the anthropometric variety in finger size. This issue is tried to be overcome by fastening with the Velcro straps, but the lack of a firm finger grip creates gaps between the phalanges and the finger pieces. These gaps were filled with foam pieces, which frequently came loose, disrupting the testing procedure.
- The pulling force on the middle phalanx led to discomfort and redness on the skin. This was also expected to be prevented with an additional foam layering between the finger and finger piece.
- The skeleton part of the device is lightweight since it is made of 3D-printed plastics. However, the motor and actuator mechanism creates load

and discomfort on the wrist and arm during use (Figure 5). Half of the participants expressed discomfort after nearly 15 minutes of use. Although they didn't think the pain and fatigue were unbearable, the impact could be worse for patients during longer rehabilitation sessions.

- Since comfort was not a priority in the current design, the robotic mechanism was attached to a standard wrist splint. Although it is easy to put on the splint, fastening it requires rotating the hand and wrist around their own axis. Some participants found this uncomfortable due to the weight of the mechanism. Additionally, some reported that the synthetic material of the splint causes sweating after use.
- During the tests, a researcher helped don the device, and the participants were asked to take it off themselves. Although doffing the device is relatively easy, correct placement and fastening of the finger pieces took time and effort during donning the device (Figure 6).
- Wearability is related to comfort in longer use as much as it does to ease of donning and doffing the device. The above-mentioned issues with weight balance, material and form are factors affecting wearability.



**Figure 5:** The weight of the motor and actuator creating load on the wrist and arm.



**Figure 6:** Donning with help (left) and doffing without help.

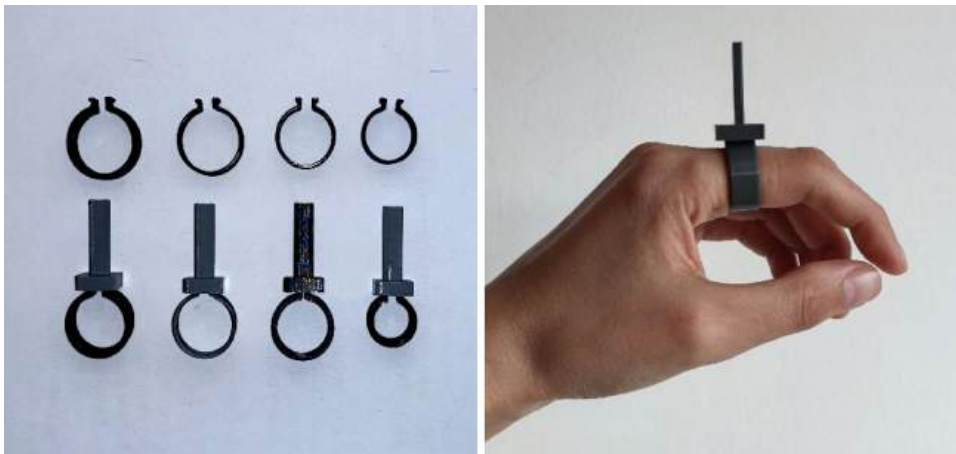
### Simplicity and Perceived Safety

Other issues identified are related to the perception of the device rather than physical interaction:

- Participants drew attention to the visual issues, particularly the fully open structure of the skeleton, motor and actuator mechanism, and the fact that the cable connections are disorganised. Some of the phrases they used to describe the visual language of the setup are: “too mechanical”, “messy”, “exposed”, and “difficult to understand”.
- Some participants emphasised the disorderly look of the protrusions of the finger pieces to fasten the Velcro straps.
- Despite the critique of the look, some participants stated that these were their first impressions, and it did not matter as much while using it because they discovered that it was not as scary as it looked and while playing the game they no longer looked at the skeleton.
- A few participants stated that the sudden finger pull of the skeleton at the beginning of the test felt unsafe. Although it is not part of the exoskeleton design, it has an impact on the user experience.

### ACTIONS FOR IMPROVEMENT

Focusing on comfort and wearability, we prioritised the development of a ring system for easy donning and doffing, secure fastening, and preventing discomfort. The system consists of individual rings worn in fingers, inserted in the nests located on the skeleton by leveraging material flexibility and interlocking tabs (Figure 7). We have completed first iterations and currently adapting the nesting piece to the exoskeleton.



**Figure 7:** First iterations of the ring system design.

As adjustability to various hand sizes is vital, the following steps will include designing a set of ring components based on an existing data set for hand anthropometry (Cakit et al., 2012). The data set will be used to develop an ergonomic design approach combining ‘design for extremes’

(5th percentile female and 95th percentile male) and ‘design for adjustable range’ strategies to accommodate the largest possible population (McCauley, 2012).

Further improvements can be made for a balanced distribution of the weight of the robotic setup to mitigate the load and discomfort on the wrist and arm. In addition to relocating the motor and actuator mechanism without obstructing the proper functioning, it is possible to place an adjustable base to the setup to support the wrist and minimise the load of the exoskeleton.

The standard wrist splint will be replaced by a custom-designed, fingerless guard glove to mitigate sweating and weight-pulling. A breathable material such as knitted polyester can be used as the main material, supported by 3D-printed attachments to fasten the exoskeleton. Polyester is widely preferred in sports gear due to its ability to facilitate sweat dispersion.

Visual simplicity and acceptability are important concerns raised by rigid exoskeletons as also emphasised in the literature. Being a rehabilitation device, system performance is naturally prioritised in the design and development of wearable rehabilitation robots, therefore device form is often overlooked. Although the skeleton structure cannot be interfered with simply based on visual concerns, disguising, collecting and orderly connecting the cables to the system components and power source can significantly contribute to the visual simplicity. Finally, replacing the current finger pieces with the newly developed ring system design is expected to improve the aesthetic perception of the device.

## **CONCLUSION**

This paper presents our findings of the formative usability assessment of a fully actuated rehabilitative hand exoskeleton. The user evaluations were conducted simultaneously with the motor tests carried out with an initial setup. Early involvement of users helped identify major usability problems that can be solved and integrated into the upcoming system revisions. This required the collaboration of a team of engineers and industrial designers, creating opportunities for the improvement of the user-friendliness of the device without compromising its performance. Future work will include integrating the planned changes into the new prototype iterations, and structured usability tests to assess comfort, wearability and perception of the device.

## **ACKNOWLEDGMENT**

This research was funded by the Scientific and Technological Research Council of Türkiye (TÜBİTAK), grant number 121E107. The authors would like to thank Amr Okasha for his contributions to the testing procedure.

## REFERENCES

- Agarwal, P. & Deshpande, A. D. (2015). Impedance and force-field control of the index finger module of a hand exoskeleton for rehabilitation. *Proceedings of IEEE International Conference on Rehabilitation Robotics*, 85–90.
- Almenara, M., Cempini, M., Gómez, C., Cortese, M., Martín, C., Medina, J., Vitiello, N., & Opisso, E. (2017). Usability test of a hand exoskeleton for activities of daily living: an example of user-centered design. *Disability and rehabilitation: Assistive technology*, 12(1), 84–96.
- Ambrosini, E., Ferrante, S., Rossini, M., Molteni, F., Gföhler, M., Reichenfeller, W., ... Pedrocchi, A. (2014). Functional and usability assessment of a robotic exoskeleton arm to support activities of daily life. *Robotica*, 32(8), 1213–1224.
- Baltrusch, S. J., Houdijk, H., van Dieën, J. H., van Bennekom, C. A. M., & de Kruif, A. J. T. C. M. (2020). Perspectives of End Users on the Potential Use of Trunk Exoskeletons for People with Low-Back Pain: A Focus Group Study. *Human Factors*, 62(3), 365–376.
- Béjot, Y., Bailly, H., Durier, J. & Giroud, M. (2016). Epidemiology of stroke in Europe and trends for the 21st century. *Presse Medicale*, 45(12), 391–398.
- Cakit, E., Durgun, B., Cetik, O., & Yoldas, O. (2012). A Survey of Hand Anthropometry and Biomechanical Measurements of Dentistry Students in Turkey. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 24(6), 739–753.
- Ferguson P. W., Shen, Y. & Rosen, J. (2019). Hand Exoskeleton Systems—Overview. In Jacob Rosen & Peter Walker Ferguson (eds.), *Wearable Robotics: Systems and Applications* (pp. 149–175). Elsevier Academic Press.
- Hill, D., Holloway, C. S., Morgado Ramirez, D. Z., Smitham, P., & Pappas, Y. (2017). What are user perspectives of exoskeleton technology? A literature review. *International Journal of Technology Assessment in Health Care*, 33(2), 160–167.
- Kabir, R., Sunny, M. S. H., Ahmed, H. U. & Rahman, M. H. (2022). Hand Rehabilitation Devices: A Comprehensive Systematic Review. *Micromachines*, 13(7):1033.
- Kobbelgaard, F. V., Kanstrup, A. M., & Struijk, L. N. S. A. (2021). Exploring User Requirements for an Exoskeleton Arm Insights from a User-Centered Study with People Living with Severe Paralysis. In C. Ardito, R. Lanzilotti, A. Malizia, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, & K. Inkpen (Eds.), *Proceedings of the Human-Computer Interaction – INTERACT 2021, Part I* (pp.312–320). Springer.
- Levin, M. F. & Demers, M. (2020). Motor learning in neurological rehabilitation. *Disability and Rehabilitation*, 43(24): 3445–3453.
- Majidi Fard Vatan, H., Nefti-Meziani, S., Davis, S., Saffari, Z. & El-Hussieny, H. (2021). A review: A Comprehensive Review of Soft and Rigid Wearable Rehabilitation and Assistive Devices with a Focus on the Shoulder Joint. *Journal of Intelligent & Robotic Systems*, 102:1–24.
- Marchal-Crespo L, Reinkensmeyer DJ. Review of control strategies for robotic movement training after neurologic injury. *NeuroEngineering and Rehabilitation*, 6: 1–15.
- Martinez-Hernandez, U., Metcalfe, B., Assaf, T., Jabban, L., Male, J., & Zhang, D. (2021). Wearable Assistive Robotics: A Perspective on Current Challenges and Future Trends. *Sensors*, 21(20), 6751.
- McCauley, P. (2012). *Ergonomics: Foundational Principles, Applications, and Technologies*. CRC Press.
- Prange, G. B., Jannink, M. J., Groothuis-Oudshoorn, C. G., Hermens, H. J., & Ijzerman, M. J. (2006). Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke. *Journal of rehabilitation research and development*, 43(2), 171–184.

- Sans-Muntadas, A., Duarte, J. E., & Reinkensmeyer, D. J. (2014). Robot-assisted motor training: assistance decreases exploration during reinforcement learning. In *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3516–3520).
- Wafa, H. A., Wolfe, C. D. A., Emmett, E., Roth, G. A., Johnson, C. O., & Wang, Y. (2020). Burden of Stroke in Europe: Thirty-Year Projections of Incidence, Prevalence, Deaths, and Disability-Adjusted Life Years. *Stroke*, *51*(8), 2418–2427.
- Washabaugh, E. P., Treadway, E., Gillespie, R. B., Remy, C. D., & Krishnan, C. (2018). Self-powered robots to reduce motor slacking during upper-extremity rehabilitation: A proof of concept study. *Restorative neurology and neuroscience*, *36*(6), 693–708.
- Zhu M, Biswas S, Dinulescu SI, et al. Soft, Wearable Robotics and Haptics: Technologies, Trends, and Emerging Applications. *Proceedings of the IEEE* (pp. 246–272).

# Exploring the Potentials of Wearable Devices in Research on Wellbeing and Stress in Workplaces

**Margherita Pillan and Isabella Ruina**

Dipartimento del Design, Politecnico di Milano, Via G. Durando 10, 20158 Milan, Italy

## ABSTRACT

Work-related wellbeing and stress are major research issues nowadays. Stress at work affects employee's mental and physical health and reduces productivity. Since the definitions of stress and wellbeing are various, different strategies to investigate the problem and promote solutions have been taken. This study was developed within the innovation ecosystem MUSA (Multilayered Urban Sustainability Action) in the challenge of Spoke 2, Big Data-Open Data in Life Sciences, which aims to create solutions for the collection, conservation, and process of big data to improve lifestyle, prevention, and treatment. The study is grounded on the requirement to meet the complexity of stress and wellbeing conditions, therefore integrating qualitative (EMA questionnaire diary) with quantitative (wearable device Fitbit) data. This article presents the preliminary studies with the principles and recommendations leading to the final design of the experiments included in the research.

**Keywords:** Work-related stress, Wellbeing, Monitoring system, Wearable devices, Interaction design, Design principles

## INTRODUCTION

Work-related stress has emerged as a significant concern in contemporary society, affecting nearly one in three workers across Europe (Leka, 2010).

Stress at work impacts the mental and physical well-being of employees and leads to diminished productivity within companies. Work-related mental and physical wellbeing is important for individuals and organizations, impacting life quality, long and short-term health conditions, performances, commitment, availability, and creativity (WHO, 2022).

Research from different disciplines aims at identifying strategies to reduce strain and stressors, improve work organization and conditions, and favour wellbeing. The literature indicates a variety of situations and definitions associated with health, wellbeing, and stress, distinguishing hedonic from eudaimonic wellbeing, eustress from distress, i.e. positive and negative work engagement (Sonnetag, 2022). Several factors can potentially act as stressors and negatively affect health (Michie, 2002; Davis, 1989). After the COVID pandemic (EUA, 2024), the importance of researching for work wellbeing became even more cogent. The changes of the working processes and the

pervasive use of digital tools also impact wellbeing (Juchnowicz, 2021; Tams, 2020).

The investigation of the work-related effects on wellbeing is an important research focus, asking for new approaches for collecting objective data to allow a deeper understanding of the correlation between stressors and stress. According to The Lancet, “work and employment are an underutilised lever to influence population health” (The Lancet, ed., 2023). A specific goal for Italy, is “*adapt monitoring of working conditions to the changing world of work, focusing on psychosocial risk factors. ii) Harmonise data on working conditions from records and registers for use in the national Information System for Prevention in the Workplace and expand the system’s capture of psychosocial risk factors*” (Pega, 2023). Research for happiness at work is needed (Misra, 2023) for individual and common advantage.

This study was developed within the innovation ecosystem MUSA – Multilayered Urban Sustainability Action – project, funded by the European Union – NextGenerationEU, under the National Recovery and Resilience Plan (NRRP) Mission 4 Component 2 Investment Line 1.5: Strengthening of research structures and creation of R&D “innovation ecosystems”, set up of “territorial leaders in R&D”. The Spoke 2 of the project, Big Data-Open Data in Life Sciences, aims at developing solutions to enable the collection, conservation, and processing of big data to improve lifestyle, prevention, and treatment. This document reports the preliminary study producing the strategies for conducting experimental investigations on the acceptability and efficacy of wearable devices in real office work environments. The experiments aim to collect and correlate objective physiological data with subjective experience during working hours. In doing so, it becomes crucial to find the right balance between gathering a sufficiently extensive dataset and avoiding imposing excessive burdens on the participants (Weale, 2023). The document includes a chapter dedicated to the complexity of defining and measuring wellbeing and stress; a part dedicated to the potential role of wearable devices in the experiments; design principles and strategic drivers for the research.

## **THE COMPLEXITY OF DEFINING STRESS AND WELLBEING**

Wellbeing and stress have been studied for a long time, yet their relationship is still complex. A univocal definition of wellbeing is currently missing (Juchnowicz, 2021), and the same is true for wellbeing in the work context. Different approaches have been used for framing wellbeing: some define it by positive/negative outcomes (e.g. job satisfaction, work engagement, ...) (Bordi, 2018); others distinguish from subject to objective dimensions (Bertoloni, 2016); and some consider its multidimensional structure composed by physical, psychological and social factors (Zani, 1999). Wellbeing and stress are two intertwined concepts, but the presence of one does not involve the absence of the other. According to the World Health Organization, “health is a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity” (WHO, 1948). Shifting to the work context, one diffused conceptual model is the



J-DR by Demerouti et al. (Demerouti, 2001). According to the authors, work characteristics can be classified into two categories: demands (from the environment) and resources (owned by the individual). The equilibrium between demands and resources leads to *wellbeing*, while the disequilibrium is associated with *illbeing* (stress) (Bordi, 2018). Adding to this model a deeper interpretation, Lu et al. (Lu, 2021) distinguish between *sustress* (inadequate stress), *eustress* (good stress), and *distress* (bad stress). Eustress is marked by an increased pulse rate without any underlying feeling of threat or fear (Gedam, 2021). If one demand is perceived as a mild challenge, where resources are sufficient to deal with it, the body's response is positive, generating eustress. This means that stress is not always bad and strengthens the idea that wellbeing and stress are not two opposite concepts. Consistent with this approach, Abreu et al. (Abreu, 2002) consider stress as the "psychological and physical state that results when the resources of the individual are not enough to deal with the demands and pressures of the situation". This means that, since stress is a psychophysiological state, the integration of physical and psychological measures is crucial for the assessment of mental health. According to this idea, this article proposes an integrated approach based on EMA questionnaire and the use of the wearable device Fitbit. EMA, Environmental Momentary Assessment (Stone, 1994), is a research method for contextual measurement while reducing recall bias (Shiffman, 2008). Physiological data are collected through the Fitbit device to integrate the body's physical response into the framework.

## WORK-RELATED STRESSORS

The identification of the job stressors is supported by theories and the scientific literature deploys a wide list of possible stressors in the work context. Here is a list of the three leading models to define stress:

- *J-DR model* by Demerouti et al. (Demerouti, 2001), mentioned above.
- *Effort-reward imbalance – ERI model* by Siegrist et al. (Siegrist, 1996). According to this theory, a job is a contract on what the worker gives (effort, time), and what he/she receives (money, esteem, career opportunities). When this relationship is imbalanced, distress might arise.
- *The transactional model* by Lazarus (1966) suggests that stress results from a dynamic interaction between the individual and the environment. The model starts from the same point as the J-DR model. Still, it adds, as an additional factor, the subjective perception of the worker about work demands and personal capabilities as well as the resources to deal with these demands. Since this perception can vary between individuals (Probst, 2010) and is influenced by personal traits, previous experience, and other factors, any aspect of the work environment can be perceived as a stressor.

Job stressors can be categorized from different perspectives. Thilagavathy and Geetha (Thilagavathy, 2021), distinguish individual, organisational, and social factors influencing wellbeing at work. The *Office Environment Model* (Bluyssen, 2011), identifies factors related to work context/physical environment; work context/social environment; worker health context.

Lukan et al. (Lukan, 2021) add a distinction between daily stressors and structural conditions. Lists of the main aspects influencing wellbeing at work have been reported by several authors (Misra, 2023; Susanto, 2022; Lilja, 2020; Tams, 2020; Bertoloni, 2016; Thilagavathy, 2021). The lists include job characteristics; Job satisfaction/content; job performance; meaningfulness of work; work variety; work role (e.g. role ambiguity, role conflict); effort-reward imbalance; work-life balance; engagement in work; relationship with others (colleague and supervisor) (e.g. social support); career opportunities/development; job control and autonomy/decision; physical environment; work-load. Finally, a relatively new factor generated has to be considered: the digitalisation of work. Digital tools make work more flexible while increasing autonomy (e.g., smart working). On the other hand, some relevant issues emerge, related to the possibility of being always connected and available, which blurs the line between work and private life (Bordi, 2018). The digitalisation of work is associated with a new term, technostress, a specific type of work stress that can cause anxiety, fatigue, skepticism, and inefficacy associated with the use of technology (Salanova, 2014).

### **CAN LOW-COST WEARABLE DEVICES BE EFFECTIVE IN RESEARCH ABOUT STRESS AND WELLBEING AT WORK?**

Due to size/cost reduction and improved ease of use (Di Flumeri, 2019), the use of wearable devices to monitor and assess physiological and mental activity has recently spread (Haghi, 2017). Investigations on stress response often take place in laboratory settings where participants are exposed to standardized stressors. Conducting such experiments is costly and demanding for participants, limiting the feasibility of large-scale assessments (Pakhomov, 2020). On the other side, wearable devices can be easily used to record bio signals without interfering with participant activities (Giorgi, 2021), making it possible to operate long-term and real-life stress monitoring (Stojchevska, 2022). Compared to gold-standard equipment, consumer wearable devices showed similar accuracy in measuring different biomarkers in different conditions (Menghini, 2019). Studies that aim to identify stress commonly depend on physiological reactions of the sympathetic nervous system triggered by stress, such as alterations in Heart Rate (HR), Heart Rate Variability (HRV), Skin Temperature (SK), and Electrodermal Activity (EDA) (van Kraaij, 2020; Hickey, 2021). Integrating more of these parameters is preferable to obtaining valid results (Gedam, 2021). Among all the physiological parameters, HRV is the most studied one (Peake, 2018), as it is the most useful physiological metric for stress detection (Hernando, 2018; Hong, 2010; Rodrigues, 2018). Unfortunately, some devices on the market use average HR to monitor stress conditions, albeit this parameter is not as accurate as HRV. This is the case of Fitbit, an accessible fitness tracker used in different research to assess stress (Pakhomov, 2020; Chalmers, 2022). Although the authors of such studies proved that HR measurements obtained with Fitbit increase as expected in response to stressors, they were not able to determine the accuracy of the measure (Pakhomov, 2020) and

failed to identify consistent HR patterns changes during stress (Chalmers, 2022). Giorgi and colleagues (2021) aimed to assess the reliability of two other wearable devices (i.e., Empatica E4) in detecting different mental states. Even if the result was positive, some consideration had to be given. The Empatica E4 is a high-level wearable device, and it is sold at a high price. This condition may represent a constraint, limiting its adoption to large-scale scientific research.

Two important factors have to be considered when designing a protocol to investigate mental health conditions using the wearable device: (1) to obtain a valid result, the physiological parameters must be measured for a long real-time period; (2) the signal recorded can be altered by contextual factors, such as posture, temperature, and physical activity (Wijsman, 2011). For example, physical exercise can increase HR and change ST, even if it is not correlated with stress triggers (Nelson, 2014). This is the case of Fitbit, when the participant remained stationary and could not differentiate between stress caused by physical activity or mental burden (Pakhomov, 2020). The use of wearable devices to detect stress conditions represents a major opportunity due to their low impact and accuracy compared to gold-standard equipment (Menghini, 2019). Some limitations still emerge when choosing which type of device, and the selection of the suitable device should accord with the following requirements:

- The acceptability of the technology, directly connected with the perceived usefulness and the perceived ease of use, as described in the Technology Acceptance Model (TAM) (Davis, 1989).
- The cost of the device, which can limit the scalability of the experiment to an extended cohort.
- The reliability of the device: different wearables monitor different parameters to assess stress; HRV vs HR are physiological measures with different effectiveness as indicator of stress.

Based on the literature analysis summarized in this document, the possibility of obtaining accurate stress measurement through low-cost wearable devices seems limited. Despite that, experiments to assess stress conditions employing low-cost devices are interesting because they support the assessment of methodologies that can be adopted in large cohort future experiments with more expensive devices (e.g. Fitbit for wristband-type devices). As technology advances, the costs are expected to decrease while keeping the quality of the data collected (Haleem, 2023), eliminating the trade-off between the reliability of the measures and the scalability of experiments.

### **AWARENESS AND BEHAVIOURAL CHANGE INDUCED BY SELF-MONITORING SYSTEMS**

Experiments based on wearable devices collecting personal data can affect participants' behaviour and mental health self-awareness due to the self-monitoring effects. The potential of self-monitoring lies in the possibility of mirroring themselves (Varisco, 2019), making people more aware of

personal habits. Using a device to monitor personal health conditions can help people to correlate physiological measures with subjective experience, thus promoting self-awareness and stimulating reflection (Chianella, 2021). Moreover, thanks to the collection of data framing past and current personal health states, individuals can identify and modify potentially unhealthy behaviours (Mercer, 2016). Indeed, in line with Oinas-Kukkonen, definition (Oinas, 2013), wearables can be described as “*socio-technical information systems with psychological and behavioral outcomes designed to form, alter or reinforce attitudes, behaviors or an act of complying without using coercion or deception*”.

When considering behaviour change, both positive and negative outcomes should be considered. On the one hand, the possibility of setting goals, receiving motivational reminders, tracking progress, and obtaining contextualized user data can lead to positive behaviour change (Kang, 2022). Conversely, the constant surveillance imposed by the wearable can also bring negative outcomes, such as feelings of disempowerment.

The collection, processing, conservation, and sharing of personal data can impact individual and collective levels. Sharing personal data can change the image of people in the community where they live or work and it can, therefore, impact personal relationships and opportunities (Varisco, Pillan, 2019; Pillan, 2017) and the exercise of privacy rights. The acceptability of smartwatches for medical purposes has also been investigated. Al-Marroof et al. (Al-Marroof, 2021) point out the importance of the perceived usefulness and ease of use for the acceptability of these devices. Shandi et al. (Shandi, 2024) reported data on the dependability of the acceptance and adoption of these solutions on factors such as age, education, occupation, economic status, fitness levels, and health conditions.

## **PARTICIPATORY DESIGN FOR RESEARCH**

Research on stress and wellbeing involving personal data in working environments is a complex task, calling for the accurate design of experiments with a focus on the fair involvement of people. The development of new approaches for research on health and wellbeing, allowing the collection of functional data on bio-parameters correlated to contextual events, requires suitable strategies addressed to the acceptability of the final users. In this scenario, the adoption of participated design methodologies and codesign are promising to cope with the delicacy of collecting and using personal data through wearable devices, as reported by some authors (Perego, 2022; Bajaj, 2023). Participatory design for health research can involve therapists, technicians, and final users of the solutions adopted in the investigations. It can concern factors such as the contents and goals of the experiments, the design of devices and applications, and the requirements for accessibility, usability, and security (Jones, 2020). This approach can be successfully adopted in the design of programs for eHealth and wellbeing (van Hierden, 2021) and in the reflection of the ethical principles in research involving different actors (Sleigh, 2022). The long tradition of user studies and participatory design in the Interaction Design discipline provides

the theoretical background and expertise on prototyping and tests to create optimal participated design processes.

## CONCLUSION

Wearable devices have great potential in research about stress and in collecting physiological measurements that may indicate unhealthy conditions. To effectively investigate wellbeing at work, the physiological data should be associated with the qualitative data on users' experiences. Wearable devices alone, in fact, could be ineffective in catching mental stress, that can instead be better investigated through the correlation of qualitative with quantitative data.

When selecting the type of device, the important factors to be considered are their acceptability and cost, and the type of data collected. These factors can influence the participation to the experiments, the scalability of the measurement, and the reliability of the data collected. Considering acceptability by the users, some elements may act as barriers to the adoption and prolonged use of the device: the motivational profile (degree of autonomy and motivation) (Friel, 2020), design aspects (Auerswald, 2020), technical issues (Coughlin, 2020) and privacy concerns. Regarding the latter, collecting and processing personal data may impact individuals' sense of self, the perception of personal status, perceptions of contexts, and behaviours. Envisioning a successful experiment, the solutions and the processes for the collection of quantitative and qualitative data should be designed according to ethical guidelines and principles. To this purpose, participated design methodologies and codesign strategies can represent an effective way to increase the acceptability and effectiveness of the whole process.

## REFERENCES

- Abreu, K. L. de, Stoll, I., Ramos, L. S., Baumgardt, R. A., Kristensen, C. H. (2002). Estresse ocupacional e Síndrome de Burnout no exercício profissional da psicologia. *Psicol. cienc. prof.* 22, 22–29.
- Al-Marouf, R., Alhamad, A. Q., Aburayya, A., Salloum, S. (2021). User Acceptance of Smart Watch for Medical Purposes: An Empirical Study. *Future Internet* 13, no. 5: 127.
- Auerswald, T., Meyer, J., von Holdt, K., Voelcker-Rehage, C. (2020). Application of Activity Trackers among Nursing Home Residents-A Pilot and Feasibility Study on Physical Activity Behavior, Usage Behavior, Acceptance, Usability and Motivational Impact. *Int J Environ Res Public Health* 17, 6683.
- Bajaj, Ruhi Kiran, Rebecca Mary Meiring, and Fernando Beltran, (2023). Co-Design, Development, and Evaluation of a Health Monitoring Tool Using Smartwatch Data: A Proof-of-Concept Study. *Future Internet* 15, no. 3: 111.
- Bertoloni, M. (2016). Sei fattori per il benessere organizzativo. *Rivista degli infortuni e delle malattie professionali.* 14.
- Bluyssen, P. M., Aries, M. and Dommelen, P. Van (2011). Comfort of workers in office buildings: The European HOPE project, *Building and Environment.* Elsevier Ltd, 46(1), pp. 280–288.

- Bordi, L., Okkonen, J., Mäkineniemi, J.-P., Heikkilä-Tammi, K., (2018). Communication in the Digital Work Environment: Implications for Wellbeing at Work. *Nordic Journal of Working Life Studies* 8.
- Chianella, R., Mandolfo, M., Lolatto, R., Pillan, M., (2021). Designing for Self-awareness: Evidence-Based Explorations of Multimodal Stress-Tracking Wearables, in: Kurosu, M. (Ed.), *Human-Computer Interaction. Interaction Techniques and Novel Applications*. Springer International Publishing, Cham, pp. 357–371.
- Chalmers, T., Hickey, B. A., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C. S., Clifton-Bligh, R., Morley, J., Lal, S. (2022). Stress Watch: The Use of Heart Rate and Heart Rate Variability to Detect Stress. *Sensors* 22, 151.
- Coughlin, S., Caplan, L., Stone, R. (2020). Use of consumer wearable devices to promote physical activity among breast, prostate, and colorectal cancer survivors: A review of health intervention studies - PubMed [WWW Document].
- Davies, A. C. L. (2022). Stress at Work: Individuals or Structures?. *Industrial Law Journal*, Volume 51, Issue 2, June 2022, Pages 403–434.
- Davis, F. (1989). Technology Acceptance Model. PDF Download Link Free [WWW Document]. Education Consultant. <https://globalassistant.info/technology-acceptance-model-davis-1989-pdf-download-link-free/> (accessed 4.5.24).
- Demerouti, E., Bakker, A. B., Nachreiner, F., Schaufeli, W. B. (2001). The job demands-resources model of burnout. *J Appl Psychol* 86, 499–512.
- Di Flumeri, G., Aricò, P., Borghini, G., Sciaraffa, N., Di Florio, A., Babiloni, F. (2019). The Dry Revolution: Evaluation of Three Different EEG Dry Electrode Types in Terms of Signal Spectral Features, Mental States Classification and Usability. *Sensors (Basel)* 19, 1365.
- EUA 2024 - European Agency for Safety and Health at Work. Mental health at work after the COVID-19 pandemic – What European figures reveal. ISBN 978-92-9402-238-7 doi: 10.2802/151862.
- Friel, C. P., & Garber, C. E. (2020). An Examination of the Relationship Between Motivation, Physical Activity, and Wearable Activity Monitor Use. *Journal of Sport and Exercise Psychology*, 42(2), 153–160.
- Gedam, S., Paul, S. (2021). A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques. *IEEE Access* 9, 84045–84066.
- Giorgi, A., Ronca, V., Vozzi, A., Sciaraffa, N., Di Florio, A., Tamborra, L., Simonetti, I., Aricò, P., Di Flumeri, G., Rossi, D., Borghini, G. (2021). Wearable Technologies for Mental Workload, Stress, and Emotional State Assessment during Working-Like Tasks: A Comparison with Laboratory Technologies. *Sensors* 21, 2332.
- Haghi, M., Thurow, K., Stoll, R. (2017). Wearable Devices in Medical Internet of Things: Scient. Res. and Commercially Available Devices. *Health Inform Res* 23, 4–15.
- Haleem, A., Javaid, M., Singh, R. P., Suman, R., Khan, S. (2023). Management 4.0: Concept, applications and advances. *Sustainable Operations and Computers* 4, 10–21.
- Hernando, D., Roca, S., Sancho, J., Alesanco, Á., Bailón, R., (2018). Validation of the Apple Watch for Heart Rate Variability Measurements during Relax and Mental Stress in Healthy Subjects. *Sensors (Basel)* 18, 2619.
- Hickey, B. A., Chalmers, T., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C. S., Clifton-Bligh, R., Morley, J., Lal, S. (2021). Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review. *Sensors (Basel)* 21, 3461.
- Hong, S., Yang, Y., Lee, J., Yang, H., Park, K., Lee, S., Lee, I., Jang, Y. (2010). Ambulatory stress monitoring with a wearable bluetooth electrocardiographic device. *Stud Health Technol Inform* 161, 66–76.

- Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, Kahn J, Simpson SA, Alvarez-Jimenez M, Rice F, Evans R, Merry S. (2020). Practitioner review: Co-design of digital mental health technologies with children and young people. *J Child Psychol Psychiatry*. 2020 Aug; 61(8): 928–940.
- Juchnowicz, M., Kinowska, H., (2021). Employee Well-Being and Digital Work during the COVID-19 Pandemic. *Information* 12, 293.
- Kang, H. S., Exworthy, M. (2022). Wearing the Future-Wearables to Empower Users to Take Greater Responsibility for Their Health and Care: Scoping Review. *JMIR Mhealth Uhealth* 10, e35684.
- Lazarus, R. S. (1966). *Psychological Stress and the Coping Process*. McGraw-Hill.
- Leka, S., Jain, A., (2010). *Health Impact of Psychosocial Hazards at Work: An Overview*.
- Lilja, J. (2020). *Digitalisation and well-being at work. Understanding work transformation and the role of acceptance through thematic narrative analysis [thesis]*. Tampere: Tampere University.
- Lu, S., Wei, F., Li, G., 2021. The evolution of the concept of stress and the framework of the stress system. *Cell Stress* 5, 76–85.
- Lukan, J., Bolliger, L., Clays, E., Mayora, O., Osmani, V., Luštrek, M. (2021). Participants' Experience and Adherence in Repeated Measurement Studies Among Office-Based Workers. *Adjunct Proc. of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proc. of the 2021 ACM International Symposium on Wearable Computers, UbiComp/ISWC '21 Adjunct*. Association for Computing Machinery, New York, NY, USA, pp. 528–531.
- Menghini, L., Gianfranchi, E., Cellini, N., Patron, E., Tagliabue, M., Sarlo, M. (2019). Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions. *Psychophysiology* 56, e13441.
- Mercer, K., Li, M., Giangregorio, L., Burns, C., Grindrod, K., 2016. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. *JMIR mHealth and uHealth* 4, e4461.
- Michie S. (2002). Causes and management of stress at work. *Occupational and Environmental Medicine*. 2002; 59:67–72.
- Misra, N., Shobhna S. (2023). *Happiness at Work: A Psychological Perspective. Happiness and Wellness - Biopsychosocial and Anthropological Perspectives*. IntechOpen. doi: 10.5772/intechopen.108241.
- Nelson, M. J., Thomson, R. L., Rogers, D. K., Howe, P. R. C., Buckley, J. D. (2014). Maximal rate of increase in heart rate during the rest-exercise transition tracks reductions in exercise performance when training load is increased. *J Sci Med Sport* 17, 129–133.
- Oinas-Kukkonen, H. (2013). A foundation for the study of behavior change support systems. *Pers Ubiquit Comput* 17, 1223–1235.
- Pakhomov, S. V. S., Thuras, P. D., Finzel, R., Eppel, J., Kotlyar, M. (2020). Using consumer-wearable technology for remote assessment of physiological response to stress in the naturalistic environment. *PLOS ONE* 15, e0229942.
- Peake, J. M., Kerr, G., Sullivan, J. P. (2018). A Critical Review of Consumer Wearables, Mobile Applications, and Equipment for Providing Biofeedback, Monitoring Stress, and Sleep in Physically Active Populations. *Front Physiol* 9, 743.
- Pega F, Momen, N. C., Abubakar, A., Al-Emam R., Hassan, M. N., Howard, J. (2023). *Monitoring worker's health: Focus on rights, determinants, and equity*. Vol. 402, Issue 10410, Oct. 2023.

- Perego, P., Scagnoli, M., Sironi, R. (2022). Co-design the Acceptability of Wearables in the Healthcare Field. In: Spinsante, S., Silva, B., Goleva, R. (eds) IoT Technologies for Health Care. HealthyIoT 2021. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecom. Eng., vol. 432. Springer, Cham.
- Pillan, M., Varisco, L., Bertolo M. (2017). Facing Digital Dystopias: A Discussion about Responsibility in the Design of Smart Products. Proc. of the Conference on Design and Semantics of Form and Movement - Sense and Sensitivity, DeSForM 2017. InTech.
- Probst, T. M. (2010). Multi-level models of stress and well-being. *Stress and Health* 26, 95–97.
- Rodrigues, S., Paiva, J. S., Dias, D., Aleixo, M., Filipe, R. M., Cunha, J. P. S. (2018). Cognitive Impact and Psychophysiological Effects of Stress Using a Biomonitoring Platform. *Int J Environ Res Public Health* 15, 1080.
- Salanova, M., Llorens Gumbau, S., Ventura, M. (2014). Technostress: The Dark Side of Technologies, in: *The Impact of ICT on Quality of Working Life*. pp. 87–103.
- Shiffman, S., Stone, A. A., Hufford, M. R. (2008). Ecological momentary assessment. *Annu Rev Clin Psychol* 4, 1–32.
- Shandhi, M. M. H., Singh, K., Janson, N. (2024). Assessment of ownership of smart devices and the acceptability of digital health data sharing. *npj Digit. Med.* 7, 44.
- Siegrist, J., 1996. Adverse health effects of high-effort/low-reward conditions. *J Occup Health Psychol* 1, 27–41.
- Sleigh, J., Amann, J. (2022). Fostering ethical reflection on health data research through co-design: A pilot study. *International Journal of Ethics Education* 7, 325–342.
- Sonnentag S., Tay L. (2023). A review on health and well-being at work: Morethan stressors and strains. *Personnel Psychology*. 76:473–510.
- Stone, A. A., Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine* 16, 199–202.
- Stojchevska, M., Steenwinckel, B., Donck, J., De Brouwer, M., Goris, A., De Turck, F., Hoecke, S., Ongenaes, F. (2022). Assessing the added value of context during stress detection from wearable data. *BMC Medical Informatics and Decision Making* 22.
- Susanto, P., Hoque, M. E., Jannat, T., Emely, B., Zona, M. A., Islam, M. A. (2022). Work-Life Balance, Job Satisfaction, and Job Performance of SMEs Employees: The Moderating Role of Family-Supportive Supervisor Behaviors. *Front Psychol* 13, 906876.
- Tams, S., Ahuja, M., Thatcher, J., Grover, V. (2020). Worker stress in the age of mobile technology: The combined effects of perceived interruption overload and worker control. *The Journal of Strategic Information Systems*. Vol. 29, Issue 1, 101595.
- The Lancet, editorial. (2023). The future of work and Health. Vol. 402, Issue 10410, Oct. 2023.
- Thilagavathy, S., Geetha, S. N. (2021). Work-life balance -a systematic review. *Vilakshan - XIMB Journal of Management* 20, 258–276.
- van Kraaij A. W. J., Schiavone G., Lutin E., Claes S., Van Hoof C. (2020) Relationship Between Chronic Stress and Heart Rate Over Time Modulated by Gender in a Cohort of Office Workers: Cross-Sectional Study Using Wearable Technologies *J Med Internet Res* 2020;22(9): e18253.
- van Hierden, Y.; Dietrich, T.; Rundle-Thiele, S. (2021). Designing an eHealthWell-Being Program: A Participatory Design Approach. *Int. J. Environ. Res. Public Health* 2021, 18, 7250.



- Varisco A – Varisco, L., Colombo, S., Casalegno, F. (2019). Designing with data. Anticipating the impact of personal data usage on individuals and society. Intelligent Human Systems Integration 2019 – Proc. of the 2nd Int. Conf. on Intelligent Human Syst. Integration IHSI 2019, Advances in Intelligent Systems and Computing 870–876.
- Varisco B – Varisco, L., Pavlovic, M., Pillan M. (2019). Anticipating Ethical Issues When Designing Services That Employ Personal Data. In *Design, User Experience, and Usability. Design Philosophy and Theory*, ed. by A. Marcus and W. Wang, 11583:113–31. Lecture Notes in Computer Science. Cham: Springer International Publishing.
- Weale, V., Love, J., Clays, E., Oakman, J. (2023). Using EMA and Physiological Data to Explore the Relationship between Day-to-Day Occupational Stress, Musculoskeletal Pain and Mental Health among University Staff: A Study Protocol. *Int. J. Environ. Res. Public Health* 2023, 20, 3526.
- World Health Organization. (2022). WHO guidelines on mental health at work. World Health Organization. <https://iris.who.int/handle/10665/363177>. License: CC BY-NC-SA 3.0 IGO
- WHO 1948 - World Health Organization, 1948. Constitution of the World Health Organization [WWW Document]. <https://www.who.int/about/accountability/governance/constitution> (accessed 4.5.24).
- Wijsman, J., Grundlehner, B., Liu, H., Hermens, H., Penders, J. (2011). Towards mental stress detection using wearable physiological sensors. *Annu Int Conf IEEE Eng Med Biol Soc* 2011, 1798–1801.
- Zani, B., Cicognani, E. (1999). *Le vie del benessere* [WWW Document]. Carocci editore. URL <https://demo.carocci.it/prodotto/le-vie-del-benessere> (accessed 4.5.24).

# Evaluating Ergonomic Design: A User Command Interface for Industrial Exoskeletons

**Olmo Alonso Moreno Franco<sup>1</sup>, Chiara Lambranzi<sup>1</sup>,  
Roberto Pitzalis<sup>1</sup>, Matteo Sposito<sup>1</sup>, Mahnaz Asgharpour<sup>2</sup>,  
Daegeun Park<sup>1</sup>, Christian Di Natali<sup>1</sup>, Luigi Monica<sup>3</sup>,  
Darwin G. Caldwell<sup>1</sup>, and Jesus Ortiz<sup>1</sup>**

<sup>1</sup>Advanced Robotics Department, Istituto Italiano di Tecnologia, Genova,  
Via Morego 30, 16163, Genova, Italy

<sup>2</sup>Piazzetta Giulio Marchi 7, 16163, Genova, Italy

<sup>3</sup>Department of Technological Innovation and Safety Equipment, Italian Workers'  
Compensation Authority, Monte Porzio Catone, 00078, Rome, Italy

## ABSTRACT

Industrial workers perform daily activities with a high risk of musculoskeletal disorders. Diverse studies have reported high rates of musculoskeletal disorders among distinct industry professionals, with values exceeding 75% for most occupations considered. Commonly affected body areas include the back (particularly the lower back), shoulders, and lower limbs. A potential solution to reduce the risk of injury among industrial workers is the use of exoskeletons in the workplace. This wearable suit improves ergonomics depending on the body part it supports. From the actuation point of view, exoskeletons can be categorised in three branches: passive, active, and quasi-passive (semi-active). Active exoskeletons contain sensors, actuators, and electric controller boards; these characteristics make them more versatile for adapting the control strategy to the required task. The wearer of an active exoskeleton, needs of a human-machine interface to modify parameters that impact the exoskeleton control strategy. The user command interface is a wearable device that allows easy adjustments when an interaction occurs. Experiments were conducted with 20 participants to evaluate the physical ergonomic attributes of four different versions of the human-machine interface and a mobile phone as a standard device. Results showed that there are significant differences in the comfort and size attribute between UCI interfaces V2.0 and V2.2. Despite the mobile phone obtaining the highest scores in terms of aesthetics, comfort, durability, and safety; the UCI V2.2 interface presented a tendency of improvement in these attributes.

**Keywords:** Human machine interface, Industrial exoskeletons, User command interface

## INTRODUCTION

Musculoskeletal disorders (MSDs) are physical diseases provoked in many cases by overexertion of muscles at specific joints, being the back the most commonly affected region. MSDs can result from factors such as incongruous postures, handling heavy loads, and repetitive lifting (Poliero

et al., 2021). Industrial exoskeletons offer several benefits in the workplace, including the reduction of biomechanical strains, prevention of MSDs, enhanced worker safety, increased productivity, and improved ergonomics. Exoskeletons can help reduce physical exertion by providing external support and assistance during physically demanding tasks. By reducing muscle demand and biomechanical stresses on the body, exoskeletons prevent work-related MSDs such as low back pain and shoulder tendinopathy (Theurel et al., 2019).

In addition, exoskeletons improve worker safety by minimising the risk of injuries associated with repetitive or strenuous tasks, such as manual material handling (MMH). By decreasing fatigue and physical effort on workers, exoskeletons play an important role in improving productivity and efficiency in the workplace. The use of these wearable devices promotes better posture and movement patterns, leading to improved ergonomics and mitigating the risk of musculoskeletal injuries (Theurel et al., 2019).

An active exoskeleton is an electromechanical structure worn by an operator that mimics the shape and functions of the human body. It is designed to augment the abilities of the human limb or trunk or to assist in the prevention of MSDs (Khairul et al., 2012). The main applications of active exoskeleton robots discussed include their use as assistive devices, rehabilitation devices, human amplifiers, and haptic interfaces (Gopura et al., 2019). From the actuation point of view, active exoskeletons use electrical or pneumatic actuators together with sensors and control boards (Lazzaroni et al., 2019).

These components make active exoskeletons more versatile to adapt to the user and the tasks to perform; allowing to achieve proper force modulation according to the control strategy adopted (Poliero et al., 2021). A human-machine interface (HMI) is required to provide communication between a human operator and the wearable device to modulate and adjust the parameters of the active exoskeleton. The HMI is the basis of cognition, communication, and interaction (Gong et al., 2009). HMI ergonomics takes a holistic, human-centred approach when designing and evaluating systems. The traditional domain of specialisation within ergonomics comprises physical ergonomics. This domain is primarily concerned with human anatomical, anthropometric, physiological, and biomechanical characteristics as they relate to physical activity. By improving HMI and human-computer interactions, ergonomics enhances system performance and user experience (Karwowski et al., 2005).

In this paper, we present an ergonomic assessment of the User Command Interface (UCI), a wearable HMI device used to configure the industrial exoskeleton XoTrunk. The interface plays a crucial role in addressing the challenges faced by developers in optimising industrial exoskeleton capabilities by offering adaptability, control, usability and performance enhancement features. This electromechanical device attached to the exoskeleton provides a solution for achieving user interaction. However, human factors regarding physical ergonomics have not been addressed with exoskeleton' users when the interface is in use. The UCI interface has been integrated into the back-support XoTrunk (Moreno et al., 2023) and upper-limb Shoulder-SideWINDER (Moreno et al., 2024) exoskeletons.

Experiments were conducted with 20 subjects. Ergonomics assessment was performed by comparing the physical attributes of four different versions of the UCI interface and a smartphone, considering the last one a standardised device.

## METHODOLOGY

Highlighting the difficulties faced is important when analysing design requirements in wearable devices, particularly in terms of measuring attributes such as comfort. For instance, this term may be defined differently in various studies, sometimes as a standalone design requirement and in other cases as part of a group of requirements. However, comfort was found to encompass aspects such as freedom from discomfort and pain, acceptable temperature, texture, shape, weight, and tightness, all of which contribute to the overall comfort and usability of a device (Francés-Morcillo et al., 2020). To assess the interface, we performed a comparison test of three physical ergonomic attributes: comfort, durability, and safety. Using the mapping wearable design requirements method, five shape-like interfaces were evaluated. Four out of five interfaces are previous versions according to the evolution of our interface, and the last one is a mobile phone. This approach of quantifying and analysing design requirements helps in understanding the complex relationships between different terms and ensures a more systematic and thorough evaluation of design aspects in wearable devices.

### System Description

The UCI is designed to enhance the functionality and versatility of industrial wearable robots, specifically the XoTrunk and Shoulder-sideWINDER exoskeletons. The UCI serves as a control interface that allows users to interact with and customise the settings of the exoskeleton system. It provides a user-friendly platform for users to access and modify various parameters related to the exoskeleton's operation, such as secure identification, signal monitoring, user management, control strategy adjustments, session management, task-specific configurations, and user profiles. The UCI includes features such as menus, submenus, cards, and decks to facilitate user interaction and task execution. In addition, the UCI is designed with the principles of security and interaction in mind, featuring a navigation wheel and buttons for user input and control (Moreno et al., 2022). Figure 1 depicts the interface, which has a display with a resolution of 800×480 pixels and a colour screen.

Navigation through the menu occurs when the user spins the lateral wheel and presses to select the desired option. The device fits in one hand and can be used by both left- and right-handed people. When the user is not interacting with the device, the interface is attached to the front of the XoTrunk exoskeleton. Currently, UCI V2.0 and V2.1 are in operation along with the XoTrunk and Shoulder-SideWINDER exoskeletons. Although these are similar in dimensions, UCI V2.1 differs because it has a holder component at the top of the case to allow the interface to be attached to the exoskeleton

with less weight. UCI V1.0 was the first prototype designed to test the low-level control layer for electronic components in a housing case. UCI V2.2 is a proof of concept that does not use the reduced size from the previous version.



**Figure 1:** User command interface system. The UCI is a wearable gadget that fits in the palm of the hand and is attached to the exoskeleton.

### Evaluation Metrics

The assessment metric used in this study comes from the co-evaluation checkpoint presented by Francés-Morcillo in the wearable design requirements, parameters, and definitions table (Francés-Morcillo et al., 2020). The evaluation for this study consisted of 20 items, including aesthetics, comfort, durability, and safety. It uses a Likert-scale option from 1 (“Totally disagree”) to 5 (“Totally agree”) and includes an “not applicable” (N/A) option. The metric can be found in the Sec. Appendix.

## EXPERIMENTAL EVALUATION

The experiment assessed the ergonomics and attributes of the User Command Interface related to comfort, durability, and safety. The newest version of the UCI is compared with three older versions and a commercial cell phone used as a reference device.

### Participants

A group of 20 subjects participated in the experiments; among the participants, 8 were females and 12 were males. Two participants were left handed. The experiment was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Liguria (protocol no.: CER Liguria 001/2019).

### Experiment Design

The experiment evaluated the differences in the physical and ergonomic aspects of the diverse versions of the UCI by holding and using the physical interfaces. Figure 2 shows the five devices: a) UCI V1.0, b) UCI V2.0, c) UCI V2.1, d) UCI V2.2, and e) smartphone Xiaomi Redmi Note 9.



**Figure 2:** Experimental devices for comparison. a) UCI V1.0, b) UCI V2.0, c) UCI V2.1, d) UCI V2.2, and e) smartphone.

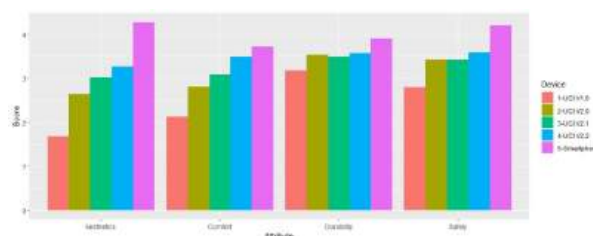
Because the UCI is a wearable device that is similar in shape to a mobile phone, a smartphone was used as a pivot to compare the attributes of the UCI and a standardised device. Although the shape is similar, the functionality of the UCI and the smartphone is far from being comparable. The UCI was designed to navigate using a rotating wheel coupled with push buttons and not to operate with a touch screen. This feature is due to the potential existence of dust, grease, water, or gloves on the user's hands. In contrast, navigation with the smartphone is performed using the touch screen. The experiment was conducted in a sequence of three tasks. First, each participant for all five devices held the device in the left hand, and only for the UCI versions rotated the wheel and pressed the push button for at least 5 s. For the smartphone, the user was required to unlock the screen. Second, the same task as the previous one was repeated, but this time the participants were using the right hand. Finally, each participant held the device in both hands, rotated the wheel, and pressed the push button for at least 5 s using the UCI version. For the smartphone, the participant had to unlock the screen. Table 1 shows the physical characteristics, such as the dimensions and weight of each device.

**Table 1.** Experimental device physical characteristics.

Device	Length (mm)	Width (mm)	Height (mm)	Weight (gm)
UCI V1.0	128.00	84.65	38.42	298
UCI V2.0	138.63	84.90	56.05	381
UCI V2.1	138.00 + 25.00 (holder)	85.27	56.05	333
UCI V2.2	121.76	74.98	53.79	149
Smartphone	162.3	72.2	8.9	199

## RESULTS AND DISCUSSION

The 20-item results of the wearable design requirements co-evaluation checkpoint were grouped into four categories, as shown in Fig. 3. Results show that the smartphone obtained the highest grade in ergonomics attributes, including aesthetics, comfort, durability, and safety, followed by the UCI V2.2. In contrast, UCI V1.0 had the lowest score among all previously mentioned attributes.



**Figure 3:** Experimental results of the wearable design requirements co-evaluation checkpoint. The ergonomics attributes such as aesthetics, comfort, durability and safety were evaluated using five different interfaces: 1) UCI V1.0, 2) UCI V2.0, 3) UCI V2.1, 4) UCI V2.2, and 5) smartphone.

Because UCI V2.0 and UCI V2.1 are the current operative interfaces, the characteristics presented in Table 1 show that the difference between these interfaces is related to the weight and extra length of UCI V2.1 from the top holder component. This extra length does not affect grip on the user's hand. Mann–Whitney U test was conducted to determine whether there is a perception in difference in weight between the UCI V2.0 and the UCI V2.1. The results indicate a nonsignificant difference,  $W = 230$ ,  $p\text{-value} = 0.4011$ . Therefore, we must not reject the null hypothesis and conclude that there is no difference in weight between these interfaces.

The UCI V2.2 was designed with a reduction in size to improve the operability of the interface with smaller hands. We conducted a Mann–Whitney U test to find a significant statistical difference in the concept of size from the comfort attributes between UCI V2.0 and UCI V2.2. The results indicate a significant difference,  $W = 323$ ,  $p\text{-value} = 0.0005914$ ; therefore, we reject the null hypothesis. Therefore, we have sufficient evidence to say that there is a significant difference in the size of UCI V2.0 and UCI V2.2.

## CONCLUSION

Ergonomic aspects such as aesthetics, comfort, durability, and safety have an impact on the usability and functionality of human–machine interfaces. We emphasise in the need for user-friendly interfaces that prioritise comfort, durability, and safety. The results of the study indicate that UCI V2.2, designed with a reduction in size, showed a significant difference compared with UCI V2.0. This finding underscores the importance of considering user comfort and adaptability in the design of exoskeleton interfaces. The incorporation of a smartphone in the experiment demonstrated that the UCI devices in all their versions have a distinction of ergonomics aspects such as aesthetics and safety. In addition, this study used a systematic evaluation approach to quantify and analyse design requirements, providing valuable insights into the complex relationships between different ergonomic terms. By incorporating user feedback and conducting rigorous testing, future developments in exoskeleton technology can further enhance worker well-being and productivity in industrial environments.

## APPENDIX

**Table A1.** Wearable design requirements co-evaluation checkpoint: attributes, concepts and items (Francés-Morcillo et al., 2020).

Attribute	Concept	Item
Aesthetics	Customization	The device is customizable
	Fashion	The device is coherent to the aesthetical and fashion that have been defined
		The device is appealing to use

Continued

**Table A1.** Continued.

Attribute	Concept	Item
Comfort	Breathability	The device is breathable and it avoids the accumulation of sweat
		The device has some slack to circulate air without compromising fit hand
	Hygiene	The device can be washed
	Movement	The device is sufficiently flexible to allow the natural movement of the body region
		Obstrusiveness
	Shape	The device enables the natural body movements
		The adjustment to the body region is the proper one
		The device is properly attached to the user and there is no danger of losing it
Sizing	The device fits the shape of the body region	
Temperature	The device adapts to my hand size	
	The temperature does not increase above the recommended value	
Durability	Weight	The device is light
	Resistance	The device seems to be resistive for its life cycle
		The device is properly protected from external elements (eg. hand tools)
Safety	Harm	The device is safe it does not cause pain to my hand
		All the device components are properly attached
		Heat dissipating devices are separated from the user skin

## ACKNOWLEDGMENT

This work was supported by the Italian Workers' Compensation Authority (INAIL) and Istituto Italiano di Tecnologia (IIT) within the project Sistemi Ciberneteci Collaborativi - Esoscheletro Collaborativo 3.

## REFERENCES

- Anam, K., & Al-Jumaily, A. A. (2012). Active Exoskeleton Control Systems: State of the Art. *Procedia Engineering*, 41, 988–994.
- Gong, C. (2009). Human-Machine Interface: Design Principles of Visual Information in Human-Machine Interface Design. *2009 International Conference on Intelligent Human-Machine Systems and Cybernetics*, 2, 262–265.
- Lazzaroni, M., Tabasi, A., Toxiri, S., Caldwell, D. G., Momi, E. D., Dijk, W. van, Looze, M. P. de, Kingma, I., Dieën, J. H. van, & Ortiz, J. (2020). Evaluation of an acceleration-based assistive strategy to control a back-support exoskeleton for manual material handling. *Wearable Technologies*, 1, e9.
- Francés-Morcillo, L., Morer-Camo, P., Rodríguez-Ferradas, M. I., & Cazón-Martín, A. (2020). Wearable Design Requirements Identification and Evaluation. *Sensors*, 20(9), Article 9.
- Karwowski, W. (2005). Ergonomics and human factors: The paradigms for science, engineering, design, technology and management of human-compatible systems. *Ergonomics*, 48(5), 436–463.



- Moreno Franco, O. A., Ortiz, J., & Caldwell, D. G. (2022). Evaluation of the User Command Interface, an Adaptable Setup System for Industrial Exoskeletons. *2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 01–07.
- Moreno, F. O. A., Crespo, J., Di Natali, C., Ortiz, J., & Caldwell, D. G. (2023). Integration of the User Command Interface to the Industrial Exoskeleton XoTrunk. *2023 IEEE/SICE International Symposium on System Integration (SII)*, 1–7.
- Moreno, F. O. A., Park, D., Di Natali, C., Caldwell, D. G., & Ortiz, J. (2024). Integration and Task Assessment of the User Command Interface to the Occupational Exoskeleton Shoulder-sideWINDER. *2024 IEEE/SICE International Symposium on System Integration (SII)*, 1176–1182.
- Poliero, T., Iurato, M., Sposito, M., Natali, C. D., Toxiri, S., Anastasi, S., Draicchio, F., Monica, L., Caldwell, D. G., Sanguineti, V., & Ortiz, J. (2021). A case study on occupational back-support exoskeletons versatility in lifting and carrying. *The 14th PErvasive Technologies Related to Assistive Environments Conference*, 210–217.
- Theurel, J., & Desbrosses, K. (2019). Occupational Exoskeletons: Overview of Their Benefits and Limitations in Preventing Work-Related Musculoskeletal Disorders: IISE Transactions on Occupational Ergonomics and Human Factors: Vol. 7, No. 3–4. *IISE Transactions on Occupational Ergonomics and Human Factors*, 7(3–4), 264–280.

# Haptic (Tactual), Portable, Hands-Free Communication for Body Compliant Interfaces

Nils-Krister Persson<sup>1,2</sup>, Amelie Olesen<sup>1,2</sup>, Eva Lindell<sup>1,2</sup>, and Li Guo<sup>1,2</sup>

<sup>1</sup>Swedish School of Textiles, Polymeric E-textiles, University of Borås, Sweden SE-501 90 Borås, Sweden

<sup>2</sup>Smart Textiles, Smart Textiles Technology Lab, University of Borås, Sweden SE-501 90 Borås, Sweden

## ABSTRACT

There is a growing number of technical communication devices, not least wearables, which take use of the haptic sense(s). Then tactors (vibrotactile elements, heating elements, cooling elements, pressure generators, active indentators, electro-stimulating electrodes etc.), are employed. Haptic technologies are often limited to binary, point-wise actuation (one vibrator). However, as we discuss, in a semiotic sense, this can only generate a representamen that is symbolic, thus only also concerning symbolic communication. For a richer communication coming closer to what exists for visual and audial displays also haptic communication that is semiologically iconic and indexical are of interest. We here present a classification of tactile displays. For this, we make a distinction between the stimulus (the tactor characteristics and relationship to human reception) and the spatial arrangement i.e. the geometrical placement of tactors. For wearables, in the latter case, the human body shape and anatomy is taken into concern. From this, we build two (partially ordered) hierarchies, the stimulus richness hierarchy and the spatial ordering hierarchy, respectively. Combining these hierarchies gives a (partially ordered) hierarchy of tactile displays for the human body. We show that the informatical richness is fast growing with placement complexity. However, such displays need space. Hands and fingers are sensitive for haptic stimuli but are better reserved for active touch. Instead, other body parts might be used. For this, textiles are employed. We demonstrate a chairable i.e. a portable textile based haptic display for communication to (deafblind) humans, arena spectators etc., that can be applied to furniture, thus enriching the Umwelt of the users.

**Keywords:** Haptic displays, Haptics, Tactile, Haptic communication, Smart textiles, Deafblindness

## INTRODUCTION

Haptics offer a number of interesting possibilities for communication; an extra channel when the other senses are occupied and saturated; the potential of a hidden communication channel from the outside; opening up for affective communication, bridging somato-mental treatment and

communication etc. It is also the case that there are individuals that need to completely rely on haptic communication for their interaction with others. These are persons having deafblindness.

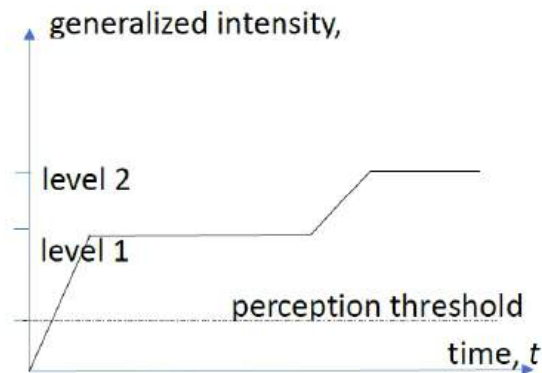
A haptic display is any device that is able to artificially generate stimulus perceived as haptual (touch, heat etc.) Haptic displays (HD) could be designed along different strategies and have a variety of features. Here we discuss ways of ordering such displays. For many applications of wearables, it is favorable with hands-off solutions, as hands should be reserved for other purposes such as carrying a bag or maneuvering a cane. Active touch is related to finger and hand movements and might better be reserved for that rather than becoming involved in and disturbed by haptic displays. Instead merely  $2 \text{ m}^2$  of body skin is at disposal for communication.

We start by making a distinction between the *stimulus* and its characterisations, quality, quantities and modes on one hand side and its number and *spatial* arrangement with geometric considerations on the other. We then have a model where the central distinction is about a *stimulus domain* and a *spatial domain*. The latter is a purely geometrical description of how a display could be arranged. From this, we build two (partially ordered) hierarchies, the *stimulus richness hierarchy* of the stimulus domain and the *spatial ordering hierarchy* of the *spatial domain*, respectively.

Combining these hierarchies gives a (partially ordered) *hierarchy of haptic displays* for the human body.

## STIMULUS DOMAIN

The stimulus is in a haptic and tactile display generated by a tactor. A *tactor* (from Latin *tactile* feeling) is any physical device that by some (or several) mechanical-electrical (occasionally biochemical) mechanisms affect any human receptor so that human perception is possible. A tactor is giving some physical quantity, force, pressure, vibration frequency, voltage, electrical current etc. We generalize this to an intensity,  $Q$ , see Figure 1. The quantity must be such that it stimulates human receptors in the skin. They are of many different types, here not further discussed; Connect with the next phrase “Pacinian ...”



**Figure 1:** Stimulus requirement for perception.

Pacinian corpuscles, Meissner corpuscles, Skip the phrase, Merkel complexes, Ruffini corpuscles, C-fiber LTM (low threshold mechanoreceptors), etc.

For creating the most simple of signals one need to have contrast, see Figure 1. Level 1 is meant to create a sensed signal. It then needs to be above any perception threshold. For creating nil can either use “non presence” of signal i.e. something below the perception threshold for this quantity or having two levels, level 1 and level 2 both above the perception threshold for this quantity. The latter case is typically more difficult to discriminate and perceive.

A tactor is defined in general manner (“is possible”) but of course, only tactors that generates stimulus that is sensed by human receptors and perceived by the Central Nervous System are of practical interest. Several categories of tactors have been developed; foremost of these are vibrotactile elements (Table 1). The tactor is such that it gives many types of stimuli (pressure, vibration, skewing etc.) at each point.

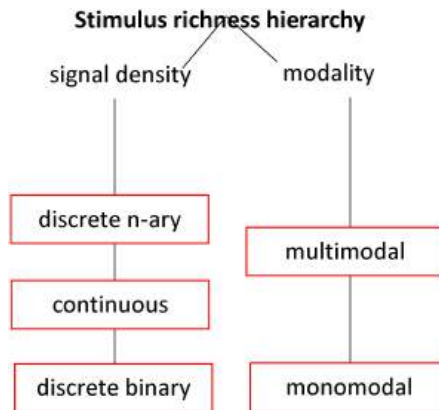
**Table 1.** Overview of tactors suitable for integration into wearables and smart textiles.

Type	Mechanism
<b>Vibrotactile elements</b>	
Coin vibrators	Nonconcentric
Cylindrical vibrators	Nonconcentric
<b>Pressure and skewing</b>	
Pneumatic compressors	Pumping out to under-pressure
McKibben adstringent	McKibben effect: pressure is causing a locked, tubular bladder expand radially and therefore contract axially
Adstringent band	Electrical DC motor
<b>Pointwise pressure</b>	Electromagnetic break
<b>Actuators</b>	A wealth of mechanisms incl. electroactive polymers
<b>Electrostimulation</b>	
Stimulating electrodes	Polarisation of the skin

We could divide tactors also based on their output. Modality is here regarded as a type of communication channel differentiable from other channels so that it has the potential of being sensed and being based on a certain physical quantity.

There could be just one output, then  $Q$  is *monomodal*. For example, vibration with one single frequency, pressed with a constant pressure to the skin och one single position. But it could also be that there are several types of output,  $Q$  being a vector  $(Q_1, Q_2, \dots, Q_n)$  in which the case it *multimodal*. Thus, there could be alternating current (AC) with current intensity,  $I$ , frequency,  $f$  etc.  $Q$  could also be *discrete* i.e. step wise or *continuous*. In the former, the case binary case (0 and 1) is the most simple. Then one could

generalize this to three intensities and further on to *n-ary signals*. It is not self-evident but it might be that many factors are not perfect and that in fact it is signals with intensities in a band rather than at discrete levels that are sent in. In devices presented it is also the case that it is common to employ only one type of discrete signal, not several types.



**Figure 2:** Stimulus richness hierarchy.

Thus we place discrete n-ary above continuous in terms of complexity. We then get the *stimulus richness hierarchy* of the stimulus domain. It consists of two legs, see Figure 2 *signal density* i.e. if it is discrete or continuous and *modality* if there are many types of Q:s. The relationship between the signal density and modality legs is not specified here. To the stimulus domain, we could also count how the stimulus could be conceived. Used for communication haptics are conveying *signs*. A sign is whatever entity bearing meaning and that is representing anything else than itself. With Peirce there are three classes of signs namely *symbols*, *icons*, and *indices*. A *symbol* is something to which there is a social-societal-cultural agreement, a convention. Examples include a letter of an alphabet symbolizing a certain linguistic sound. An *icon* expresses similarity to another entity. Examples include a photo that resembles the person in question. An *index* is something that is indicating that *w* is around, *w*, being some ontological entity; event, thing etc. An index is something that is there due to causality or otherwise bear some information on something else. Examples include touch indicating presence of another person.

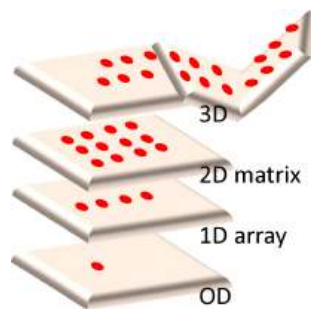
## SPATIAL DOMAIN

The *placement* i.e. spatial arrangement of factors in and on the device is here of concern. They should be related to the geometrical position on the body. This we denote *location* thus being an anatomical measure. The spatial domain consists of geometrical aspects only, no perceptual. Still of course, how factors are arranged will be important when one adds human spatial resolution capabilities and the uneven distribution over the body of this.

The placement with a set of factors could be regarded as ordered or not. In the former case – we denote that a *heap* – they are potentially able to convey more information as the number of combinations is growing fast. In the latter case individual factors are not meaningful and the factors are effectively regarded as one.

We denote one single factor as *zero dimensional*, 0D and we notice that this is a much-used situation in wearables or cell phones.

Arranging actuators in a linear order, *one dimensional*, 1D, opens up for temporal sequencing of the actuators. We differentiate between array consisting of  $n$  distinct elements and matrix. *Two-dimensional actuation*, 2D, is the arrangement of actuators on a limited part of the body so that the set of actuators still could be said to be in a plane or having maximum one single curvature. Actuators in a 2D arrangement could be in any non-array formation but a matrix arrangement is perhaps first at hand, by this forming columns and rows.



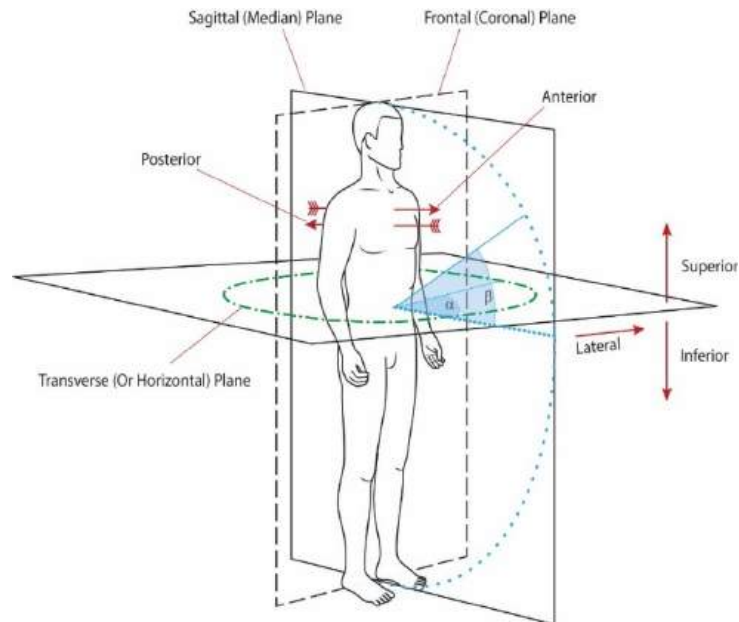
**Figure 3:** Spatial ordering hierarchy.

If the arrangement is having a more factors complicated geometry, switching between concave and convex shape, (*pseudo*) *three dimensional*, 3D, could be the term used. This also includes the case when covering different body parts. A switch between convexity and concavity might provide a surface where it is difficult to maintain mechanical contact with the skin for wearables.

It can be noted that full-fledged three-dimensional actuator arrangements are probably of less interest, as humans cannot, by the tactile sense, detect things outside of each other and actuators risk to cover each other. Still, for future research it is interesting to open up for this.

Awaiting introducing human capability aspects from a purely geometrical point of view the potential for communication is enormous. Assuming a  $m \times n$  matrix system with freedom for the distances in row direction, all being  $d_1$  and distances in the vertical direction,  $d_2$ . The  $m \times n$  matrix has every factor (commonly vibro-tactile elements) in either of say two states; active ('high voltage sent') and non-active ('low voltage applied or nothing sent'). Each factor could be active for an individually controllable period. There are  $m \cdot n$  factors. There are 2 times 2 times 2 times... = for  $m, n = 10$ , say (which could be feasible on the back of a person), this is a very high

number of combinations, i.e. signs, possible to generate, around  $1.26 \cdot 10^{30}$ . This is of course far outside of any human discrimination capability.



**Figure 4:** Human and planes, sections and directions used within medicine. Factors could be placed in many ways. Arranging them in the lateral, sagittal and coronal plane enables some directionality. It could be advantageous to avoid hands (red).

Many senses are directional. Vision gives information of where the dorsal object is, i.e. information on from which direction the flow of photons is coming from. Thanks to doubled ears audial stimulus in general show the same characteristic. For haptics, this is even more pronounced. Due to the fact that even in passive mode we know not only *that* someone touched us but also *where* on the body haptic stimulus occurred, the haptic sense is directional. *Directional faithfulness* is the characteristic of a sense that a proximal stimulus convey information about the direction (i.e. some angle and some distance) of the distal stimulus. In fact, the human body could sense haptic stimulus from beneath to the foot, normal to the legs and torso and from above on the head or shoulders or stretched-out arms. This is 180 degree in a vertical plane and 360 degree in a horizontal plane, see Fig 4. This could be called 580 degree sensing. Thus, body worn haptic displays should beneficiary have the input capability of directional faithfulness. By 2D arrangement, directional information is possible. By activating some of the actuators in a sequence this could mimic from where a certain phenomenon is coming (left-right, up-down).

## HAPTIC DISPLAY HIERARCHY

Of course merging two partial ordered structure is not possible to perform unambiguous. Here we suggest a hierarchy primary based on spatial dimensionality, and secondary on stimulation modality.

The simplest haptic display is one where there is a tactor at one single anatomical position that is able to one and only one stimulating quantity,  $Q$ . This is what is found in an alarm clock that besides giving sound – not interesting in our context - is shaking or vibrating. It could also be the vibrating function such as in a smart watch or vibrator found in cell phones.

This stimulus has two levels. “low” or “passive” or “zero” or “0”, silent” on one hand side and “alarming”, “active”, “on”, “1” on the other. Thus, it is binary 0-1. No other modulation is possible in this case. In order to be interesting from an informatical point of view the low level must be beyond any receptor detection level or be regarded as nothing more than a noise level. The Active level must be above any human threshold limit. The semantic content in itself is very limited, just “on” typically used as an “alarm”. There need to be a convention introduced to contextualize what it is to be meant when the tactor is active, if it is to be a fire alarm, burglary alarm, iron overheating alarm or hindrance detection signaling for people having deafblindness. Or is it a notification for an incoming mail in a smart watch? Thus, conventionalizing, this type of haptic display is a *symbol*. However, it could be argued that there is also an indexical component in it. If there is an alarm clock on the bedside table shaking to such an extent it is causing the person in the bed to be impacted and awakening it is an index for wakening up.

The tactor could then be exchanged to give not a binary signal but a continuous one. We then define this as the next level in the hierarchy for haptic display, Table 2. Of course, another alternative could have been chosen for the next step such as having two binary tactors. However, we keep to the strategy of working through a spatial dimensionality before going to the next level.

Next level is to let the tactor be such that many modalities are generated. This is typically technical complicated and seldom used. Instead, several tactors with different modalities are used in array or matrix form. We then introduce an array of tactors i.e. having what we denote a one-dimensional arrangement (class IV). The most simple of such is to have a binary stimulus 0-1. One could argue that one could also have an unstructured heap of a number of tactors is the most immature version of the 1D case. However, as discussed in the chapter earlier this is of less interest. So keeping the 1D spatial arrangement but switching the tactor to a continuous stimulus (class V) and then to a multimodal stimulus (class VI) comes next. In the last case, the stimulus could be either binary or continuous.

We now place the tactors spatially in a matrix, the 2D situation. We follow the same strategy as in 1D case. First, we create a class with binary tactors (class VII) then with continuous tactors (class VIII) and then allowing for multimodal stimulus (class IX).



If we now place the factors (1D or 2D, but unified and for generality denoted 2D in Table 2.) curvilinear and distributed for individual perception it is possible to obtain directionality. This can be done in three planes as in Fig. 4. For navigation, tactile navigation belts have been developed. They are a system with factors worn around the waist or on the shoulders, sometimes on the arms and hands. For geographical guiding a left-right indication is the most interesting the variant of “Partly directional haptic display” with factors in the - transverse plane (fig 4) being the most relevant. The two other potential variants with factors in the sagittal and coronal plane, respectively are included for completeness but are probably of less use and are placed below the transverse. All is denoted class X.

Generalizing this is gives a full body haptic device (class XI). Such is covering in a reasonable part of directions of the human body. An imaginable grid is wrapped around the body and at some crossings point a factor is placed. We could differentiate between simple factors each one being binary (class XI), which is technically more simple to make say by vibrators; or multimodal and continuous (class XII). We then open up for full body haptic displays (class XIII) where the aim is to mimic skin both so that the stimulus is resembling natural environment such as wind and cooling and by having a similar spatial resolution of the stimulus given. This call will probably be interesting to aim for from a scientific perspective whereas class XII is more realistic from a technical and product point of view.

Finally we denote ‘hyperskin’ (class XIV) as potentially a good term for a Haptic Display that goes beyond a multimodal full body skinlike HD. Here new modalities are added that gives new sensations using sensor switching. In addition, stereoscopic stimulus could be applied. We do not discuss hyperskin HD further in this text, but probably this will be employed for scientific interest in the future. It is also possible to expand here with new classes.

**Table 2.** Overview of factors suitable for integration into wearables and smart textiles. haptic display hierarchy ordered in a tentative complexity from the lowest (less informatical content) to higher (more informatical content). HD = haptic display.

	Suggested name of the display type	Spatial extension	Stimulus	Stimulus richness	Directional?	
XIV	Hyperskin	3D, full body	continuous	Multimodal and adding new senses	Yes. Detects up-down, back-front and left-right	iconic
XIII	Skinlike haptic display	3D, full body		Multimodal mimicking humans cutaneous senses	Yes. Detects up-down, back-front and left-right	iconic
XII	Full-body haptic display	3D	Continuous or binary	multimodal	yes	
XI	Full-body haptic display	3D	binary	monomodal	yes	

Continued

**Table 3.** Continued

	Suggested name of the display type	Spatial extension	Stimulus	Stimulus richness	Directional?	
Xa	Partly directional - transverse	bent 2D			Partly. Detects left-right	
Xb	Partly directional -sagital	bent 2D			Partly. Detects back-front	
Xc	Partly directional - coronal	bent 2D			Partly. Detects up-down	
IX	Haptic matrix with several tactors with different modalities, continuous or binary	2D	Binary or continuous array	multimodal	no	
VIII	Haptic matrix or Continuous matrix haptic display	2D	Continuous matrix	monomodal	no	
VII	Haptic matrix or Binary matrix haptic display	2D	Binary matrix	monomodal	no	
VI	Haptic array with several tactors with different modalities, continuous or binary	1 D	Binary or continuous array	multimodal	no	
V	Haptic array or Continuous array haptic display	1D	Continuous array	monomodal	no	
IV	Haptic array or Binary array haptic display	1D	Binary array	monomodal	no	
III	Point wise multimodal haptic display or zero dimensional multimodal HD	0 D	continuous	multimodal	no	
II	Point wise haptic display or zero dimensional HD	0 D	continuous	monomodal	no	
I	Point wise haptic display or zero dimensional HD or binary single point type	0 D	binary	monomodal	no	symbolic

As mentioned, 0D, can only convey some information that is synonymous with “Something is happening” but nothing more (like *what* is happening, if several things are happening, from where somethings is happening etc.). 0D is necessary indexical and symbolic in a semiotic parlour, whereas 2D could potentially also mimic reality in some very simple meaning and being iconic.

## APPLICATION

Textiles with its characteristics of being able to cover surfaces and volumes such as the human body is mostly used for garments. The enrichment of textiles with sensors and actuators, including haptic devices, is referred to as smart textiles. A novel type, a *chairable*, has been constructed which is a communication device covering furniture but aimed for humans, such as people having deafblindness or spectators at a sports event. It is a device that is close to the body without being a traditional worn wearable. Thanks to being textiles it fits on a broad variety of chairs, sofas, arm chairs etc. It has a 2D arrangement of tactors (vibrators) and is a binary matrix haptic display (class VII) in our HD hierarchy. In a pocket powering and electronics is gathered, not further described here. The chairable is wireless receiving signals. Incoming signals are translated to stimulus applied from the back

---

to the sitting user. Thus the chairable is allowing for having the hands and fingers free. It is a fully portable device that can even be folded. It can be mass-produced for arena or domestic use besides its main application as assistive device for persons having deafblindness.

### **ACKNOWLEDGMENT**

This work has been partially funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 780814 SUITCEYES and Familjen Kamprad stiftelse under LISA agreement.

# Gamification to Enhance the Mini-Conference: A Case Study From Researching Digital Cultural Heritage

Gautam Vishwanath and Lily Diaz-Kommonen

Aalto University School of Arts, Design and Architecture, Espoo, 02150, Finland

## ABSTRACT

Recent trends since the COVID-19 era indicate a rapid increase in remotely organized conferences; however, the remote-based nature of these types of events has gained notoriety for a lack of appeal particularly due to long and exhausting sessions without physical contact. On the other hand, previous research has demonstrated several benefits to well-designed games and gamification such as its ability in creating a state of flow by instilling motivation and rewarding its participants for overcoming challenges. As a case study in the cultural heritage sector, we organized a day-long remote conference known as the Mini-conference held among members of a consortium to which we belong. The Mini-conference applied gamification in combination with other relevant methodologies using online collaboration tools. This remotely organized event deployed four (4) collaborative workshops conducted by the different consortium partners. The results of the workshops presented in this paper demonstrate the use of gamification. The methods used were based on qualitative engagement, group participation and outputs emanating from each of the activities. They indicate positive effects of gamification in drawing interest and enhancing engagement among the remotely located participants. The contributions made by this paper include unique insights to the ongoing research on remotely organized conferences, especially through the purview of gamification. In addition, the paper also sheds light on methods that could be employed by museums within the domain of cultural heritage. These might be relevant to the “post-COVID-19 era” in which hybrid engagement consisting of physical and remote collaboration is becoming the norm.

**Keywords:** Collaborative workshops, Digital cultural heritage, Gamification, Motivational affordances, Remote conferencing

## INTRODUCTION

In the last decade, remote conferences have surged, boosted by technological advancements. The COVID-19 pandemic further fuelled this trend, as lockdowns and travel restrictions made physical gatherings challenging (Bozkurt and Sharma, 2020; Hacker et al., 2020). Online collaboration tools have made these remote gatherings more accessible, proving ideal for consortiums and case studies.

Poorly designed conferences and lengthy meetings often suffer from a reputation for being dull, monotonous, and exhausting for attendees (Lu and Abella, 2021). Conversely, well-designed games, including board games, card

games, and digital ones, offer continuous engagement, motivation through competition, obstacles, challenges, and rewards for overcoming them. **Gamification, the application of game-like elements such as competition and rewards in non-game settings** like remote conferences, has shown benefits across various domains (Hamari et al., 2014; Koivisto, 2017). Can it be developed further as a design tool to mitigate the challenges of lengthy and draining remote conferences? This paper explores this question and showcases the use of gamification in our Mini-conference held within a research project.

The project, SPICE (Social cohesion, Participation, and Inclusion through Cultural Engagement), operated within the H2020 European research framework, aiming to integrate advanced information processing tools with cultural heritage to foster social cohesion, participation, and inclusion among museum audiences (Bruni et al., 2020). It involved partners from research universities, museum institutions, and industry, focusing on cultural heritage, game design, and language technologies. The Mini-conference, co-hosted by project partners, specifically targeted case study museums within the project (Díaz-Kommonen et al., 2024).

In late 2020, during the initial phase of the project when physical meetings were impossible due to travel restrictions and lockdowns, we adapted by organizing this online event. The event featured four (4) collaborative workshops led by various consortium partners. The outcomes, discussed in this paper, utilized gamification and centred on qualitative engagement, group participation, and outputs from each activity.

The findings highlight the impact of gamification in generating interest and boosting engagement among participants dispersed remotely. This paper offers valuable contributions to the ongoing research on remotely organized conferences, particularly through the lens of gamification. These findings are particularly relevant in the evolving landscape of the “post-COVID-19 era”, where hybrid engagement involving both physical and remote collaboration is increasingly common (Sneader and Sternfels, 2020).

## **FIELDS OF STUDY AND APPROACH**

It is the case that an overabundance of research has examined gamification in collaborative workshops, often in relation to learning scenarios and the education sector (Dicheva et al., 2015; Stott and Neustaedter, 2013). However, **there has been a lack of study researching as well as critically examining the application of gamification exclusively within remotely organized conferences.**

A conference can be defined as a formal event wherein many people meet to discuss and engage with a specific topic, typically lasting several hours or days (van Venter, 2019). From our perspective, we see gamification as the incorporation of specific elements of games into non-game contexts including in conferences as well as in other work-related events such as, for example, in collaborative workshops. We are also aware that during the past decade, gamification has been gaining increasing prominence not only within

academia but in several sectors such as museums and cultural heritage (Bujari et al., 2017; Hammady et al., 2016; Ioannides et al., 2017).

### **Remote Conferences**

Videoconferencing traces its origins back to the early 1930s with AT&T's 'two-way television-telephone' system (MacDougall, 2006). However, it wasn't until the early 1990s that videoconferencing with multiple participants gained momentum. PictureTel conducted the first world's largest global video conference in 1995, involving around 50 sites worldwide (Lasic-Lazic et al., n.d.). The COVID-19 pandemic from early 2020 accelerated the adoption of cloud videoconferencing software like Zoom, driven by the increase in remote work and the shift of conferences from physical to virtual spaces (Hacker et al., 2020). Given this surge, this paper advocates for research into integrating gamification with remotely organized events, such as conferences.

### **Gamification**

Games are commonly associated with amusement and entertainment, structured by rules toward a goal. However, they also hold potential for motivating players and addressing real-life challenges (Marshall, 1999). Key elements of games include players, rules, mechanics, environment, and interface (Järvinen, 2008). Gamification extends these elements to non-game contexts (Deterding et al., 2011). The growing digital gaming industry and widespread adoption of gaming illustrate their impact (Şener et al., 2021; Sneader & Sternfels, 2020). McGonigal's book *Reality is Broken* highlights games' ability to foster play, curiosity, motivation, meaning, and reward, thereby enhancing participant engagement (McGonigal, 2011). To evaluate the latter, we utilized video recordings and qualitatively analyzed the workshop outputs.

### **Approaches Used in the Mini-Conference**

The Mini-conference relied on two multimedia conferencing components: videoconferencing for all attendees and real-time online collaboration among participants in workshop groups (Marshall, 1999; Schooler et al., 1991). Zoom was chosen for videoconferencing due to its widespread familiarity among participants and features like "breakout rooms" for dividing participants into groups. Additionally, Flinga facilitated real-time online collaboration (NordTouch n.d.). By integrating videoconferencing and gamified workshop activities, our goal was to keep participants engaged and motivated with frequent creative tasks, rather than passively listening throughout the potentially lengthy event.

Each workshop, co-designed by different consortium partners, aimed to foster inclusion and participation aligned with project goals. Gamification strategies included: setting clear goals, providing agency to participants through moderate facilitation, ensuring tasks were challenging yet achievable, implementing time-limits and limited selectable items for competition, using intuitive collaborative online tools for a seamless user experience, and

rewarding participants with acknowledgment and constructive criticism. Through this gamified approach, the workshops aimed to bridge physical gaps between partners by connecting them with case studies through collaborative creative endeavours.

## **METHODOLOGIES EMPLOYED**

In the Mini-conference, methodologies were carefully chosen to align with co-designed workshop contents and the multimedia conferencing context. These methods included gamification for motivation, collaborative brainstorming for creativity, and qualitative analysis of video recordings for output evaluation. Lessons from the Mini-conference experience are invaluable for future remote conferences, particularly those involving target audiences of case studies. To address potential shortcomings, participants were provided with pre-requisite tasks or homework to familiarize them with topics, concepts, and online tools, aiding in the successful implementation of methodologies.

### **Gamification**

Gamification was carefully integrated into the workshops to align with the cultural heritage-based goals of the research project. Rather than imposing it on tasks, specific game elements were selectively used to match the framework, design, content, and scope of the workshops. Additionally, the Mini-conference incorporated trivia game sessions between workshops to inject fun and humor into the event. Each session featured an entertaining multiple-choice question presented to all attendees, who were prompted to type the correct answer using Zoom's chat feature and promised a reward.

### **Group Formations**

Group discussions are a valuable source of knowledge, stimulating conversation and collaborative development of ideas (Flick, 2006). This dynamic is also pertinent to gamified activities, where participants collaborate and focus on common topics. At the Mini-conference, approximately thirty (30) participants were grouped based on their involvement with specific case studies, resulting in five (5) groups for each workshop activity, named after consortium case studies. Each group was facilitated by a moderator who guided discussions and occasionally assisted participants. Moderators were instructed to minimize interruptions, allowing participants to engage in gameplay and achieve workshop objectives. This approach intended to ensure participant comfort, maintain task control, foster new insights, and meet workshop requirements.

### **Video Recordings**

Facilitators were tasked with recording sessions held in their respective breakout rooms to ensure accessibility of data (video recordings) for all participants, facilitators, and organizers. On the positive side, analyzing this data proved fruitful, allowing observation of participant

engagement and the effects of gamification. Screen recordings of Flinga Whiteboard sessions provided insights into the brainstorming process. Despite potential concerns about data privacy under GDPR laws, no significant changes in participant behavior due to camera shyness were observed (GDPRhub n.d.). This lack of change could be attributed to the remote nature of the activities, subtle camera presence, and participants' familiarity with each other and the video recording tools.

## **DESIGN OF THE MINI-CONFERENCE**

The primary goal of the one-day Mini-conference was to advance the collaborative design of case studies, which had been hindered by the inability of consortium partners to meet physically due to COVID-19. This objective was achieved through a series of engaging thematic co-design workshops led by various consortium partners. The workshops were facilitated using multimedia conferencing and other online collaborative tools.

### **Use of Online Collaboration Tools**

For video conferencing, Zoom was selected as the primary tool, while Flinga facilitated group collaboration and gamified tasks during workshop activities. These tools were chosen based on specific criteria:

1. Facilitating collaborative whiteboarding for participants to generate ideas rapidly during brainstorming sessions.
2. Featuring simple and intuitive interfaces for ease of use, particularly for newcomers without the need for tutorials.
3. Accessibility via web browsers without requiring offline installation.
4. Adherence to ethical norms, including personal data protection and anonymity for participants by not mandating account registration.

In the Mini-conference, a Flinga Whiteboard supported participants' collaborative activities and brainstorming sessions enabling them to view each other's work, as well as facilitated competition for resources from a shared pool.

### **Outline of the Collaborative Workshops**

The activities in the workshops and their objectives were developed and refined collaboratively with other consortium members. Gamification and game elements were not always explicitly embedded in every activity. Rather, they were strategically blended within the workshops and other conference activities. However, a time limit for each activity was used as a 'common game obstacle' for all the participants.



**Table 1.** Planned workshop activities along with game elements in the Mini-conference.

Workshop (W) objectives	Activities (A) in separate groups	Incorporated game elements
W1 – Creating interpretive and fictional narratives about heritage artefacts and sharing the stories.	A1 - Presenting separate artefacts to the groups and tasking each participant to write a short fictional story based on how they might have interpreted the artefact in their preadolescent and adolescent years. A2 - Sharing stories amongst group members in the breakout rooms.	Rules: Participants presented with different sets of images of artefacts. Mechanics: Writing a short fictional story based on several images of artefacts and participants guessing which ones were presented while listening to stories. Goal: Completing the story. Rewards: Correctly guessing, receiving acknowledgement, and relating to each other's stories.
W2 – Persona design for different museums and end-user communities.	A1 – Elaborating on the aspects of a case study's pre-requisite persona designs. A2 - Selecting an unexpected or "surprise" visitor from a pool of fresh personas and brainstorming their attributes. A3 – Choosing characteristics of a case study's user model.	Components: Cards used in Flinga. Rule: Selecting a visitor. Mechanics: Collaborative design of a persona and selecting images from a common resource pool. Goal: Completing all the attributes for the persona within a time-limit. Reward: Most interesting and detailed group's persona awarded.
W3 – Curating an imaginary exhibition.	A1 - Curating using the cards in the UX map and the personas as references. A2 - Selecting a set of artefacts from a pool to be used for an imaginary exhibition. A3 - Organizing gathered artefact samples by arranging them to denote a storyline. A4 - Sharing exhibitions in the main room and narrating stories.	Components: Cards representing artefacts and a pre-defined UX map template presented in Flinga for participants to develop. Rule. Maximum limit of 12 for selection. Mechanics: Collaboratively arranging artefacts and developing the attributes of the UX map. Goal: Completing the UX map and curating an imaginary exhibition in a time-limit.
W4 – Evaluating an interface for an exhibition.	A1 – Discussing, rating, and selecting interfaces to be used by museum visitors.	Mechanics: Collaboratively rating pre-defined interfaces. Goal: Evaluating and selecting intended interfaces.

## RESULTS

Participant engagement was assessed through multiple measures:

1. Video documentation: Reviewing recorded videos provided insight into participant involvement and the impact of gamification on collaboration.
2. Analysis of workshop outputs: Reviewing the materials created by participants during the workshops offered tangible evidence of engagement and productivity.
3. Feedback analysis: Gathering feedback from participants after each workshop allowed for a direct assessment of their experience and engagement levels.

By combining these approaches, a comprehensive understanding of participant engagement was achieved, informing future conference planning and implementation strategies. The outputs derived from each workshop are summarized in Table 2.

**Table 2.** The outputs derived from each workshop along with the noticeable effects of gamification.

Workshop (W) objectives	Output(s)	Incorporated game elements
W1 – Narrative development	Narrative methods and artifact analysis allowed participants to share personal and fictional childhood stories. This established a benchmark for the potential application of these methods in workshops conducted by case study museums with their target audiences (Vishwanath, 2023).	The interaction among participants in groups revealed that creating and sharing stories around artifacts was engaging, relatable, challenging, and rewarding.
W2 – Persona development along with users and communities	Persona development focused on curating and evaluating systems for museum audiences' needs and testing proposed solutions. Using an unexpected visitor for the case study encouraged creative brainstorming by introducing a fresh perspective. Participants analyzed their museums to promote inclusivity and accessibility. Case studies employed unique personas, necessitating further testing and development.	Participants in groups cooperatively played to match and develop persona attributes. Creating attributes for an unexpected visitor prompted participants to imagine and craft new personas. After persona development, participants reconvened in the main room to share feedback, including acknowledgment, compliments, and constructive criticism.
W3 – Activities and contexts through pre-defined UX maps	Within their respective group's UX maps, participants identified barriers in the provided templates and iterated solutions focused on physical and cultural accessibility and inclusion. In addition to artifacts aligning with the theme and genre of a case study's museum, groups selected various other artifacts, indicating the museum's openness to multiple interpretations. Different order types, including non-linear storytelling, emerged in the arrangement of artifacts for the exhibition and these stories narrated by group representatives contained intriguing new ideas and fostered a creative outlook.	Groups competitively selected artifacts from a shared pool for their exhibition to prevent others from reserving them. Collaborative decision-making and arranging artifacts based on a story for the exhibition proved to be a playful process. Presenting the exhibition to all attendees by narrating its story upon completion was perceived as a rewarding experience.
W4 – User interface design	Each case study group selected their preliminary sets of interfaces for visits, interpretations, and reflections, and rated the majority of the interfaces presented to them.	The outcomes of gamification were limited to collaborative selection and rating within a defined time limit, with no explicit additional results.

## **DISCUSSION AND FUTURE WORK**

Reviewing the Mini-conference from a case study perspective, enabled us to identify specific characteristics of multimedia conferencing and the impact of gamification when integrated with other methodologies. These types of remotely organized conferences, utilizing gamification as a methodology, are likely to be employed in co-design workshops conducted with the target groups of the case studies.

### **Gamification**

Even though McGonigal's (2011) theories, drawing from Csikszentmihalyi (1990), were crucial in designing the Mini-conference, further validation regarding motivational aspects of gamification was needed. A literature review by Hamari et al. (2014) underscored the lack of coherent understanding and empirical works on gamification's motivational affordances and psychological outcomes. They noted positive results in various contexts, including intra-organizational systems, but cautioned that descriptive studies couldn't infer direct effects of gamification (Hamari et al., 2014). Similarly, while the Mini-conference participants responded positively to gamification, this paper aimed to quantify workshop outputs, revealing gamification's pivotal role in fostering motivation and engagement alongside other methodologies.

### **The Future for a Hybrid Mode of Collaboration**

A hybrid collaboration model, combining remote sessions and limited physical engagement, is anticipated to become commonplace post-COVID-19, notably in the cultural heritage sector (Bujari et al., 2017; Zbucnea et al., 2021). For instance, the Ars Electronica festival in 2020 hosted its first remotely organized exhibition using Mozilla Hubs (Doyle, 2020). Their Kepler Garden installation offered diverse 3D experiences accessible worldwide, spanning artistic, scientific, political, and cultural heritage themes. These immersive encounters were easily accessible via Mozilla Hubs on various devices, including desktops, laptops, smartphones, tablets, and VR headsets. This suggests vast potential for future remotely organized conferences to create novel, far-reaching experiences accessible to broader audiences.

## **LIMITATIONS**

This study has limitations primarily due to the qualitative nature of the workshop activities' outputs and the remote conferencing environment.

### **Remote Nature of Collaboration**

Participants' feedback highlighted that some tasks lacked clarity initially, requiring further clarification to align with the workshops' goals and rules. While having facilitators in each breakout room helped clarify tasks and improved online collaboration, it may have increased the need for facilitation in some activities and reduced voluntary participation. Additionally,

collaborative writing of lengthy texts within time constraints proved challenging for participants. As an alternative, we suggest implementing gamified mechanics such as responding with keywords or working on shorter texts in groups.

### **Gamification**

While participants found the activities engaging, as demonstrated by video recordings and other workshop outputs, increased engagement may not always directly correlate with gamification. Various factors, such as the effects of other methodologies or participants' familiarity with the topics, could influence engagement. Therefore, further research is needed to empirically validate the direct correlation between gamification and the enhancement of remote conferences.

### **Video Recordings**

Recording videos can pose challenges related to camera-consciousness and privacy concerns. Additionally, according to our data management plan, access to video data is strictly limited to consortium members. We explicitly informed all attendees that video recordings would be used solely for research purposes, such as analyzing the development process of workshop activities. However, some participants may feel uncomfortable being on video. To accommodate this, participants were given the option to turn off their camera and mute their microphone while engaging anonymously through online collaboration tools. Nonetheless, this may have hindered their participation. Fortunately, none of the participants expressed a desire to do so.

### **CONCLUSION**

Within our research project, we successfully designed and conducted a remotely organized Mini-conference incorporating multimedia conferencing characteristics. We suggest that implementing gamification in workshop activities enhances participants' motivation and contributes to valuable project outputs. The results offer unique insights into remote conferences and multimedia conferencing with gamification. While the empirical correlation between gamification and participant engagement was demonstrated, further follow-up is warranted. Nevertheless, the prevalence of remotely organized conferences is increasing, making the use of gamified strategies in events and collaborative workshops appealing to the Human Interaction and Emerging Technologies communities.

### **REFERENCES**

- Bozkurt, A., & Sharma, R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education*, 15(1), i-vi.
- Bruni, L. E., Daga, E., Damiano, R., Diaz, L., Kuflik, T., Lieto, A.,... Wecker, A. (2020). Towards advanced interfaces for citizen curation.

- Bujari, A., Ciman, M., Gaggi, O., & Palazzi, C. E. (2017). Using gamification to discover cultural heritage locations from geo-tagged photos. *Personal and Ubiquitous Computing*, 21(2), 235–252.
- GDPRhub. (n.d.). Article 2 GDPR. Retrieved from [https://gdprhub.eu/index.php?title=Article\\_2\\_GDPR](https://gdprhub.eu/index.php?title=Article_2_GDPR).
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990). New York: Harper & Row.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In *CHI'11 extended abstracts on human factors in computing systems* (pp. 2425–2428). ACM.
- Díaz-Kommonen, L., Svinhufvud, L., Thiel, S. and Vishwanath, G., 2024. Enriching Museum Collection with Virtual Design Objects and Community Narratives: Pop-up-VR Museum. *Collections*, p.15501906241233817.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society*, 18(3), 75–88.
- Doyle, D. (2020). Virtual creativity in COVID times. *Virtual Creativity*, 10(2), 137–140.
- Flick, U. (2006). *An introduction to qualitative research*. London: Sage Publications.
- Hacker, J., vom Brocke, J., Handali, J., Otto, M., & Schneider, J. (2020). Virtually in this together—how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. *European Journal of Information Systems*, 29(5), 563–584.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). Does gamification work?—a literature review of empirical studies on gamification. In *2014 47th Hawaii international conference on system sciences* (pp. 3025–3034). IEEE.
- Hammady, R., Ma, M., & Temple, N. (2016, September). Augmented reality and gamification in heritage museums. In *Joint International Conference on Serious Games* (pp. 181–187). Springer, Cham.
- Ioannides, M., Magnenat-Thalmann, N., & Papagiannakis, G. (2017). *Mixed reality and gamification for cultural heritage* (Vol. 2). Cham, Switzerland: Springer.
- Järvinen, A. (2008). *Games without frontiers: Theories and methods for game studies and design*. Tampere University Press.
- Koivisto, J. (2017). *Gamification: A study on users, benefits and literature*.
- Lasic-Lazic, J., Stancic, H., & Banek, M. (n.d.). Audio-video conferences in the business and educational environment. In *3rd CARNet Users Conference*, Zagreb, Croatia.
- Lu, Y., & Abella, D. (2021). The Race to Fix Virtual Meetings. *The New York Times*. Retrieved from <https://www.nytimes.com/2021/02/17/magazine/video-conference.html>.
- MacDougall, R. (2006). Long lines: AT&T's long-distance network as an organizational and political strategy. *Business History Review*, 297–327.
- Marshall, D. (1999). *Multimedia Conferencing*. Retrieved from <http://users.cs.cf.ac.uk/Dave.Marshall/Multimedia/node474.html>.
- McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.
- NordTouch. (n.d.). *Flinga*. Retrieved from <https://flinga.fi/tools>.
- Şener, D., Yalçın, T., & Klokmose, C. (2021). The Impact of Covid-19 on the Video Game Industry. Available at SSRN 3766147.

- Schooler, E. M., Casner, S. L., & Postel, J. (1991, January). Multimedia conferencing: Has it come of age?. In *Proceedings of the Twenty-Fourth Annual Hawaii International Conference on System Sciences* (pp. 707–716). IEEE.
- Sneider, K., & Sternfels, B. (2020). *From surviving to thriving: Reimagining the post-COVID-19 return*. McKinsey Global Institute.
- Stott, A., & Neustaedter, C. (2013). *Analysis of gamification in education*. Surrey, BC, Canada, 8(36).
- van Venter, R. d. (2019). Is there a difference between a congress, conference, symposium, seminar, journal club, and workshop in terms of continuing professional development activities? *South African Radiographer*, 57(1), 24–26.
- Vishwanath, G., 2023, June. *Enhancing Engagement through Digital Cultural Heritage: A Case Study about Senior Citizens using a Virtual Reality Museum*. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* (pp. 150–156).
- Zbucnea, A., Romanelli, M., & Bira, M. (2021). *Through the public's Lens: Are museums active members of society? An investigation during the COVID-19 pandemic*. *Cultural Initiatives for Sustainable Development: Management, Participation and Entrepreneurship in the Cultural and Creative Sector*, 61.

# UX Sustainability in AI-Infused Objects: A Systematic Literature Review of Available Tools for Designers

**Alice Paracolli and Venanzio Arquilla**

Design Department, Politecnico di Milano, Milano, Italy

## ABSTRACT

The rapid integration of AI-infused objects into our daily lives, as part of the growing Internet of Things (IoT) ecosystem, is transforming common appliances into sophisticated and interconnected systems (ITU, 2020). With projections indicating an increase from 5 billion objects in 2020 to over 200 billion by 2030 (CISCO, 2020), these AI-infused objects create expansive networks of data-consuming devices that persist indefinitely (Crawford, 2018). This surge necessitates a deeper understanding of their ongoing environmental impact, particularly during the use phase. Recognizing the potential for user experience Designers to adjust interactions to mitigate the environmental impact during the use phase of AI-infused objects, we conducted a systematic literature review to pinpoint the Design tools that can assist Designers in this effort. Our systematic literature review aims to identify Design tools that evaluate the sustainability of User Experience in IoT products. We analyzed 24 sources dedicated to sustainability from a User Experience perspective, and 22 that assess UX in IoT devices. The findings reveal a strong focus on product-focused evaluation tools, with general emphasis on User Experience and the usage ecosystem of these objects. As AI-infused objects become increasingly prevalent, it is essential for Designers to gain a comprehensive understanding of the environmental impacts and their cause. This awareness could lead Designers to integrate both technological advancements and environmental considerations effectively into their Design process.

**Keywords:** User experience, Sustainability, AI systems, Design, HCI, Tools, IoT

## INTRODUCTION

The concept of User Experience (UX) has greatly evolved since it was coined by Norman, now serving as a critical lens for both evaluating and designing products, according to Kerr (2015). Through the lens of UX, designers have the opportunity to assess and shape their work, taking into account not only the user's context and needs, but also considering the broader environmental impact of interactions (Paracolli, 2023). The environmental impact of UX, particularly with AI-Infused Objects, spans both direct effects, like energy consumption during use, and indirect effects, such as lifestyle changes induced by the product, as outlined by Ligozat (2022), Shehabi (2017), and Pohl (2019). However, Pohl et al. highlight a gap in incorporating user-related and behavioural effects into life cycle assessments (LCA), underscoring the need for a more holistic approach to sustainability in Design. The User Experience Environmental Impact is the environmental impact generated by

the user's adopting and interacting with the AI-infused object ecosystem. The impact of user interaction with digital products academically has been explored in Sustainable Interaction Design over the last 15 years (Blevis, 2007), which divides the field into Sustainability in Design and Sustainability through Design. However, these two perspectives could benefit of the comprehensive vision that the lens of UX provides, through analysing user context and needs. Hence, a shift in perspective is required: it's not enough to design highly efficient or minimally impactful products or objects that guide users towards more sustainable behavior, but understand how the objects are used to reduce the impact of that phase while supporting the user in their daily life. This research intends to bridge by providing a systematic review of Design tools that consider sustainability in UX Design, offering practitioners resources to create more sustainable experiences in AI-Infused objects ecosystems within the domestic settings. We consider Design tools as all the sources available to Designers that can support practical activities, such as toolkits, tips, checklists, guides, guidelines, manifestos etc. Design tool is defined as a physical or digital object used to achieve a specific goal, resulting from intentional human transformative processes, applied through a set of instructions. Instructions ensure replicability in various situations, establishing the Design tool's validity as a method. Notably, Design agencies like IDEO, MJV, and Frog often organise purpose-specific tools into toolkits, complete with guidelines on their application (Bruno & Mattioli, 2022). Toolkit research aims to simplify the Design process, empower new users, and foster innovation (Ledo, 2019). Toolkits serve as crucial educational and collaborative resources, especially in fields like sustainability and IoT Design since they require a systematic thinking. However, the wide variety and lack of critical testing of these tools, as noted by Nebeling (2017) and Remy (2018), create a challenge for practitioners, leading to an entropy of sources. This study present a systematic analysis of Design tools dedicated to "UX and Sustainability" and "UX and IOT" to help Designers navigate this entropy, focusing on guiding AI-Infused Objects Designers to assess and design with sustainability in mind. Highlighting the Design community's responsibility towards environmental impacts, it draws on the principles of responsible Design advocated by Papanek (1995) and Monteiro (2019). Despite the wealth of tools for UX Design, there's a gap in addressing the environmental consequences of user interactions in Design processes.

## **METHODOLOGY**

We conducted a comprehensive review of existing literature, employing a multivocal literature review (MLR) approach following Garousi et al. (2019). This methodology extends beyond traditional systematic literature reviews by incorporating both gray literature — including non-peer-reviewed materials like industry toolkits, expert blogs, seminars, materials from professional practices and manifestos — and white literature, which encompasses peer-reviewed sources such as articles from conferences and journals, and frameworks. Our goal is to achieve a more comprehensive understanding of a the perspective of UX to design more sustainable AI-Infused objects. This expansive approach was critical for understanding the breadth of strategies



in both the academic and practical realms. The resulting analysis provided a foundation for selecting frameworks and methods that align UX in IoT with environmental sustainability, to create user experiences that are conscious of their ecological footprint.

### Information Sources and Inclusion Criteria

Academic literature was sourced from Google Scholar, Scopus, and the IEEE library. To incorporate practical insights and approaches emerging directly from industry practitioners, additional searches were conducted using Google Search and LinkedIn. The research was guided by two main pathways: (1) approaches and tools focused on enhancing user experience for more environmental sustainable design. (2) Approaches for designing AI-infused objects within domestic settings through the lens of UX. Sources that aimed at supporting the Design process within either of these pathways were considered for inclusion. Criteria for selection included the requirement that sources be in English, fully accessible, and validated—either through peer-review for academic contributions or via practical case studies for practitioner-derived insights, or both; understandable and usable by Designers during the Design process. This stringent selection process was designed to ensure the reliability and relevance of the information incorporated into our study.

### Search Terms for Queries

Derived from the research question (What approaches are available to analyze and design the use phase of AI-Infused Objects with environmental sustainability in mind?), search terms were chosen to cover two main areas of study: user experience (UX) and its intersection with sustainability; the Internet of Things (IoT) from a UX perspective. Notably, the scope of the research was broadened to encompass IoT due to the relative scarcity of literature specifically addressing the term “AI-Infused Objects,” which falls under the broader IoT domain. Table 1 presents these terms along with their synonyms.

**Table 1.** Research terms, inclusion criteria and no of sources considered.

	“UX & IOT”	“UX & Sustainability”
Search in	Title, Abstract, Keywords	Title, Abstract, Keywords
1. Terms & synonyms	lot, Internet Of Things, AI Infused Objects, User Experience, UX,	Sustainability, Environmental Impact, Sustainable, Human Computer Interaction, SHCI, UX, User Experience,
2. Terms & synonyms	Tool, Toolkit, Methodology, Method, Guidelines, Manifesto, Framework	Tool, Toolkit, Methodology, Method, Guidelines, Manifesto, Framework
Source available at	IEEE, Scopus, Google Scholar, LinkedIn, Google Search	IEEE, Scopus, Google Scholar, LinkedIn, Google Search
Inclusion Criteria	English; understandable by designer; practical to be inserted in design process	English; understandable by designer; practical to be inserted in design process
No of sources considered	22	24

### Overview of Sources

Our literature review encompassed a total of 46 sources, divided into two main themes: “UX & IoT” and “UX & Sustainability”. For UX & IoT, we analysed 22 sources, including 20 from the white literature and 2 from the gray literature, spanning the years 2012 to 2023. There is an abundance of manifestos in the literature but Fritsch et al. (2018) revised 28 different IoT manifestos, identifying a notable absence of thorough context analysis for sustainability during the use phase, additionally a UX perspective gap is highlighted in the manifesto by De Roeck (2012). Although these manifestos offer insights that could guide Designers toward a better future with IoT, they generally lack practical, step-by-step guidance for Designers. The manifesto we found dedicated to sustainability in IoT, focuses mainly on physical characteristics considering end-of life and repair of the product, rather than on user impact during the usage phase (Stead, 2019).

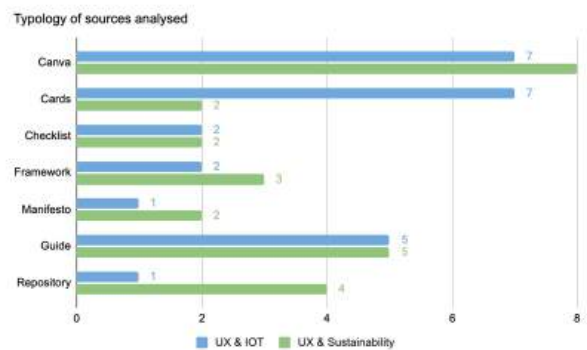


Figure 1: Typology of sources analysed. NB: each source might contain more tools.

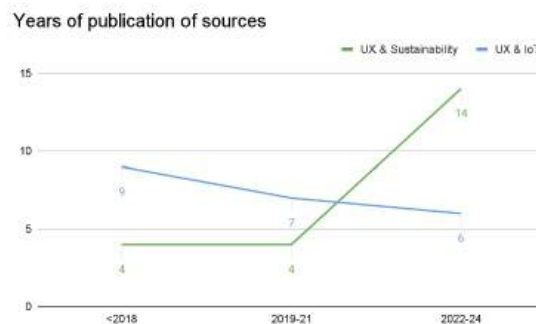


Figure 2: Sources by year of publication.

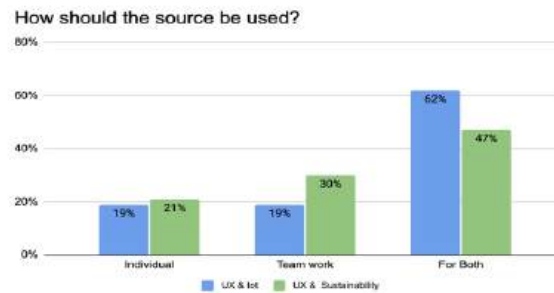
For UX & Sustainability, our review included 24 sources, with 6 from the white literature and 18 from the grey literature, published between 2010 and 2024. We observed a significant increase in grey literature from 2022

to 2024, reflecting the growing adoption of Sustainable UX practices among practitioners. This trend is especially evident in the design of web pages aimed at being effective and responsive while reducing load times. Such sustainable practices are not only effective but also necessary, prompting practitioners to develop a variety of toolkits to address these challenges. None of the sources referred to AI-infused objects or more in general to IoT.

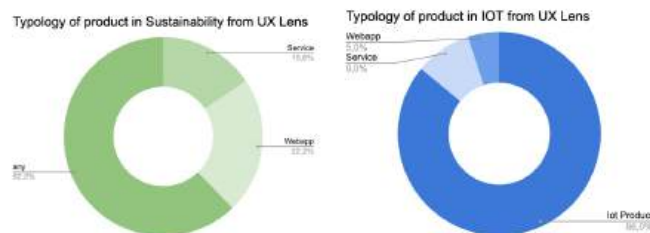
### CATEGORIES AND SOURCES ANALYSIS

Our analysis of Design approaches has been readapted using the categories established by Hsieh (2023). The first is the “creative inspiration” category of Design tools. These tools are intended to stimulate creativity and encourage out-of-the-box thinking by providing provocative prompts or visuals to ignite ideas (Appendix A & B). Right after, the category “individually or in teams” enabling Designers to swiftly determine the appropriate mode of usage based on the project’s goals (Fig. 3).

We assessed the types of products that could be analysed (web apps, IoT, services), helping Designers navigate the array of tools available and quickly identify the most suitable for their specific design objective (Fig. 4, 5). During the analysis, we categorized the tool by indicating in which phase of the Design process each resource would be useful. We followed and partially adapted the Design process phases proposed by Hsieh (2023) (Table 2), adding “Redesign”.



**Figure 3:** Utilisation of sources: “individual vs teamwork”

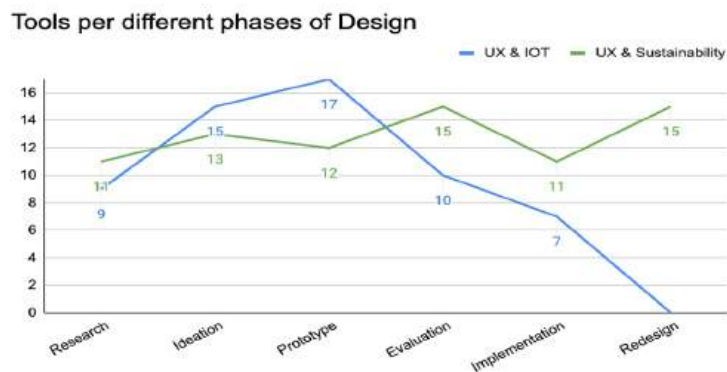


**Figure 4 & 5:** Kind of product to be analysed by sources

**Table 2.** Design process phase and description.

Process Phase	Description
Meta-Design	Meta-Design goes beyond traditional process models by facilitating environments where users also function as designers, fostering creativity beyond conventional boundaries (Hsieh, 2023).
Research	This initial phase involves a deep dive into the problem space to understand the stakeholders and their challenges. Insights gained here are crucial for informing subsequent Design decisions.
Ideation	Designers engage in creative brainstorming to generate ideas. It includes Tools that provide inspiration or insights about human behavior and problems.
Prototyping	Practitioners convert their concepts or potential solutions into low or high fidelity prototypes to explore the problem space the Design aims to address and evaluate their feasibility.
Evaluation	Focuses on testing how well the Design solutions work with actual users and context, aiming to check usability and effectiveness.
Implementation	Involves detailing the prototype for production, collaborating with manufacturers and developers to finalize the design.
Redesign	It aims to review and refine existing products, high-fidelity prototypes, or well-defined concepts. Tools in this section should facilitate understanding of potential pain points in the product/prototype/concept idea and enable targeted adjustments for improvement.

An overview of the source distribution is visible in Fig 6. The “meta-Design” category is not included because only one source fell into this category: Ambe (2019).



**Figure 6:** Source distribution per design process phase.

The graph demonstrates that “UX & IoT” tools are more concentrated in the early stages of the Design process, particularly during ideation, while UX & Sustainability tools have a more consistent distribution, emphasizing

the evaluation phase and maintaining presence through the implementation and redesign phases. This could imply that sustainability considerations in UX are relevant throughout the entire Design process, whereas the focus on IoT within UX seems to diminish as projects move towards implementation and post-launch stages. The data presented in the graph indicates a well-established foundation of tools for sustainability and UX in the early phases of design. However, there appears to be a notable decline in tools addressing these aspects in the later stages of IoT projects, specifically during prototype evaluation, implementation, and redesign. Recognizing this trend, there is an opportunity to develop a unified toolkit that can lead Designers to re-evaluate existing IoT projects through a sustainability-focused UX lens. Such a toolkit would capitalize on the iterative essence of UX practices, facilitating the refinement of IoT designs to reduce environmental impact.

### **UX Approaches to Sustainability**

In digital products, the use phase's environmental impact is often overlooked, as evidenced by prevailing sustainability strategies. Circular approaches, for instance, typically concentrate on product reuse and recycling but tend to fail to address energy consumption during usage. Modular and end-of-life tactics likewise focus on tangible impacts, neglecting operational environmental effects. Sustainable behavior strategies, while promoting eco-friendlier actions, lack robust assessment methods to verify genuine enhancements in sustainability. Even in life cycle assessments (LCAs), though designed to consider the use phase, this is frequently ignored, underscoring a widespread gap in evaluating environmental impacts comprehensively (Pohl, 2019).

The Sustainable UX Network, established in 2021 by a group of practitioners (Jonas, 2021), is dedicated to promoting sustainability in UX and through UX. It aligns with the UN's Sustainable Development Goals (SDGs) and non-profit objectives, advocating for a reduction in negative impacts such as carbon footprint, energy usage, and waste, while encouraging positive outcomes like green solutions and fair practices. This perspective has been anticipated in the academic field of Sustainable Human-Computer Interaction (SHCI), where there is a distinction between sustainability in Design (SiD) and sustainability through Design (StD). SiD adheres to Blevis's (2007) concept of minimizing environmental impact through product design, whereas StD aims to modify user behaviors and enhance awareness. However, many StD projects face challenges in empirically assessing behavior change. Nonetheless, Basyouny & Männik (2023) highlight a knowledge gap among young Interaction Designers in incorporating sustainability into digital artifact Design processes, which contributes to greater energy usage and carbon emissions. This underscores the urgent need for educational modules on sustainable digital Design practices and practical tools to aid Designers.

We categorized our source analysis following the perspectives of SiD and StD and then specified if each source had focus on product (more product-oriented viewpoints focusing on what serves the Designer in creating the

product) or UX (aiming to support the designer in understanding the user’s context and needs).

We identified in SiD perspectives a greater focus on product-oriented tools (n°7), suggesting an emphasis on the product’s inherent design sustainability. However, only a few tools concentrate on UX (n°3), indicating room for growth, especially given the use phase often-neglected impact. The absence of tools addressing both UX and Product points to a lack of holistic approaches within this category. The analysis indicates a limited focus on Designing for sustainable behavior, likely because such considerations often fall outside the traditional scope of UX Design, thus fewer studies of this nature are represented. Consequently, there’s a clear necessity for additional UX tools that evaluate the environmental impact of a product from a UX standpoint.

The perspective and focus used in Sustainability sources

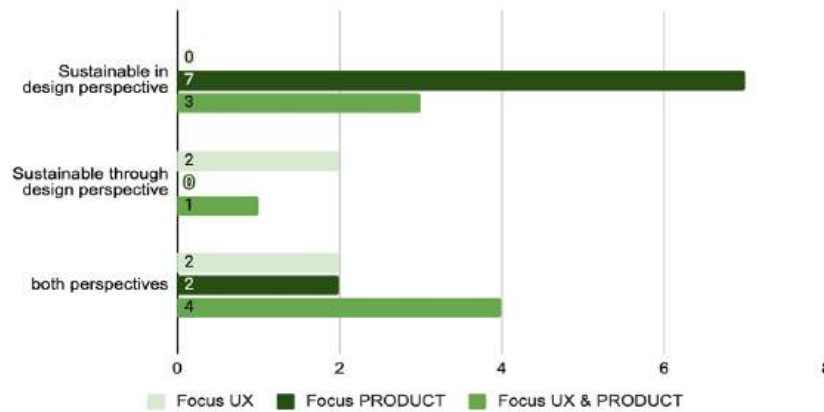


Figure 7: Perspectives (SiD,StD) & focus (UX, Products) in “UX & Sustainability” sources.

### UX Approaches in IoT

In the realm of IoT, one of the most significant challenges is the inherent complexity related to the number of involved technologies, interfaces, data points, and interactions among various devices, platforms, and stakeholders. Additionally, implementing the appropriate UX for IoT systems becomes particularly challenging due to the diverse needs of different stakeholder groups. This complexity is discussed in works such as Trendowicz (2023) for IoT and Paracolli (2023) when applied to sustainability evaluations. Indeed, evaluating the sustainability of IoT designs is further complicated by factors such as user behavior, including the frequency and scope of system use, as well as the volume of data processed. These variables heavily depend on the context in which the user operates, usually observable only once the product is in use. Consequently, the evaluation and redesign phases become crucial for assessing the environmental impact during the use phase of AI-infused objects. This is the domestication process, Design and domestication are the two components of innovation. Domestication is anticipated by Design, and

Design is completed in domestication, in which the Designer outline the object's meaning for the user and how it will be used (Silverstone, 1996). However, the object acquires whole meaning after, in the domestication process, which must be studied and then manipulated by the Designer with the aim to reduce the impact of the user experience, making the context of use fundamental part of the sustainability evaluation.

The role of data acquires an important facet both environmentally and economically, among the Design tools we analyzed, "The Definitive Ecosystem Design Toolkit" (2024) (Appendix B) is particularly notable. Although primarily business-oriented, this toolkit can be adapted for IoT applications, focusing on stakeholder ecosystems for digital services. It encourages reflection on the role of data not only economically but also in the design of efficient IoT ecosystems where devices communicate securely and collaborate to achieve common goals. Such approach can enhance interoperability, energy efficiency, and security of the IoT ecosystem. However, there is still a need for tools that help designers quantify and understand the impact generated by data.

As shown in Fig. 6, most sources in "UX & IoT" prioritize the early stages of the Design process, often neglecting the evaluation and redesign aspects, which are essential for sustainability. These phases are crucial for reducing object production, possibly by updating existing products rather than creating new ones, and altering user experiences to minimize environmental impact.

Benchmarking aids in comprehending existing market precedents to prevent redundancy in product/service creation and to learn from previous successes and failures. Among the resources reviewed, Vitali (2018), De Roeck (2019), and Aspiala (2016) offer tools ideal for this analysis, emphasizing market trends, user context, and IoT ecosystem assessment respectively. Furthermore, benchmarking highlights how similar products have been utilized, having the potential to encourage a reconsideration of a more sustainable user experience. Yet, a structured system for evaluation or comparison is lacking, which is crucial to support Designers in their decision-making processes. In practice, especially within companies, Designers rarely develop entirely new products from scratch; they often must consider multiple factors such as feasibility, budget, and stakeholder demands (e.g. in Barbosa-Hughes, 2019). This highlights the need for comprehensive, practical tools that encompass both ideation and thorough user research, involving stakeholders without neglecting the users themselves (e.g. in Ambe, 2019).

## CONCLUSION

According to Maeda (2019), key aspects for practitioners in Computational Design include a profound understanding of computation, a critical approach to technology, the ability to apply classic Design principles focused on functionality and industrial concerns, and the incorporation of Design Thinking to meet user needs and develop feasible products. Importantly, designers should also take responsibility for the environmental impacts of

their creations. Specifically for digital products, it is crucial to consider the environmental impacts during the user phase. In our study, we observed a significant gap in Design tools that help designers evaluate the environmental impacts of their user experiences, particularly within IoT contexts. To effectively assess these impacts, it is essential first to understand the technological foundations of the ecosystem of connected objects. Analyzing user context to fully comprehend how users interact with these systems, including the frequency and duration of interactions and what motivates these interactions. Context also helps interpret the data transmitted within the ecosystem of connected objects, such as what data users check, use, and why. Indeed, the duration and frequency of user interactions, as well as how often users access or utilize data, directly influence the environmental impact of the user experience. Therefore, there is a pressing need for tools that enable designers to assess environmental impacts in a comprehensive and intuitive manner. Our analysis also highlighted a demand for tools that support designers in redesigning or updating IoT products. Such tools would encourage designers and businesses to update existing solutions, making them more meaningful for users through UX analysis and, simultaneously, more sustainable.

## APPENDICES

### Appendix A

Source	Type of source	180-81															Sustainability			Other			Prospective	
		Carbon footprint	Individual items	Embodied energy	Resource use	Waste	Plastology	Materials	Substrates	Materials	Design	By Design	Other	Design	Other	Other	Other	Other						
Thomson, T. (2004). "A 100-Year History of the Environmental Design Process." <i>Environmental Design Process</i> , pp. 10-15. URL: <a href="#">http://www.environmental-design.com/100-year-history-of-the-environmental-design-process</a>	Journal	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
Wang, C. (2021). "The Future of Sustainable Design." <i>Environmental Design</i> , pp. 1-10. URL: <a href="#">http://www.environmental-design.com/the-future-of-sustainable-design</a>	Journal	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2021). "2021 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2020). "2020 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2019). "2019 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2018). "2018 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2017). "2017 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2016). "2016 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2015). "2015 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2014). "2014 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2013). "2013 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2012). "2012 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					
McGraw-Hill Construction Research & Analytics (2011). "2011 GreenSource Report." URL: <a href="#">https://www.mcgraw-hill.com/resources/green-source-report</a>	Report	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes					

Figure 8: Analysed sources “UX & Sustainability.”



## Appendix B

Source	Typology	Type of access	Academics	Fiat	WIoT				WHIoT					Perspective		
					Individual	Team	Professional	Company	Project	Process	Subject	Product	Interaction	Information	Feedback	Field of product
Benabib, A., Strain, E. C., Herzig, J., Shelton, J., Bhatt, H., Stroh, S., & Kwon, T. (2023). User experience and performance indicators for intelligent systems. <i>AGATHÓN International Journal of Digital Business</i> , 15(2023).	Article	Free	Yes	Co-creation with stakeholders												
The definable ecosystem design model (EDM) is available at <a href="https://www.designswarm.com/know-cards/also-touched">https://www.designswarm.com/know-cards/also-touched</a> .	Article	Free	Yes	Both												
User Experience Toolkit for Insights (IoT and Industry IoT) [Document]. Available at: <a href="https://www.designswarm.com/know-cards/also-touched">https://www.designswarm.com/know-cards/also-touched</a> .	Repository	Free	Yes	Co-creation with stakeholders												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												
Chen, Y., & Paracoli, A. (2023). User experience and performance indicators for intelligent systems. In <i>Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems</i> (pp. 503-512).	Article + Cards	Free	Yes	Both												

Figure 9: Analysed sources “UX & IoT.”

## REFERENCES

Ambe, A. H., Breteron, M., Soro, A., Chai, M. Z., Buys, L., & Roe, P. (2019, May). Older people inventing their personal internet of things with the IoT un-kit experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–15).

Arquilla, V., & Paracoli, A. (2023). User experience design and sustainability of AI-infused objects. *AGATHÓN International Journal of Architecture, Art and Design*, 13, 259–268. doi: 10.19229/2464-9309/13222023.

Aspijala, Deschamps-Sonsino, (2016), Learn. Play. Collect. Know Cards. url: <https://www.designswarm.com/know-cards/> ALSO TOUCHED <https://www.diva-porta.l.org/smash/get/diva2:1588866/FULLTEXT01.pdf>

Basyouny, M., & Männik, A. (2023). Unleashing the Awareness of Sustainable Human-Computer Interaction (HCI) Among Youth: Bridging the Knowledge Gap and Empowering Future Sustainable Designers.

Blevins, E. (2007, April). Sustainable interaction design: Invention & disposal, renewal & reuse. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 503–512).

Bruno, C., Gramegna, S. M., Mattioli, F., Monna, V., & Rossi, M. (2022). Design Tools: A reflection towards the definition of a taxonomy (p. 94). FrancoAngeli.

Crawford, K., & Joler, V. (2018). *Anatomy of an AI System*. Retrieved September, 18, 2018.

- De Roeck, D., Slegers, K., Criel, J., Godon, M., Claeys, L., Kilpi, K., & Jacobs, A. (2012, October). I would DiYSE for it! A manifesto for do-it-yourself internet-of-things creation. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (pp. 170–179).
- Deroeck, D. (2019). Ideas of things: The IOT design Kit. In *DIS '19 Companion: Companion publication of the 2019 on designing interactive systems conference*, pp. 159–163. <https://doi.org/10.1145/3301019.33238>
- Fritsch, E., Shklovski, I., & Douglas-Jones, R. (2018, April). Calling for a revolution: An analysis of IoT manifestos. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–13).
- Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and software technology*, 106, 101–121.
- Hsieh, G., Halperin, B. A., Schmitz, E., Chew, Y. N., & Tseng, Y. C. (2023, April). What is in the cards: Exploring uses, patterns, and trends in design cards. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–18).
- ITU-T Rec. L.1410 (12/2014) Methodology for environmental life cycle assessments of information and communication technology goods, networks and services available at: [https://www.itu.int/rec/dologin\\_pub.asp?lang=f&id=T-REC-L.1410-201412-I!!PDF-E&type=items](https://www.itu.int/rec/dologin_pub.asp?lang=f&id=T-REC-L.1410-201412-I!!PDF-E&type=items)
- Jonas, T. (2021) *The Sustainable UX Network* available <https://sustainableuxnetwork.com/>.
- Ledo, D., Houben, S., Vermeulen, J., Marquardt, N., Oehlberg, L., & Greenberg, S. (2018, April). Evaluation strategies for HCI toolkit research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–17).
- Ligozat, A. L., Lefèvre, J., Bugeau, A., & Combaz, J. (2022). Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions. *Sustainability*, 14(9), 5172.
- Maeda J., *Design in Tech Report*, (2019) p. 11, available at <https://designintechreport/>
- Monteiro, M. (2019). Ruined by design: How designers destroyed the world, and what we can do to fix it. *Mule Design*.
- Nebeling, M. (2017, May). Playing the Tricky Game of Toolkits Research. In workshop on HCI. Tools at CHI.
- Papanek, V. (1971). *What to design and why*. Selskabet for Industriel Formgivning.
- Pohl, J., Hilty, L. M., & Finkbeiner, M. (2019). How LCA contributes to the environmental assessment of higher order effects of ICT application: A review of different approaches. *Journal of cleaner production*, 219, 698–712.
- Remy, C., Bates, O., Dix, A., Thomas, V., Hazas, M., Friday, A., & Huang, E. M. (2018, April). Evaluation beyond usability: Validating sustainable HCI research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).
- Silverstone, R., & Haddon, L. (1996). Design and the domestication of information and communication technologies: Technical change and everyday life.
- Shehabi, A. (2017). Data Clouds and the Environment. In *Routledge Handbook of Sustainable Design* (pp. 170–178). Routledge.
- Stead, M., Coulton, P., & Lindley, J. (2019). Spimes not things. Creating a design manifesto for a sustainable internet of things. *The Design Journal*, 22(sup1), 2133–2152.
- Vitali, I. (2018) *Mapping the IoT*, available at <https://rri-tools.eu/-/mapping-the-iot-toolkit>.

# Data Visualization in the Public Energy Sector: A Study on User Experience and Satisfaction

**Amanda Andrade Lentez and Gabriela Botelho Mager**

State University of Santa Catarina (UDESC - CEART), Florianópolis, SC 88.035-901, Brazil

## ABSTRACT

This study delves into the analysis of how data visualization impacts user experience (UX) when interacting with public electrical energy data. In the era of big data, the volumes of information on energy generation, distribution, and consumption are expanding exponentially. This surge underscores the urgency for presentation techniques that not only simplify complex datasets for the lay audience but also improve engagement and comprehension. Thus, the study aims to bridge the existing knowledge gap by identifying effective data visualization strategies that enhance the public's ability to understand intricate data, thereby supporting informed decision-making and heightening awareness about energy sustainability. Adopting a mixed-methods approach, the study integrates extensive literature reviews with empirical usability testing involving 30 participants. To complement the quantitative findings, qualitative insights were extracted from interviews and focus groups, aiming to capture user preferences, challenges encountered, and suggestions for improvement. This analysis covered the effectiveness of various visualization components, including filters, information hierarchies, graphical elements, and data diversity, in facilitating an intuitive grasp of electrical energy data. The study showed a correlation between intuitive visualization techniques and the improvement of UX metrics such as engagement, comprehension, and satisfaction. Key findings emphasized that features such as interactive filters and good information hierarchies are instrumental in empowering users to effectively navigate and interpret electrical energy data. The study culminates in the formulation of eleven targeted guidelines for the development of user-centric data visualizations within the public energy sector.

**Keywords:** Data visualization, Human factors, Cognitive psychology, User experience, Public energy data

## INTRODUCTION

In the contemporary landscape marked by the continuous proliferation of data, the quantity of accessible information is unparalleled. This phenomenon is particularly pronounced in sectors where public data influences social decision-making and awareness, such as the energy sector. Herein, data pertaining to energy generation, distribution, and consumption is not only voluminous but complex, posing a challenge for the lay audience to comprehend without the aid of effective presentation strategies (Tufte,

1997; Few, 2009). The complexity accentuates the need for advanced data visualization techniques that refine intricate datasets, thereby augmenting public engagement and comprehension. Data visualization transcends mere aesthetic presentation; it emerges as a facilitator of user experience (UX), enabling informed decision-making and raising awareness on critical matters like energy sustainability (Cairo, 2012; Yau, 2013).

The present study explores the impact of data visualization's on UX, specifically in the context of interactions with public energy data. The study is dedicated to identifying data visualization strategies that increase the public's capability to understand complex public energy data. Concentrating on UX metrics such as engagement, comprehension, and satisfaction, the research endeavors to discover insights on optimizing data visualization to foster public understanding and participation in energy sustainability dialogues.

Employing a mixed-methods approach, the research integrates an extensive literature review with an empirical usability testing involving thirty participants (Nielsen, 1993; Shneiderman, 1996). This methodology facilitates a thorough exploration of the influences exerted by various data visualization elements—such as filters, information hierarchies, graphical elements, and data diversity—on user satisfaction and engagement (Lentez, 2020). Through interactive sessions of usability testing with existing data visualization dashboards displaying electrical energy datasets, performance was measured using task completion times, error rates, and subjective usability scores derived from questionnaires. Complementing these quantitative findings, qualitative insights were extracted from interviews and focus groups, in order to identify user preferences, encountered challenges, and recommendations for improvement of the dashboards.

This methodical approach not only elucidates the efficacy of diverse visualization components in enabling an intuitive comprehension of electrical energy data but also highlights the role of user-centered design in devising data visualizations for the public sphere (Lentez, 2020). By bridging the gap between complex datasets and the general population, well-conceived data visualizations hold the potential to promote the discourse on energy issues. Thus, the study contributes to the burgeoning domain of data visualization by furnishing academically and practically pertinent insights, guiding future design initiatives to resonate with user needs and preferences more closely in the realm of public energy data (Norman, 2004; Zimmerman, Forlizzi, & Evenson, 2007).

## **DATA VISUALIZATION AND PUBLIC DATA**

Data visualization is a critical tool in making complex information accessible and understandable to a broader audience. As noted by Tufte (1997), effective visualization helps users analyse and reason about data and information graphically. According to Fry (2008), visualization serves as a translation between the abstract and complex data into a form that human senses can understand and interact with, thereby supporting decision-making processes.

The role of interactive data visualization is increasingly emphasized in literature, particularly in how it facilitates deeper engagement with data.

Shneiderman's (1996) visualization mantra —“overview first, zoom and filter, then details-on-demand”— highlights the importance of providing users with intuitive navigation and interaction mechanisms to explore complex datasets efficiently.

In this context, cognitive ergonomics plays a significant role in the design of data visualizations by focusing on the human factors that influence information processing. As outlined by Norman (2004), the cognitive and emotional responses of users to visualization interfaces can significantly impact their satisfaction and overall experience. Ergonomic principles ensure that visualizations are not only functional but also address the psychological and cognitive needs of users, promoting ease of use and reducing cognitive load (Wickens et al., 2004).

As such, the usability of data visualizations is directly linked to user satisfaction and effectiveness in data interpretation and, therefore, to its UX. As suggested by Nielsen (1993), usability encompasses attributes such as learnability, efficiency, memorability, errors, and satisfaction. High usability in data visualization ensures that users can interact with data without undue frustration and with increased productivity, which is crucial in environments like public energy data where decisions based on these visualizations can have extensive implications when used by public managers.

Data visualizations in the public sector, therefore, must not only be accurate but also designed with the lay public in mind, ensuring that information is presented clearly and can be interpreted correctly without specialized knowledge.

## **METHODOLOGY**

This study employs a mixed-methods approach, integrating both quantitative and qualitative methodologies to investigate the impact of data visualization on user experience, specifically in the context of interacting with public electrical energy data (Lentez, 2020). This approach is necessary for exploring the multifaceted nature of data visualization's influence on user comprehension, engagement, and satisfaction (Creswell & Creswell, 2017). The methodology encompasses an extensive literature review, usability testing, and the extraction of qualitative insights through interviews and focus groups.

An extensive literature review constituted the initial phase of the research, aiming to identify existing knowledge gaps in the domain of data visualization and its effects on user experience. The review focused on scholarly articles, books, and case studies pertaining to data visualization, user experience design, and the specific context of public electrical energy data.

The empirical study involved a purposive sample of 30 participants, selected to encompass a diverse range of experiences with data visualization tools. This sample included 15 participants identified as data visualization specialists, possessing expertise in the field, and 15 non-specialists, representing the lay audience likely to interact with public electrical energy data visualizations. The inclusion of both specialist and non-specialist

participants allowed for a comprehensive exploration of data visualization strategies from multiple perspectives (Palinkas et al., 2015).

Usability testing sessions were conducted to quantitatively assess the effectiveness of various data visualization components, including filters, information hierarchies, graphical elements, and data diversity. Participants interacted with interactive dashboards displaying public energy datasets, and their performance was measured using task completion times, error rates, and subjective usability scores obtained from questionnaires (Lentez, 2020). This method allowed for the objective evaluation of visualization techniques in facilitating user comprehension and navigation of complex datasets (Nielsen, 1993).

To complement the quantitative findings from usability testing, qualitative data were collected through semi-structured interviews and focus groups (Morgan, 1997). These discussions aimed to capture participants' preferences, challenges encountered during the usability tests, and suggestions for enhancing the data visualization dashboards. The qualitative insights provided depth and context to the quantitative results, offering a holistic understanding of the user experience with public energy data visualizations.

Quantitative data from usability tests were analysed statistically to identify significant patterns and correlations between visualization techniques and UX metrics. Qualitative data from interviews and focus groups were subjected to thematic analysis, enabling the identification of recurring themes related to user preferences, challenges, and recommendations for data visualization improvements. This dual analysis approach ensured a comprehensive examination of the study's research questions (Braun & Clarke, 2006).

## RESULTS

The usability testing conducted as part of this research offers a deep insight into how data visualization strategies influence user experience (UX) when interacting with public energy data. Participants included 15 data visualization specialists and 15 non-specialists, allowing for an analysis of both expert and lay user interactions with public energy data dashboards.

The quantitative findings reveal significant disparities in task success rates and interaction times between specialists and non-specialists. For specialists, the task completion success rate averaged at 80%, whereas for non-specialists it was slightly lower, indicating the complexity of the visualizations used and their varying accessibility to different user groups (Lentez, 2020).

The average task completion time for specialists was noted to be approximately 2 minutes and 30 seconds, significantly faster than the 3 minutes and 15 seconds recorded for non-specialists. This suggests that familiarity with data visualization concepts positively impacts the efficiency of data interaction.

Satisfaction, measured through post-interaction surveys, indicated a generally positive response with some notable exceptions. While specialists expressed high satisfaction rates, scoring an average of 4.2 out of 5 in

terms of ease of use and data comprehension, non-specialists reported lower satisfaction, averaging 3.5 out of 5. This highlights a gap in visualization design, underscoring the need for more intuitive interfaces that can cater effectively to a broader audience (Lentez, 2020).

Qualitative data collected from interviews and focus groups shed light on specific elements that affected user satisfaction. Participants frequently cited the presence of interactive filters and a clear hierarchy of information as crucial for enhancing their understanding and engagement with the data. Graphical elements such as dynamic charts and maps were praised for their ability to illustrate complex data in an accessible format. However, some users noted the lack of contextual information and guidance on how to interpret the data effectively.

The analysis confirmed that intuitive design features such as interactive filters, graphical diversity, and clear informational hierarchies substantially improve user satisfaction and engagement.

The performance metrics (task completion time and error rates) correlated strongly with user satisfaction, indicating that more efficient interfaces lead to higher user contentment and better overall UX.

These results demonstrate the significant role of tailored visualization strategies in enhancing the public's ability to interact with and understand complex public energy data. The findings also highlight the necessity for design adjustments to accommodate non-specialist audiences, ensuring that public data visualizations are inclusive and effectively serve their educational and informational purpose.

## **DISCUSSION**

The investigation into the influence of data visualization on user satisfaction, particularly within the public electric sector, reinforce the important role of user-centered design principles. The mixed-methods approach, combining extensive literature review and empirical usability testing, facilitated a comprehensive understanding of how various data visualization components — filters, information hierarchies, graphical elements, and data diversity — affect user experience (UX).

### **Influence of Visualization Components on UX**

The study's quantitative findings reveal a significant correlation between the intuitiveness of visualization strategies and UX metrics improvement, including engagement, comprehension, and satisfaction. This aligns with previous research emphasizing the importance of interactive and intuitive design in data visualization for enhancing user interaction with complex datasets (Nielsen, 1993; Shneiderman, 1996). Notably, interactive dashboards, dynamic charts, and graphical summaries significantly outperformed traditional tabular presentations, echoing the authors such as Cairo (2012) and Yau (2013) regarding the efficacy of visually compelling data presentations in boosting user engagement and comprehension.

Qualitative insights from interviews and focus groups further corroborated these findings, highlighting user preferences for visual cues and interactive exploration capabilities that cater to individual needs and preferences.

### **Implications for Public Discourse on Energy Issues**

The findings reinforce the role of sophisticated data visualization techniques in elevating public discourse on energy issues. By demystifying complex datasets related to energy generation, distribution, and consumption, well-designed visualization, such as dashboards, can significantly enhance public engagement in discussions about energy sustainability. This is particularly relevant in the context of growing environmental concerns and the urgent need for informed public participation in sustainable energy initiatives.

### **Eleven Guidelines for User-Centric Data Visualization Development**

This research has identified several key factors that impact user satisfaction in the interaction with public electrical energy data visualizations. Based on the empirical usability testing and comprehensive analysis, eleven targeted guidelines (Lentez, 2020) are proposed for the development of future user-centric data visualizations in the public energy sector. These guidelines aim to facilitate the design of intuitive, engaging, and informative visualizations that cater to a diverse user base, including both specialists in data visualization and the lay audience.

1. **Incorporate Interactive Filters:** Allow users to customize their data exploration experience through interactive filters. This empowers users to focus on data that are relevant to their specific interests or concerns, enhancing their overall engagement and satisfaction (Nielsen, 1993).
2. **Implement Clear Information Hierarchies:** Design visualizations with clear information hierarchies to guide users effortlessly through complex datasets. A logical structure aids in the intuitive understanding and navigation of the information presented.
3. **Utilize Diverse Graphical Elements:** Employ a variety of graphical elements (e.g., charts, graphs, maps) to represent different data types and relationships. This diversity in presentation helps cater to different user preferences and enhances the interpretability of data (Tufte, 1997).
4. **Ensure Data Diversity:** Present a wide range of data to offer comprehensive insights into the electrical energy sector. This includes historical data, real-time statistics, and projections, catering to both casual users and experts seeking depth.
5. **Optimize for Usability:** Adhere to best practices in usability to ensure that the visualization is easy to use, learn, and navigate. This involves minimizing the cognitive load and providing immediate, meaningful feedback to user interactions (Norman, 2004).
6. **Design for Accessibility:** Make visualizations accessible to users with diverse abilities by following accessibility guidelines. This includes considering color contrasts, providing text alternatives for graphical information, and ensuring keyboard navigability.



7. **Enable Interactive Exploration:** Facilitate interactive exploration of data, allowing users to drill down into specifics or zoom out for an overview. Interactive elements should be responsive and provide immediate feedback.
8. **Support Multiple User Pathways:** Recognize that users may have different goals and levels of expertise. Design visualizations that support multiple pathways through the data, accommodating both novice users and experts.
9. **Provide Contextual Help and Documentation:** Include contextual help and documentation within the visualization. This assists users unfamiliar with the data or the visualization tool, enhancing their learning and engagement.
10. **Incorporate User Feedback in Design Iterations:** Engage with your user base to gather feedback on the visualization. User insights should inform ongoing design iterations, ensuring the tool evolves in alignment with user needs and preferences.
11. **Evaluate and Refine Based on User Testing:** Conduct regular user testing sessions to evaluate the effectiveness of the visualization tool. Use insights from these evaluations to refine and improve the design and functionality of the tool.

## CONCLUSION

This study explored how data visualization impacts user experience (UX) in the interaction with public energy data. The mixed-methods research approach, encompassing an extensive literature review and empirical usability testing with thirty participants, has highlighted the role of intuitive visualization techniques in enhancing UX metrics such as engagement, comprehension, and satisfaction.

The quantitative analysis underscored the efficacy of advanced visualization, such as interactive dashboards and dynamic charts, in facilitating user interaction with complex public energy data. These tools were shown to significantly outperform traditional tabular presentations in terms of task completion times, error rates, and subjective usability scores. The qualitative insights, drawn from interviews and focus groups, echoed the quantitative findings, highlighting the users' preference for visualization features that enable effective navigation and interpretation of data.

Key to this study's findings is the formulation of eleven targeted guidelines for developing user-centric data visualizations within the public energy sector. These guidelines emphasize the importance of features like interactive filters, clear information hierarchies, and the inclusion of graphical elements and data diversity. Such features are instrumental to facilitate users to navigate and interpret complex datasets confidently.

## FUTURE DIRECTIONS

Looking forward, this research opens several pathways for further exploration in the domain of data visualization and UX in public energy data interaction. The following future directions are proposed:

1. **Expanding User Studies:** Future studies could extend the participant base to include a wider demographic range, capturing a more diverse set of user experiences and preferences. This could help in refining the guidelines for user-centric design further.
2. **Longitudinal Studies:** Investigating the long-term effects of using advanced data visualization representations on user engagement and satisfaction would provide deeper insights into the sustained value of these visualizations.
3. **Cross-Sectoral Applications:** Exploring the application of the identified guidelines in other public sectors, such as water management or public transportation, could reveal universal principles of effective data visualization design.
4. **Technological Advances:** As technology evolves, so do the opportunities for innovative data visualization techniques. Future research could explore the integration of emerging technologies like virtual reality (VR) and augmented reality (AR) in public data visualization.
5. **Policy and Practice Implications:** Further research could examine how the findings of this study can influence policy-making and practical implementations in the public energy sector to foster informed public discourse and participation.

In conclusion, by bridging the gap in existing literature regarding the confluence of data visualization and UX in the electrical energy data sphere, this investigation provides profound insights into creating more accessible and impactful public data interfaces. The enhanced public engagement, informed decision-making, and heightened sustainability consciousness underscore the significant impact of data visualization on augmenting UX with public electrical energy data. This study contributes to the field of data visualization by offering academically and practically relevant insights, guiding future design efforts to align more closely with user needs and preferences in the context of electrical energy data.

## ACKNOWLEDGMENT

The authors are grateful for the support of UDESC, CEART and CAPES.

## REFERENCES

- Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). University of Chicago Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Cairo, A. (2012). *The functional art: An introduction to information graphics and visualization*. New Riders.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
- Few, S. (2009). *Now you see it: Simple visualization techniques for quantitative analysis*. Analytics Press.
- Fry, B. (2008). *Visualizing data*. O'Reilly Media.

- Lentez, A. (2020). *Visualizações de dados públicos do setor elétrico: Identificação de fatores e diretrizes que influenciam a satisfação do usuário* [Master's thesis, Universidade do Estado de Santa Catarina, Centro de Artes, Design e Moda]. Florianópolis.
- Meirelles, I. (2013). *Design for information: An introduction to the histories, theories, and best practices behind effective information visualizations*. Rockport Publishers.
- Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Sage Publications.
- Murray, S. (2017). *Interactive data visualization for the web: An introduction to designing with D3*. O'Reilly Media.
- Nielsen, J. (1993). *Usability engineering*. Academic Press.
- Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Books.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533–544.
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, 336–343.
- Tufte, E. R. (1997). *Visual explanations: Images and quantities, evidence and narrative*. Graphics Press.
- Yau, N. (2013). *Data points: Visualization that means something*. Wiley.
- Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 493–502.

# The Probable Impact of Social Media on Your Brain

Kweon Sang Hee<sup>1</sup>, Kang Bo Young<sup>2</sup>, and Ryou Jihyun<sup>3</sup>

<sup>1</sup>Department of Media & Communication, SKKU 0339, Korea

<sup>2</sup>Entertainment Management, Dong-Ah Institute of Media & Arts, Korea

<sup>3</sup>YOU Creative Communications, Korea

## ABSTRACT

This study explores to test the differences in brain waves during visual (highly liked) social media networking image fusion acceptance or viewing of social media (Youtube, Blog, and Instagram) and low liked SNS (“low liked”). The study follows a 2(low liked SNS and highly liked linked) X 3(genre-Youtube, blog, and Instagram) research design of the brain wave responses. Brain waves were measured using EEG responses by recording alpha ( $\alpha$ ) waves (8-12.99Hz) and beta ( $\beta$ ) waves (13-29.99Hz). Different parts of the brain (frontal lobe, temporal lobe, occipital lobe) were also measured in order to compare the response difference to stimulus. The experimental study was based on a statistical analysis on the EEG responses obtained through a total of 60 subjects. The brain wave difference between the low liked SNS and high liked social media were measured first. Then, responses were measured by 2X3 experimental design in order to measure the difference in brain waves according to the SNS type (Youtube, Blog, Instagram). The subjects’ brain waves were measured after viewing low liked SNS and high liked social media. Social media contents are similar message fall into the following categories: Youtube, Blog, Instagram.

**Keywords:** Exemplary paper, Brain wave, Social brain, Social modelling language

## INTRODUCTION

The social brain refers to a network of brain regions that are involved in various social cognitive processes, such as empathy and theory of mind. These processes allow us to understand and interact with others, infer their mental states, and navigate social situations. Two key aspects of the social brain are co-activation and connectivity among different networks.

Neuroscientists created an Instagram-style app to investigate the effects of social media on our brains. Through fMRI scans, they observed adolescents scrolling through photos on the app and identified the activated regions of their brains. The researchers manipulated variables such as the number of likes and the content of the photos (risky or neutral behaviours). These findings, which intrigued and concerned the scientist as a parent of a six-year-old, were subsequently replicated with young adults and in scenarios involving giving and receiving likes.

## Review of Literature

The ‘social brain’ is the network of brain regions that are involved in understanding other people, and includes the medial prefrontal cortex (mPFC) and the posterior superior temporal sulcus (pSTS). These regions are key to the process of mentalizing — that is, the attribution of mental states to oneself and to other people

The structure of the social network is correlated with activity in the amygdala, which links decoding and interpreting social signals and social values. The structure also relies on the mentalizing network, which is central to an individual’s ability to infer the mental states of others.

Seeing photographs with more likes was associated with increased activity in brain regions responsible for social cognition, rewards (the dopamine system), and attention (the visual cortex). When participants viewed photos with more likes, their overall brain activity and the activation of the visual cortex were heightened. This suggests that we pay more attention to and focus on images that receive more likes, examining them in greater detail.

To ensure that image differences did not influence the results, the researchers randomly assigned the number of likes across images and controlled for factors such as luminosity and content. The findings held true whether participants were looking at their own photos or others’ photos. This implies that when we see social media images with higher like counts, we tend to scrutinize and engage with them more intensively.

The activation of the mentalizing network and regions associated with social skills was observed when individuals viewed their own photos with more (randomly assigned) likes. Our brains process self-images within a social context, considering how others perceive us and our connections with them. Furthermore, receiving more likes on one’s own photos activates the dopamine reward system, which plays a role in pleasure, motivation, and Pavlovian responses.

Having more likes on one’s own photos activates the social brain and stimulates the mentalizing network. When individuals see photos of themselves with a higher number of likes, their brain activity increases in regions related to social skills and imitation. This suggests that when we view our own photos, our brains engage in thinking about how others perceive us and consider the social context surrounding the photos.

## RESEARCH QUESTION

In this study, to learn about the differences for the human recognition between few social network and large social network, we will proceed with the following topic of social media as social brain activities. In practice, compared to low liked SNS, the social brain activities are less than high liked SNS.

Social interactions have played a pivotal role in the evolution of human brains and behaviour. This literature review investigates the neurobiological basis of social connections, delving into the evolutionary advantages of social engagement, the neural mechanisms governing social behaviours, and the contemporary implications of social media on brain responses.

**Evolutionary Significance of Social Interaction:** Loneliness, often perceived negatively, could have an evolutionary benefit. Cacioppo and Patrick (2008) propose that loneliness persists due to its role in motivating individuals to establish, repair, and maintain social relationships. This theory is supported by research on primates, where larger brains were favoured in more social species (Dunbar, 1998), and on humans, where neocortex size correlated with social group size (Lewis et al., 2011).

**Neural Mechanisms Underlying Social Behaviours:** The human brain's response to social interactions involves the dopamine reward system. Eisenberger, Lieberman, and Williams (2003) found that social exclusion activated brain regions associated with physical pain, while the ventral striatum played a key role in romantic love (Aron et al., 2005), cooperation (Rilling et al., 2002), social comparisons (Fliessbach et al., 2007), and altruism (De Quervain et al., 2004).

**Impact of social media on Brain Responses:** The advent of social media has transformed the landscape of social interactions. Sherman et al. (2016) studied brain responses to receiving likes on social media, demonstrating its rewarding nature. Subsequent studies extended these findings to young adults (Sherman et al., 2018) and revealed brain activation during giving and receiving likes (Sherman et al., 2018). However, concerns about potential negative effects have been raised (Meshi et al., 2013; Allen, 2017).

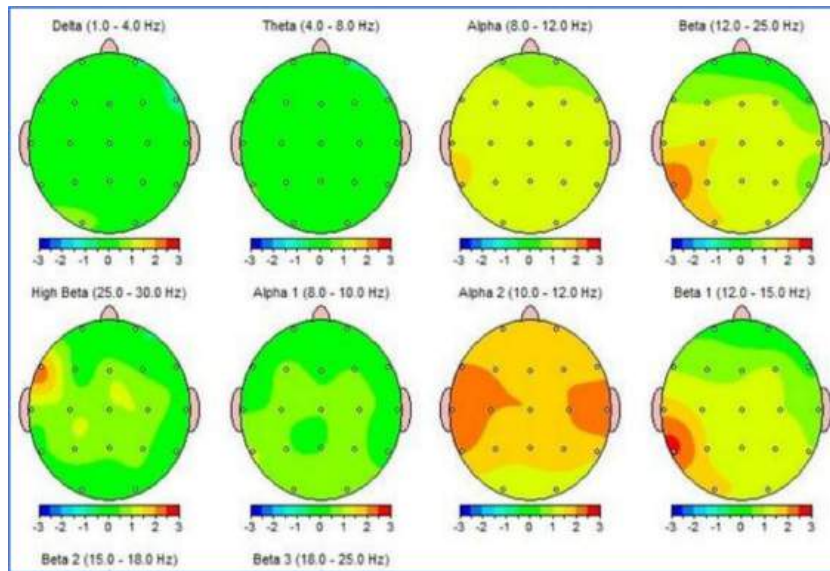
The studies reviewed underscore the fundamental importance of social connections in shaping the human brain's structure and function. Loneliness serves as an evolutionary motivator, neural mechanisms underlie social behaviours, and contemporary technology like social media influences brain responses. This review highlights the dynamic interplay between social engagement and brain processes, offering insights into the intricate relationship between human biology and social interactions.

- Research Question 1:  
How do  $\alpha$ -waves and  $\beta$ -waves differ between high-liked vs. low liked SNS and social YouTube social media (high-liked vs. low liked)?
- Research Question 2  
Will different wavelengths in two channels appear for between high liked vs, low liked SNS and social YouTube social media (high-liked vs. low liked)?

## RESEARCH METHOD

In general, brain waves are divided into delta ( $\delta$ ) waves (0.2-3.99 Hz), theta ( $\theta$ ) waves (4-7.99 Hz), alpha ( $\alpha$ ) waves (8-12.99 Hz), beta ( $\beta$ ) waves (13-29.99 Hz), and gamma ( $\zeta$ ) waves (30-50 Hz) based on the range of frequencies at which they oscillate. While electroencephalography can measure all frequencies of brain waves, only the  $\alpha$ - and  $\beta$ -waves, which are the most interpretive, are extracted and used in brainwave research.

Alpha waves are dominant in relaxed states, such as relaxation. Their amplitude increases as you become more stable and relaxed. They are generally regular, steady waves that are largest in the parietal and occipital lobes and smallest in the frontal lobe.



**Figure 1:** Brainwave measurement.

Beta waves are predominant in the frontal lobe and are present during all conscious activity, such as waking and talking. They are often seen when we are anxious, engaged in complex thinking or calculations, and when we are nervous.

The cerebral cortex beyond the surface of the head is divided into the frontal lobe, parietal lobe, temporal lobe, and occipital lobe, and each part plays a different role. The occipital lobe, located at the back of the head, contains the primary visual cortex, which is responsible for processing primary visual information, and the parietal lobe, located near the crown, contains the somatosensory cortex, which is responsible for processing motor/sensory information.

**Table 1.** Related brain waves and functions of each channel of the equipment.

Right	CH 1	CH 2	CH 3	CH 4	REF
	RED Prefrontal ( $\beta$ ) Complex thinking	YELLOW Frontal lobe ( $\beta$ ) Auditory	Green Temporal Lobe Visual + emotional thinking	Blue Occipital Lobe ( $\alpha$ ) Grounding	(Red) Earlobe
Left	CH 5	CH 6	CH 7	CH 8	GND
	Orange Prefrontal ( $\beta$ ) Complex thinking	Purple Frontal lobe ( $\beta$ ) Auditory	Gray Temporal Lobe Visual + emotional thinking	White Occipital Lobe ( $\alpha$ ) Grounding	Black Back of hand

When different types of social media (high-liked vs. low liked) are shown to subjects, wavelength changes by channel can be used to identify the structure of the brain and its developmental functions. In particular,  $\alpha$  and  $\beta$  waves have the characteristic of establishing an inverse relationship. If the  $\alpha$  wave is high after watching a social media (high-liked vs. low liked), the  $\beta$  wave is likely to be relatively low, and in this case, it can be expected that the visuals are intensively stimulated and emotional thinking becomes active. On the other hand, if the  $\beta$  wave is high and the  $\alpha$  wave is relatively low after watching a social media (high-liked vs. low liked), it stimulates the entire frontal lobe from the temporal lobe to the frontal lobe, which stimulates the auditory system and leads to planning and complex thinking rather than emotion.

## RESULT OF THE RESEARCH

First, we can analyse the results of research question 1. To analyse the results, we used paired t-test to compare the results. In other words, based on the results, we can compare the difference between  $\alpha$  and  $\beta$  waves between flat and stereoscopic images.

**Table 2.** Comparison of average brainwaves of high-liking social media and low-liking social sites.

Class		Low-Liked	High-liked	t	Sig.
$\alpha$ Wave	YouTube	17.61	10.34	4.01	.017
	Blog	16.99	10.91	2.54	.036
	Instagram	39.54	12.43	5.23	.003
$\beta$ Wave	YouTube	23.92	80.55	-8.39	.000
	Blog	26.40	76.10	-4.12	.009
	Instagram	35.04	58.90	-3.07	.035

In the above experiment, the waveforms of a flat image and a stereoscopic image are recorded. Overall, we can see that alpha waves are higher in flat images, and beta waves are higher in stereoscopic images. In the case of alpha waves, sports, animation, and promotional social media (high-liked vs. low liked) all have t-values greater than 0, and the probability of significance is lower than .05. Therefore, we can see that low-liked SNS are higher than high-liked SNS in alpha waves. This can be interpreted as low social brain of low liked SNS images. In the beta (B) wave, Youtube, blog, and Instagram all have negative t values, and the probability of significance is lower than .05. In other words, in the beta B wave, high-liked SNS have higher values than low-liked SNS. Overall, high liked SNS were found to be more stimulating.



**Table 3.** Analysis of the absolute value of wave A for low and high likes on social sites.

		Type	Low Liked	High-Liked	t	Sig
$\alpha$ Wave	YouTube	Ch 1	24.66	13.39	6.534	.000
		Ch 2	26.12	13.63	6.127	.000
		Ch 3	12.94	11.57	5.418	.000
		Ch 4	13.19	11.26	4.583	.001
		Ch 5	13.27	9.33	2.196	.056
		Ch 6	12.89	10.82	4.719	.001
		Ch 7	12.77	10.62	3.964	.003
		Ch 8	15.40	12.15	3.482	.007
	Blog	Ch 1	27.94	13.36	3.686	.005
		Ch 2	28.74	13.58	3.906	.004
		Ch 3	12.82	10.46	6.315	.000
		Ch 4	13.39	10.37	5.953	.000
		Ch 5	7.682	9.98	4.469	.002
		Ch 6	11.91	10.57	3.827	.004
		Ch 7	11.73	9.56	4.566	.001
		Ch 8	23.87	10.13	2.094	.066
	Instagram	Ch 1	45.65	12.69	1.879	.093
		Ch 2	52.23	13.06	1.697	.124
		Ch 3	19.37	10.20	2.509	.033
		Ch 4	25.06	10.24	2.012	.075
		Ch 5	53.67	8.65	1.121	.291
		Ch 6	83.19	9.74	2.144	.023
		Ch 7	18.15	9.54	5.064	.061
		Ch 8	48.90	10.12	3.144	.001
$\beta$ Wave	YouTube	Ch 1	35.17	81.78	-2.401	.040
		Ch 2	33.16	85.05	-2.533	.032
		Ch 3	17.62	59.65	-1.674	.128
		Ch 4	14.18	49.75	-1.862	.096
		Ch 5	39.37	121.73	-1.488	.171
		Ch 6	25.66	79.13	-2.173	.058
		Ch 7	11.46	41.40	-1.621	.140
		Ch 8	14.74	54.57	-2.019	.074
	Blog	Ch 1	42.57	77.02	-1.273	.235
		Ch 2	39.78	83.35	-1.110	.296
		Ch 3	20.20	40.47	-2.393	.040
		Ch 4	16.10	30.82	-3.838	.004
		Ch 5	28.95	134.80	-3.168	.011
		Ch 6	22.10	92.75	-2.673	.026
		Ch 7	15.65	25.32	-9.242	.000
		Ch 8	25.77	27.87	-13.067	.000
	Instagram	Ch 1	43.67	88.60	-2.921	.017
		Ch 2	46.55	92.54	-2.098	.065
		Ch 3	24.23	40.73	-4.228	.002
		Ch 4	22.11	30.47	-5.408	.000
		Ch 5	45.27	99.91	-2.934	.017
		Ch 6	54.43	51.68	-5.671	.000
		Ch 7	16.21	24.16	-6.311	.000
		Ch 8	27.83	26.62	-8.723	.000

## ACKNOWLEDGMENT

The authors would like to acknowledge for all research comments.

## REFERENCES

- Allen, M. (2017, November 9). Sean Parker Unloads on Facebook: ‘God Only Knows What It’s Doing to Our Children’s Brains.’ *Axios*. Retrieved from <https://www.axios.com/sean-parker-facebook-children-brains-8af0c0ab-8c4b-46eb-8a37-e237a8c5f3b1.html>.
- Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*, *94*(1), 327–337.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the Autistic Child Have a ‘Theory of Mind’?. *Cognition*, *21*(1), 37–46.
- Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *Neuroimage*, *49*(3), 2687–2696.
- Cacioppo, J. T., & Patrick, W. (2008). *Loneliness: Human Nature and the Need for Social Connection*. New York: W. W. Norton.
- Cacioppo, J. T., Cacioppo, S., & Boomsma, D. I. (2014). Evolutionary mechanisms for loneliness. *Cognition and Emotion*, *28*(1), 3–21.
- Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, AL: Harvester Press.
- De Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*(5), 178–190.
- Dunbar, R. I. M. (2011). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*(6), 469–493.
- Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, *8*(7), 294–300.
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*(5643), 290–292.
- Falk, E. B., O’Donnell, M. B., Tompson, S., Gonzalez, R., Dal Cin, S., Weber, R.,... & Strecher, V. (2015). Functional brain imaging predicts public health campaign success. *Social Cognitive and Affective Neuroscience*, *10*(3), 487–493.
- Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*(5854), 1305–1308.
- Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science*, *153*(3735), 501–506.
- Lieberman, M. D. (2013). *Social: Why Our Brains Are Wired to Connect*. New York: Oxford University Press.
- Maliske, L., Schurz, M., & Kanske, P. (2021). Interactions within the social brain: Co-activation and connectivity among networks enabling empathy and theory of mind. *PsyArXiv Preprints*. <https://doi.org/10.31234/osf.io/52z4f>
- Matthews, G. A., Nieh, E. H., Vander Weele, C. M., Halbert, S. A., Pradhan, R. V., Yosafat, A. S.,... & Tye, K. M. (2016). Dorsal raphe dopamine neurons represent the experience of social isolation. *Cell*, *164*(4), 617–631.

- Nation, D. A., Gonzales, J. A., Mendez, A. J., Zaias, J., Szeto, A., Brooks, L. G.,... & Schneiderman, N. (2008). The effect of social environment on markers of vascular oxidative stress and inflammation in the Watanabe heritable hyperlipidemic rabbit. *Psychosomatic Medicine*, 70(3), 269–275.
- Nonogaki, K., Nozue, K., & Oka, Y. (2007). Social isolation affects the development of obesity and type 2 diabetes in mice. *Endocrinology*, 148(10), 4658–4666.
- Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, 47(6), 419.
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans. & Ed.). Oxford: Oxford University Press.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, 35(2), 395–405.
- Ruan, H., & Wu, C.-F. (2008). Social interaction-mediated lifespan extension of *Drosophila* Cu/Zn superoxide dismutase mutants. *Proceedings of the National Academy of Sciences*, 105(21), 7506–7510.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews*, 42, 9–34.
- Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., et al. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, 147(1), 293–327.
- Sherman, L. E., Greenfield, P. M., Hernandez, L. M., & Dapretto, M. (2018). Peer influence via Instagram: Effects on brain and behavior in adolescence and young adulthood. *Child Development*, 89(1), 37–47.
- Sherman, L. E., Konrath, S. H., Raby, K. L., & Smalley, D. (2018). What the brain ‘likes’: Neural correlates of providing feedback on social media. *Social Cognitive and Affective Neuroscience*, 13(7), 699–707.
- Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The power of the like in adolescence: Effects of peer influence on neural and behavioral responses to social media. *Psychological Science*, 27(7), 1027–1035.
- Stranahan, A. M., Khalil, D., & Gould, E. (2006). Social isolation delays the positive effects of running on adult neurogenesis. *Nature Neuroscience*, 9(4), 526–533.

# Automated Visualization for Visual Analytics: Trends, Challenges, and Opportunities

Vinay Kumar Uppalapati and Bo Sun

Department of Computer Science Rowan University, Glassboro, NJ 08028, USA

## ABSTRACT

Visualization, as a major approach of visual analytics, involves many human interaction techniques, especially in terms of how individuals communicate, comprehend, and interpret information. Creating visualizations is a tedious process and requires skill, but automatic data visualization technologies have made it easier to create visualizations. They completely changed the landscape of data analysis and decision-making processes. As the demand for effective and efficient visualization solutions grows across diverse sectors, researchers and practitioners have developed a plethora of autonomous systems aimed at transforming raw data into meaningful visual representations. This paper investigates the methodologies utilized by these systems, categorizing them based on machine learning approaches combined with various data inputs, template-based approach, and other technique/algorithm-based approach. We collected 31 top-tier journal papers in the field and shed light on the diverse techniques employed in generating visualizations automatically, enhancing our understanding of their capabilities, compatibility, and usability across various contexts. Our survey aims to provide insights into the strengths, limitations, and potential areas for future exploration in automatic data visualization, offering guidance to practitioners, researchers, and developers in selecting appropriate techniques for their specific needs and datasets. By systematically examining these systems and pinpointing areas for improvement, we contribute to the advancement and refinement of automatic data visualization methodologies, fostering progress in this dynamically evolving domain.

**Keywords:** Automation, Infographic, Data visualization, Dashboard, Machine learning, Artificial intelligence, Natural language

## INTRODUCTION

Data visualization is a critical tool in modern data analysis and decision-making processes. Data visualizations help humans understand abstract data. As the volume and complexity of data increases significantly across different sectors, such as business, healthcare, finance, and scientific research, so does the demand for effective and efficient data visualization solutions. Despite the number of visualization tools and libraries, users usually struggle to select the appropriate charts, appropriately encode data, and create visually appealing images. Users must understand the data and be proficient in using the available tools to do the activities listed above.

Creating visualizations manually requires a significant amount of effort and time. This problem can be overcome by automation; hence there is a high need for systems that can visualize data without user intervention. In response to this demand, researchers and practitioners have created a wide range of autonomous data visualization systems that turn raw data into meaningful visual representations. Automatic data visualization has progressed dramatically over time, reflecting technological developments and an increasing demand for accessible data insights. Initially, data visualization necessitated human labor, with analysts making charts and graphs using special tools. However, as computer power and machine learning techniques improved, automated data visualization tools arose. These programs automatically analyze datasets, discover patterns, and provide visual representations, saving users time and effort. From simple bar charts, as seen in (Dibia & Demiralp, 2019), (Luo et al., 2018) and (Harper & Agrawala, 2017) to complex interactive dashboards, as seen in (Deng et al., 2022), (Wu, Wang, Zhou, et al., 2021) automated data visualization solutions today offer a range of capabilities to cater to varying data analysis demands, helping users across various industries to make informed decisions based on data-driven insights.

Understanding automatic data visualization systems requires an examination of the methodologies they utilize to create visualizations. This survey classified existing automatic data visualization systems into three primary categories based on their methodology approaches: Systems adopted machine learning such as (Dibia & Demiralp, 2019) and (Luo et al., 2018) systems that are template-based such as the ones proposed by (Harper & Agrawala, 2017), and algorithmic or technique-driven systems as seen in (Li et al., 2021). Each approach offers distinct advantages and faces unique challenges, accommodating various data traits, user needs, and application contexts.

In this survey paper, we explore the field of automatic data visualization systems to offer a comprehensive overview of the techniques they utilize. We aim to categorize these systems according to their underlying methodologies, shedding light on the various approaches employed in generating data visualizations automatically.

What sets our survey paper apart is its emphasis on a novel classification scheme that categorizes automatic data visualization systems based on their underlying methodologies. This refined classification enhances our understanding of automatic data visualization methods' capabilities, compatibility, and usability across various data sources and user interaction contexts. The objective of this survey paper is to categorize and analyze existing automatic data visualization systems based on their underlying methodologies. We also aim to pinpoint the areas the existing systems are lacking. By doing so, we provide insights into the strengths and limitations of different approaches, offering guidance for researchers, practitioners, and developers in selecting appropriate automatic data visualization techniques for their specific needs and datasets. Additionally, by pinpointing gaps and potential areas for future exploration, our goal is to encourage innovation and progress in this dynamically evolving domain. Through a

systematic examination of these systems, we contribute to the advancement and refinement of automatic data visualization methodologies, fostering innovation and improved usability in data visualization tools and systems.

In Related Work, we summarize related work in the area. In the section Survey Methodology, we detail our approach to this survey paper. Then we review the top-tier papers in the field in the section Automatic Visualization Systems and discuss challenges that current automatic data visualization systems face, and opportunities derived from our survey in the challenges and opportunities. The next section is about the discussion and limitations of our survey, and finally, our last section is the conclusion.

## RELATED WORK

Automated visualization systems are a relatively new area of research. There is a small dedicated group of researchers from visualization and graphics. The top-tier papers in this area are limited compared with other Human-Computer Interaction (HCI) research.

There are a few surveys in this field. Wu, Wang, Shu, et al. (2021) explored the concept of formalizing visualizations as a data format and examining recent advancements in applying artificial intelligence (AI) techniques to visualization data, referred to as AI4VIS. The focus is on understanding and analyzing the digital representations of visualizations stored in computers, with an emphasis on data visualization, such as charts and infographics. The paper outlines a series of typical tasks researchers utilize for visualizing data and provides an in-depth examination of AI methodologies designed to fulfil these tasks. Zhu et al. (2020) reviewed and categorized automatic tools and academic papers in their survey, which focused on providing recommendations for visualizing data in various applications. These recommendations are classified into different categories based on the type of visualization they pertain to, such as network graph visualizations, annotation visualizations, and storytelling visualizations. Wang (Wang et al., 2021) surveyed papers of machine learning techniques applied to visualizations with the agenda of answering “what visualization processes can be assisted by ML?” and “how ML techniques can be used to solve visualization problems?”.

In contrast to previous survey papers, which primarily focus on summarizing existing automatic visualization generation systems using AI, our review expanded the scope to include template-based approach and other non-AI based algorithm approach. Our objective is to provide a comprehensive overview of these systems while also identifying the research problems and gaps they face.

## SURVEY METHODOLOGY

To provide an understanding of the current Automatic Visualization studies. We have conducted an analysis of 31 related papers in the fields of Visualization, Machine Learning, and Artificial Intelligence. Each paper is classified based on the approach used by the system or model in the paper to

generate the visualization. Manual analysis is done for the collected papers to extract thorough details.

## Survey Scope

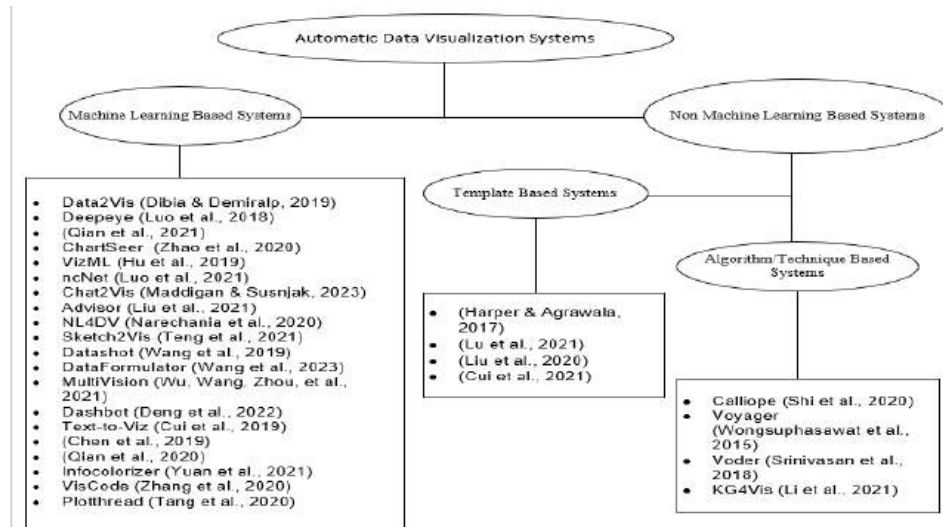
Our analysis commenced with the assembly of a comprehensive corpus of automatic visualization papers. We meticulously selected research articles published between January 2013 and December 2023 from prestigious journals and conferences associated with IEEE and ACM (see Figure 1). Leveraging Google Scholar's top results, we focused on the past decade to capture the latest advancements in automatic visualization generation research.

To expedite the process without compromising accuracy, we adopted a method akin to previous studies by (Wang et al., 2021), primarily scrutinizing paper titles and abstracts. We reserved a detailed examination of the full manuscript for instances where title and abstract information were insufficient. We have even collected research papers related to the field of visualization if the abstract is talking about visualization generation. Throughout our search, we were attentive to keywords like "Automation," "Infographic," "Visualization," "ML," "AI," and "Natural Language" to pinpoint relevant papers.

Publisher	Conference/Journal	Year of publication	Research
IEEE	Transactions on Visualization and Computer Graphics	2023	Wang et al.
		2022	Deng et al.
		2021	Cui et al., Li et al., Luo et al., Wang et al., Wu et al., Wu et al., Yuan et al.
		2020	Narechania et al., Qian et al., Shi et al., Tang et al., Zhang et al., Zhao et al.
		2019	Cui et al., Wang et al., Chen et al.
		2018	Srinivasan et al.
		2017	Harper et al.
		2015	Wongsuphasawat et al.
		2021	Lu et al., Liu et al.
		2020	Srinivasan et al.
	Pacific Visualization Symposium	2023	Maddigan et al.
		2019	Dibia et al.
		2018	Luo et al.
		2021	Teng et al.
ACM	ACM SIGKDD Conference on Knowledge Discovery & Data Mining	2021	Qian et al.
	CHI Conference on Human Factors in Computing Systems	2019	Hu et al.
Elsevier	Visual Informatics	2020	Zhu et al.

**Figure 1:** Diagram of a table showing the number of papers per conference/journal.

After following the process and filtering many papers, we included 31 top-tier papers for this review. Since the survey aims to understand the current studies in the field of automatic visualization, we focused on papers that mostly contribute to automating the process of visualization generation and papers related to automating aspects related to visualizations such as generating captions etc. Theory and evaluation papers were excluded. Then we validated the papers based on whether the paper contributes to the field in generating the visualizations automatically or automating aspects related to visualizations without the user manually taking the action.



**Figure 2:** Diagram showing the classification of automatic data visualization systems in the study.

## Paper Classification

Our classification framework revolves around the approach adopted by each paper to generate visualizations. Papers are initially categorized into two main types: ML-based and non-ML-based. Non-ML-based systems are further classified into two sub-categories: Template-Based and Algorithm/technique-based (see Figure 2). This classification schema offers valuable insights into the diverse approaches employed in automatic visualization generation, highlighting their respective strengths and limitations.

Throughout our validation process, we meticulously ensured that each paper included in our analysis indeed contributed in a way to the automatic generation of visualizations or automation of aspects related to visualizations without manual intervention. This rigorous validation step underscores the reliability and credibility of the papers included in our survey. In summary, our methodology encompasses a systematic and thorough approach to analyzing the current landscape of automatic visualization studies, providing a robust foundation for our survey's findings and conclusions.

## AUTOMATIC VISUALIZATION SYSTEMS

### ML-Based Systems

ML-supported visualization generation systems reduce manual work by automatically generating visualizations for the input data using trained ML models.

Data2Vis is an end-to-end trainable neural translation model designed to automatically generate visualizations from given datasets. It employs Vega-Lite, a declarative language, to translate data specifications into



visualization specifications. Utilizing a multilayered attention-based encoder-decoder network with LSTM units, Data2Vis streamlines the process of visualization generation (Dibia & Demiralp, 2019).

Evaluating and ranking visualizations is critical to guaranteeing the effectiveness and relevance of generated visual representations. To evaluate and prioritize visualizations, dedicated systems use machine learning techniques. Deepeye is a visualization system that solves three following problems: It recognizes whether a visualization is good or bad, it ranks which visualization is better between any 2 given two visualizations, and it finds the top k-visualizations for a given data set. It solves recognition issues by training a binary classifier to decide whether a visualization is good or bad. It solves visualization ranking issues in two ways either by using a supervised learning-to-rank model or relying on experts' knowledge to specify partial orders as rules. Then, the graph-based approach and rule-based optimization efficiently compute top-k visualizations by filtering bad visualizations (Luo et al., 2018).

Similar to Luo's approach (Luo et al., 2018), Qian (Qian et al., 2021) developed a machine-learning system to suggest visualizations for any given unseen new dataset. It automatically generates visualizations, assigns scores based on their effectiveness, and provides the user with a list of recommended visualizations in order of usefulness.

Assisting analysts with visual analysis involves automating elements of data exploration and interpretation. Machine learning systems make it easier for analysts to extract insights from complex datasets. Chartseer is a system that uses machine intelligence to assist analysts in visually monitoring their exploratory visual analysis (EVA). ChartSeer employs deep learning to analyze charts created by analysts, providing visual summaries, and suggesting relevant charts for further exploration based on user interactions. It contains a Data Storage and Analysis module, a Chart Encoder, and a Chart Decoder, which were trained using the GVAE deep learning model (Zhao et al., 2020).

Luo's work (Luo et al., 2018) emphasis on analyzing and ranking visualizations is consistent with the goals of Zhao's (Zhao et al., 2020) and Qian's (Qian et al., 2021) work, where they used the ML approach to recommend visualizations based on their effectiveness.

Design choice during visualization development is critical for guiding the automated generation process. Some systems use machine learning algorithms to assess datasets and suggest the best visualization designs. VizML is a visualization approach in which the system learns from a large collection of datasets and their corresponding visualizations. The system identifies five key design choices analysts make: at the encoding level, choosing mark types using a shared axis or not and deciding how to represent data on the X and Y-axis; at the visualization level, selecting visualization types. VizML differs from other systems by focusing on suggesting design choices. It offers easier validation, interpretable measures, and smoother integration into visualization systems. The drawback is that it only suggests how to visually represent data and doesn't provide recommendations for data queries. Information graphics are a type of data visualization that delivers

information in an engaging manner (Hu et al., 2019). Vizml's predictive approach to visualization design is similar to Qian's (Qian et al., 2021) system in suggesting visualizations based on machine learning-driven analysis.

Natural language querying enables people to engage with data and generate visualizations in a common language. Machine learning-based natural language processing systems enable users to communicate seamlessly with data visualization tools. Luo et al. (2021) present ncNet, a Transformer-based sequence-to-sequence model that supports the conversion of natural language queries on the given input dataset to visualization (NL2VIS). Additionally, it also takes in a chart template as a constraint to limit what could be visualized and then applies many visual optimizations to produce better visualization templates. Chat2VIS is a system that utilizes pre-trained large language models such as ChatGPT, GPT – 3 and Codex and converts natural language into code to generate visualizations (Maddigan & Susnjak, 2023). Advisor is an automatic system that creates visualizations with explanations to address questions from the public about tabular data. It has a pre-trained language model to convert natural language questions and table headers into vectors for a multi-task deep neural network, then identifies relevant data areas and determines the appropriate aggregation type based on these vectors. The results are presented with well-designed visualizations and annotations for different types of attributes and tasks (Liu et al., 2021). NL4DV is a toolkit that helps people use Natural language to interact with visual data for analysis. It has a query processor that infers relevant information from the input query, such as data attributes and analytic tasks. This information is then passed to a Visualization Recommendation Engine that generates a list of visualization specifications relevant to the input query (Narechania et al., 2020).

ncNet (Luo et al., 2021), Chat2Vis (Maddigan & Susnjak, 2023), Advisor (Liu et al., 2021), and NL4DV (Narechania et al., 2020) all focus on enabling natural language interaction for visualization tasks. Systems like Advisor and ncNet use explicit mechanisms to decide which visualizations to generate, while Chat2VIS addresses this gap with the help of an AI component's ability to delegate the decision-making process for chart selection (Maddigan & Susnjak, 2023).

Getting visualizations generated from hand-drawn sketches allows designers and data analysts to experiment rapidly. The Sketch2Vis system uses advanced computer techniques to automatically create computer code for making graphs and charts from hand-drawn sketches (Teng et al., 2021).

Fact sheets provide brief summaries of facts, making complex information easy to comprehend for users. Automated systems that use machine learning can simplify the process of creating fact sheets from raw data. Datashot is an automated system that creates fact sheets automatically from tabular data. uses a trained decision tree model to select the best visualization style and generate fact sheets (Wang et al., 2019).

Interactive visualization tools help users to efficiently explore and express the results of data analysis. These tools frequently use machine learning approaches to speed up the visualization generation process. DataFormulator is an interactive tool for creating visualizations. It separates the high-level

visualization ideas from the low-level data transformation steps with the help of an AI agent. The user describes the data they want to visualize in plain language. The tool's AI agent then takes care of automatically transforming the data to match these descriptions with the help of a large language model and generates the desired visualizations (Wang et al., 2023).

Dashboards give customers a full picture of data, allowing them to track trends and make informed decisions. MultiVision is a deep learning model that not only selects data columns but also recommends multiple charts to simplify the process of creating analytical dashboards for data analysis. The model is part of a mixed-initiative system, allowing users to provide input for better recommendations. It introduces passive recommendations upon request and active recommendations that adjust automatically with user changes (Wu, Wang, Zhou, et al., 2021). Dashbot is a deep reinforcement learning-based recommendation system that uses a deep neural network to analyse data and create dashboards (Deng et al., 2022).

Creating an infographic involves more than just ensuring it effectively communicates information—it also requires attention to visual aesthetics. This aspect of the process can be challenging and time-intensive for both skilled designers and individuals lacking expertise in design. As a result, there is a growing need for automated tools that can streamline the infographic design process.

Like Data2Vis (Dibia & Demiralp, 2019), Text-to-Viz is a system that uses neural networks to convert statements about simple proportion-related statistics to a set of infographics with pre-designed styles. Infographics are effective tools for presenting complex information in an eye-catching way. Text-to-Viz uses a trained CNN + CRF model to identify and extract information for infographic constructions. A framework to systematically synthesize infographics is proposed. The infographics are ranked by taking three scores, including semantic, visual, and informative, into consideration. The infographics with the best overall score values are presented to the user for any changes, and then the user can save it (Cui et al., 2019).

Chen (Chen et al., 2019) presented an automated approach for designing timeline infographics. In his approach, a deep neural network extracts global (timeline representation, scale, layout) and local (element location, category, pixels) information from a bitmap timeline. Techniques like Non-Maximum Merging, Redundancy Recovery, and DL Grab Cut are employed to create an adaptable template and then generate the visualization using the template and the input data. Qian (Qian et al., 2020) suggested a new method to automate the creation of infographics by mimicking existing examples found online. Recursive neural networks and the Monte Carlo Markov Chain (MCMC) are commonly used in this method to improve the visual appearance of the initial draft until a satisfactory result is achieved.

Infographic design platforms frequently offer a restricted range of color palette choices, which can constrain users' creative flexibility. Infocolorizer, a recommendation engine built using deep learning techniques, considers user preferences, automates the process, and tailors color palettes to the layout of elements (Yuan et al., 2021).

Enhancing visualizations entails adding additional information or increasing visual aesthetics while maintaining data integrity. Systems that enhance the visualization are as follows. VisCode is a method used to embed information into visualization images with minimal visual distortion. A Deep Neural Network model is trained on visualization images and QR Code data (Zhang et al., 2020).

Storyline visualizations show data in a narrative fashion, allowing users to follow a logical storyline while studying large datasets. Plot Thread is a tool that integrates a set of flexible interactions to support easy customization of storyline visualizations with the help of an AI agent, which also incorporates the authoring process (Tang et al., 2020).

### **Non-ML Based Systems**

Some researchers adopted templates or techniques instead of the ML approach to automate the creation of a visualization. In this section, we further classified this research into two more subcategories: template-based systems and Technique /Algorithm-based systems.

### **Template-Based Systems**

Template-supported visualization generation systems use an existing visualization as a template and generate a new visualization with the same style for a given input data. Harper (Harper & Agrawala, 2017) presented a technique for converting a basic D3 chart into a reusable style template and then generating a chart for the input data given in the style of that template. The proposed algorithm extracts components from D3 charts, and the representation aptly captures various chart structures, translatable into Vega-lite's mapping-based format.

Data-based scrollytelling has grown in popularity because of its capacity to effectively communicate data insights. Despite its success, the process can be time-consuming due to the requirement to collect, analyze, and show data in an appealing and engaging manner. Lu (Lu et al., 2021) presented an automatic method to generate expressive scroll telling visualization, which can present easy-to-understand data facts through a carefully arranged sequence of views. The proposed method uses data mining to extract data facts from the tabular data given as input. Facts are divided into two categories: presentational (i.e., value, proportion, aggregation, and rank) and computational facts (i.e., distribution, difference, trend, association, extreme, and outlier). It computes the importance of the extracted facts, and arranges them in sequence, and then uses templates to generate visualization stories.

Visualization captions could assist analysts in quickly summarizing and interpreting visualizations. AutoCaption system presented a new method for automatically creating captions for visualization charts. A multilayer perceptron classifier to identify visual marks, visual channels, and associated text information in the charts. A 1-D convolutional residual network is then used to analyze relationships between visual elements and recognize important features. Finally, a template-based approach is employed to generate full descriptions of the visual charts (Liu et al., 2020).

Infographic bar charts are extensively used to communicate numerical information, but changing or reusing them requires time-consuming and error-prone human editing. Cui (Cui et al., 2021) proposed an approach whereby, when given an infographic bar chart, visual elements are collected. Features are extracted from them, and several trained decision tree models were employed to identify corresponding information and recover data from the given chart. While users are operating the data, the algorithm detects the changes and updates the visual elements in the chart accordingly.

### **Algorithm-Based/Rule-Based Systems**

This section reviews the systems that use an algorithm or technique that is not ML-based.

Data storytelling helps people comprehend and communicate data insights. Creating engaging narratives from raw data makes complex information more understandable. Calliope is a data story generation system that creates visual data stories from an input spreadsheet through an automatic process. The system adopted a special algorithm called Monte Carlo tree search. It creates story pieces and data facts that are arranged in a logical order. The importance of the generated facts in each search step will be estimated by their information quantity, which is calculated based on the information theory and their pattern significance using auto-insight techniques (Shi et al., 2020).

Voyager is a system that aims to complement the manual construction of charts with an interactive gallery of automatically generated visualizations. Voyager employs a mixed-initiative approach, meaning it combines human input with automated assistance. It recommends charts based on statistical and perceptual measures. The system supports faceted browsing by allowing users to explore and navigate the recommended visualizations (Wongsuphasawat et al., 2015). (Srinivasan et al., 2018) presented Voder, a system that lets users interact with the automatically generated data facts in exploring both alternative visualizations and conveying a data fact presented by a set of embellishments within a visualization.

Like ML-based systems, Chartseer (Zhao et al., 2020) and MultiVision (Wu, Wang, Zhou, et al., 2021), Voyager (Wongsuphasawat et al., 2015), and Voder (Srinivasan et al., 2018) assist the users in exploring and selecting relevant visualizations.

ML-based techniques frequently function as black boxes, making it difficult to comprehend why a certain visualization is suggested, potentially restricting their widespread adoption. To fill this gap, Li (Li et al., 2021) presented KG4Vis, a Knowledge Graph (KG)-based approach for visualization recommendation. The recommendation algorithm uses embedding-based methods and the widely recognized embedding approach, TransE. TransE learns the embedding vectors of entities and relations in the KG, automatically generates meaningful rules, and then suggests suitable visualizations based on data. This approach aligns with ncNet (Luo et al., 2021) and NL4DV (Narechania et al., 2020) in leveraging semantic understanding for visualization recommendation.

## CHALLENGES AND OPPORTUNITIES

In this section, we highlight the challenges of automated visualization systems. By doing this, we aim to provide a holistic view of the current state of the field and inspire future research efforts to overcome existing limitations and advance the capabilities of automatic visualization systems.

In this study, the challenges encountered by machine learning (ML) based systems can be generalized into five categories. These include Scope limitations, Data understanding and representation, Interaction and flexibility, Creativity and Generalization and Technical limitations.

**Scope Limitations:** Many systems are constrained by their choice of visualizations, the types of datasets they handle, or the complexity of the tasks they support. For instance, Data2Vis sometimes selects fields with low information values, resulting in simplistic visualizations (Dibia & Demiralp, 2019). Deepeye only supports four types of visualizations (pie chart, bar graph, line chart, and scatterplot), limiting its adaptability to complex datasets (Luo et al., 2018). VizML favours Plotly datasets and focuses on a subset of tasks in visualization recommendation, potentially limiting its applicability to diverse datasets and tasks (Hu et al., 2019).

**Data Understanding and Representation:** Challenges arise in interpreting the semantics of the data and representing it effectively in visualizations. This includes issues such as overlooking semantic meanings between data points, bias in fact generation, and limitations in handling diverse or complex data transformations. For example, Datashot evaluates data in tables individually, ignoring semantic meanings between data during fact sheet generation, leading to potential inaccuracies (Wang et al., 2019). Data Formulator focuses on fundamental transformations, potentially limiting its ability to handle diverse and complicated data transformations (Wang et al., 2023). Chartseer recommended charts have limited effectiveness due to its inability to display temporal details (Zhao et al., 2020).

**Interaction and Flexibility:** Systems that enable natural language interaction or dashboard generation may face limitations in handling iterative queries, semantic parsing, inefficient algorithms for large datasets or providing flexibility in visualization choices. For example, ncNet is limited to processing single queries, limiting its usability for iterative or complex data analytics tasks (Luo et al., 2021). Chat2VIS and Advisor systems lack flexibility since they have semantic parsing limitations (Liu et al., 2021; Maddigan & Susnjak, 2023). NL4DV faces challenges in handling complex query structures (Narechania et al., 2020). MultiVision training time increases with column count, which can limit its scalability for large-dimensional data (Wu et al., 2021). Dashbot considers diversity and parsimony between charts but lacks consideration for the effectiveness of charts in dashboard creation, limiting flexibility in dashboard design (Deng et al., 2022).

**Creativity and Generalization:** Systems for infographic generation or visualization enhancement may struggle with creativity and generalization. This involves challenges such as limited dataset support, understanding the learning process of models, and the need for broader support for various

types of visualizations. For example, Text-to-Viz Capable of handling only proportional assertions, limiting the creativity of infographics (Cui et al., 2019). Chen (Chen et al., 2019) model's reliance on limited datasets and the learning process being a 'black box' may hinder generalization and understanding of the model. The system Qian (Qian et al., 2020) presented is limited to proportion-related infographics, requiring generalization to support other types. Storylines created by the Plotthread system lack creativity (Tang et al., 2020).

**Technical Constraints:** Limitations in handling poor-quality inputs, inefficient techniques or accommodating diverse user preferences, also play a significant role in hindering system performance and effectiveness. For example, the oversimplified scoring techniques in Qian (Qian et al., 2021) proposed system result in visualizations with the same score, making it difficult to rank them. The Viscode system struggles with deformed input images, and there is a limit to the amount of information that can fit into visualizations (Zhang et al., 2020). The Infocolorizer struggles to recommend proper color palettes for data charts in infographics (Yuan et al., 2021).

Even though each system may face a unique set of challenges, there are some common ones. For example, many systems have difficulty scaling, especially when confronted with complex tasks or large datasets. Many systems struggle with flexibility and adaptability, particularly when dealing with changing user needs, preferences, and data types.

The challenges associated with systems that do not use machine learning models are as follows. Template-based methods, in general, have a limited ability to represent diverse visualization types, lack flexibility, struggle with handling dynamic visualization elements, require manual corrections, and may produce less accurate and detailed captions. For example, the tool introduced by Harper and Jonathan (Harper and Jonathan, 2017) demonstrates limitations in representing complex visualizations and struggles with handling dynamic visualization elements. Lu's (Lu et al., 2021) approach is frequently criticized for its repetition and lack of flexibility in visualizations. AutoCaption faces issues regarding the accuracy and complexity of generated captions for visualizations (Liu et al., 2020). Cui's (Cui et al., 2021) system faces a significant challenge in managing chart-level elements and requires manual corrections.

Algorithm based/ Rule-based systems, in general, lack the adaptability to deal with complicated data or tasks that require expert knowledge. They may suffer from scalability, interactivity, and incorporation of effective semantic understanding. For example, Calliope generated captions are overly rigid and lack semantic depth (Shi et al., 2020). Datashot has scalability issues with large multidimensional datasets (Wongsuphasawat et al., 2015). Voyager systems have limited effectiveness in performing specific tasks and lack interactivity in visualizations (Srinivasan et al., 2018). KG4Vis faces issues with adaptability to deal with complicated tasks that require expert knowledge; it has limited visualization design choices and lacks flexibility (Li et al., 2021).

## DISCUSSIONS AND LIMITATIONS OF SURVEY

In this section, we discuss the findings of our survey and their implications. From closely looking at the current automated visualization systems most of them often provide predefined templates or algorithms for generating visualizations, which may limit the level of customization possible. The majority of existing autonomous data visualization systems use simple chart formats such as line charts, bar charts, pie charts, and scatter plots. While these chart formats are useful for expressing specific types of data, they may not be able to visualize complex datasets. One of the most significant shortcomings of these systems is they lack the contextual understanding of the data that human analysts possess. Ignoring context between the data might result in visualizations that fail to properly express significant aspects of the data. Without regard for the meaning behind the data points, visualizations may appear disconnected or misleading, making it difficult for users to analyze and derive useful insights. Some of the systems deal with empty values and avoid such rows from the tables, but the majority of the systems do not deal with data quality before generating the visualizations. For example, If the data is incomplete, inconsistent, or contains errors, it can lead to inaccurate or misleading visualizations. Most of the generated visualizations lack interactivity. None of the visualization systems focus on the security aspect when dealing with sensitive data. Hierarchical data visualization allows for easy comparison between different levels of hierarchy, enabling users to identify similarities, differences, and patterns across various parts of the data. Notably, the major finding of our survey is that there is a significant gap in the automated visualization of hierarchical data, which suggests promising directions for future work.

The survey we conducted is not exhaustive. We excluded theory and evaluation publications from our research. Evaluating the generated visualization is a crucial aspect in knowing the effectiveness of the visualizations. Automatic data visualization is an active research topic with significant potential in the field of visual analytics, and we anticipate many more studies will be undertaken to improve our understanding.

## CONCLUSION

In this study, we reviewed 31 top-tier papers related to current methods in automated data visualization. We focused on the aspects of automating related to data visualization to get a comprehensive overview. Then we divided the existing systems into ML-based approaches, template-based approaches, and rule-based approaches that use a certain technique/algorithm. We highlighted the challenges of the field and potential research prospects in autonomous data visualization based on our study of the papers. This survey will provide important insights into autonomous data visualization and encourage further research.



## REFERENCES

- Chen, Z., Wang, Y., Wang, Q., Wang, Y., & Qu, H. (2019). Towards automated infographic design: Deep learning-based auto-extraction of extensible timeline. *IEEE transactions on visualization and computer graphics*, 26(1), 917–926.
- Cui, W., Wang, J., Huang, H., Wang, Y., Lin, C.-Y., Zhang, H., & Zhang, D. (2021). A mixed-initiative approach to reusing infographic charts. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 173–183.
- Cui, W., Zhang, X., Wang, Y., Huang, H., Chen, B., Fang, L.,... Zhang, D. (2019). Text-to-viz: Automatic generation of infographics from proportion-related natural language statements. *IEEE transactions on visualization and computer graphics*, 26(1), 906–916.
- Deng, D., Wu, A., Qu, H., & Wu, Y. (2022). Dashbot: Insight-driven dashboard generation based on deep reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics*, 29(1), 690–700.
- Dibia, V., & Demiralp, Ç. (2019). Data2vis: Automatic generation of data visualizations using sequence-to-sequence recurrent neural networks. *IEEE computer graphics and applications*, 39(5), 33–46.
- Harper, J., & Agrawala, M. (2017). Converting basic D3 charts into reusable style templates. *IEEE transactions on visualization and computer graphics*, 24(3), 1274–1286.
- Hu, K., Bakker, M. A., Li, S., Kraska, T., & Hidalgo, C. (2019). Vizml: A machine learning approach to visualization recommendation. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.
- Li, H., Wang, Y., Zhang, S., Song, Y., & Qu, H. (2021). KG4Vis: A knowledge graph-based approach for visualization recommendation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 195–205.
- Liu, C., Han, Y., Jiang, R., & Yuan, X. (2021). Advisor: Automatic visualization answer for natural-language question on tabular data. 2021 IEEE 14th Pacific Visualization Symposium (PacificVis).
- Liu, C., Xie, L., Han, Y., Wei, D., & Yuan, X. (2020). AutoCaption: An approach to generate natural language description from visualization automatically. 2020 IEEE Pacific visualization symposium (PacificVis).
- Lu, J., Chen, W., Ye, H., Wang, J., Mei, H., Gu, Y.,... Ma, K.-L. (2021). Automatic generation of unit visualization-based scrollytelling for impromptu data facts delivery. 2021 IEEE 14th Pacific visualization symposium (PacificVis).
- Luo, Y., Qin, X., Tang, N., & Li, G. (2018). Deepeye: Towards automatic data visualization. 2018 IEEE 34th international conference on data engineering (ICDE).
- Luo, Y., Tang, N., Li, G., Tang, J., Chai, C., & Qin, X. (2021). Natural language to visualization by neural machine translation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 217–226.
- Maddigan, P., & Susnjak, T. (2023). Chat2vis: Generating data visualisations via natural language using chatgpt, codex and gpt-3 large language models. *IEEE Access*.
- Narechania, A., Srinivasan, A., & Stasko, J. (2020). NL4DV: A toolkit for generating analytic specifications for data visualization from natural language queries. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 369–379.
- Qian, C., Sun, S., Cui, W., Lou, J.-G., Zhang, H., & Zhang, D. (2020). Retrieve-then-adapt: Example-based automatic generation for proportion-related infographics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 443–452.

- Qian, X., Rossi, R. A., Du, F., Kim, S., Koh, E., Malik, S.,... Chan, J. (2021). Learning to recommend visualizations from data. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*.
- Shi, D., Xu, X., Sun, F., Shi, Y., & Cao, N. (2020). Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 453–463.
- Srinivasan, A., Drucker, S. M., Endert, A., & Stasko, J. (2018). Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE transactions on visualization and computer graphics*, 25(1), 672–681.
- Tang, T., Li, R., Wu, X., Liu, S., Knittel, J., Koch, S.,... Wu, Y. (2020). Plot Thread: Creating expressive storyline visualizations using reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 294–303.
- Teng, Z., Fu, Q., White, J., & Schmidt, D. C. (2021). Sketch2Vis: Generating data visualizations from hand-drawn sketches with deep learning. 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA).
- Wang, C., Thompson, J., & Lee, B. (2023). Data Formulator: Ai-powered concept-driven visualization authoring. *IEEE Transactions on Visualization and Computer Graphics*.
- Wang, Q., Chen, Z., Wang, Y., & Qu, H. (2021). A survey on ML4VIS: Applying machine learning advances to data visualization. *IEEE transactions on visualization and computer graphics*, 28(12), 5134–5153.
- Wang, Y., Sun, Z., Zhang, H., Cui, W., Xu, K., Ma, X., & Zhang, D. (2019). Datasheet: Automatic generation of fact sheets from tabular data. *IEEE transactions on visualization and computer graphics*, 26(1), 895–905.
- Wongsuphasawat, K., Moritz, D., Anand, A., Mackinlay, J., Howe, B., & Heer, J. (2015). Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE transactions on visualization and computer graphics*, 22(1), 649–658.
- Wu, A., Wang, Y., Shu, X., Moritz, D., Cui, W., Zhang, H.,... Qu, H. (2021). Ai4vis: Survey on artificial intelligence approaches for data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28(12), 5049–5070.
- Wu, A., Wang, Y., Zhou, M., He, X., Zhang, H., Qu, H., & Zhang, D. (2021). MultiVision: Designing analytical dashboards with deep learning-based recommendation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 162–172.
- Yuan, L.-P., Zhou, Z., Zhao, J., Guo, Y., Du, F., & Qu, H. (2021). Infocolorizer: Interactive recommendation of color palettes for infographics. *IEEE Transactions on Visualization and Computer Graphics*, 28(12), 4252–4266.
- Zhang, P., Li, C., & Wang, C. (2020). Viscoder: Embedding information in visualization images using encoder-decoder network. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), 326–336.
- Zhao, J., Fan, M., & Feng, M. (2020). Chartseer: Interactive steering exploratory visual analysis with machine intelligence. *IEEE Transactions on Visualization and Computer Graphics*, 28(3), 1500–1513.
- Zhu, S., Sun, G., Jiang, Q., Zha, M., & Liang, R. (2020). A survey on automatic infographics and visualization recommendations. *Visual Informatics*, 4(3), 24–40.

---

# Classification of Uncertainties in Agile Development of Mechatronic Systems

**Kristin Paetzold-Byhain and Franziska Scharold**

Institute of Machine Elements and Machine Design, TU Dresden, Dresden 01069, Germany

## ABSTRACT

Agile methods are increasingly important for developing mechatronic systems in dynamic and volatile environments. These methods help development teams to deal with uncertainties in the development process and utilize them as opportunities. A feature of agile methods is the active involvement of developers in the flow of communication and information, which has a positive effect on the quality of decision making in terms of the development process. It is essential that the developer understands, accepts, and respects the uncertainty in the development task/process to realistically evaluate development scenarios and make well-founded decisions. Tailor-made approaches are required to deal with uncertainties in the development process, as the causes of these are manifold. This paper is based on extensive literature research, analyses of agile development processes at industrial partners and a series of studies on the agile development of physical products that have been carried out regularly for six years. The goal is to differentiate uncertainties in such a way that the artifacts and activities adopted in agile development can be used and adapted to deal with uncertainties more effectively.

**Keywords:** Uncertainties, Agile product development, VUCA

## INTRODUCTION

To respond appropriately in the dynamic and volatile development of mechatronic systems, agile methods are becoming increasingly important. The provision of mechanisms serves to facilitate the management of uncertainties inherent to the development process, enabling their acceptance and subsequent exploitation as opportunities (Böhmer et al., 2015). One characteristic of agile development is the active integration of developers into communication and information flows. It has a positive impact on the quality of decision-making in product development. This is precisely why it is necessary to recognise, accept and respect uncertainties in the development process in order to be able to realistically assess development situations and react in a context-specific manner (Thunnissen, 2003). Product development processes are inherently characterized by uncertainties (Freisleben and Vajna, 2002). There are many reasons for this, but they are primarily based on the fact that target systems are initially only theoretical ideas that need to be successively concretised (Bender and Gericke, 2021). Mechatronic systems are developed through the division of labour. Which means that

sub-functions must be defined to implement the overall system behaviour, resulting in organizational and functional interfaces. Further interfaces may be added during the specification of the solution due to decisions on solution principles, which are often not sufficiently communicated (Ehrlenspiel and Meerkamm, 2013). Development tasks are often processed asynchronously. This means that the necessary information may not be available at the beginning of the development process and must first be supplemented by assumptions. This leads to uncertainties that are associated with risks, both for the development process itself and for the product or system behaviour (Thunnissen, 2003). From a project perspective, uncertainties arise primarily from the novelty of the development task, but also from development constraints (Rupp, 2014). Product development is aware of these risks due to the inherent uncertainty of the development process and the constraints imposed by the project's boundary conditions. The utilization of risk analysis methodologies (e.g., as exemplified in Goldberg et al. (1994)) is employed, yet these processes are time-consuming and costly, ultimately impeding the availability of development resources for the identification of solutions (Sutherland, 2014). This domain encompasses the application of agile development methodologies, which are founded upon the values postulated in the Agile Manifesto (Beck et al., 2001) and aim to minimize non-value-adding activities. Agile principles and procedures facilitate the ability of development organizations to remain flexible in dynamic and volatile conditions and to respond quickly to change (Böhmer et al., 2015).

Böhmer et al. (2015) define agility as follows: “*Agility is the capability to react, and adopt to expected and unexpected changes within a dynamic environment constantly and quickly; and to use those changes (if possible) as an advantage.*”

Uncertainty is an inherent characteristic of complex systems. A key element of agile development is that the iterative and incremental process not only evaluates the results regarding the fulfilment of the requirements, but also the goals underlying the requirements. The simultaneous consideration makes it possible to recognise discrepancies between the original goals and the current reality, thus enabling the goals to be realigned or corrected (Schrof and Paetzold, 2020). This approach provides a foundation for a new way of dealing with uncertainty.

- Agile approaches such as Scrum (Schwaber and Sutherland, 2020), which are currently widely used (Weiss et al., 2023), were originally designed to enable teams of eight to ten developers to complete development tasks in a manageable amount of time. Adaptations of agile methods appear necessary, as complex mechatronic systems are always developed based on a division of labour, whereby the individual development tasks are not independent of each other. This aspect is addressed with scaling process models such as LeSS or SAFe (Dingsøyr and Moe, 2014).
- The physicality of mechatronic systems requires suitable adaptations (Ovesen, 2012) to employ elements such as increments and prototyping. Ultimately, this is necessary in order to preserve the effects of agile methods in terms of flexibility, self-organisation, transparency and the organisation's ability to learn (Heimicke et al., 2020).

The **research question** that forms the basis of this paper is derived from the aforementioned considerations:

**What are the criteria for identifying and evaluating uncertainties in the development process in order to draw conclusions about necessary and targeted adaptations of agile methods in the specific organizational context?**

## **THEORETICAL FOUNDATIONS**

To identify and analyse uncertainties in the development of mechatronic systems, a comprehensive understanding of uncertainties is required. To this end, relevant existing approaches for describing and analysing uncertainties as well as agile development methods are presented in the following.

### **Approaches to the Identification and Classification of Uncertainties**

Uncertainty plays an important role in all types of decision-making processes, as it is not possible to assume completely deterministic conditions (Walker et al., 2003). Consequently, the management of uncertainties is an issue of concern across various scientific disciplines, including political and social sciences, economics, and engineering. The diverse perspectives on uncertainties arise from a multitude of factors, including the system under consideration, the decisions to be made, and the boundary conditions to be taken into account (Walker et al., 2003). Thunnissen (2003) provides an overview of approaches from various areas of expertise. Despite the disparate approaches employed, it becomes evident that there are similarities and patterns in their interpretation, classification, and root cause analysis. Thunnissen (2003) summarizes these findings in his classification of uncertainties for the design of complex technical systems. He differentiates between **ambiguities**, **epistemic** and **aleatory uncertainties**, as well as **interactions**, and further subdivides these (Thunnissen, 2003). **Ambiguities** describe situations in which the decision is not based on a clear idea of its effects. Such ambiguities are based not least on uncertainties regarding cause-effect relationships. As a result, it is not possible to accurately assess the impact of changing boundary conditions (Engelhardt, 2013). **Aleatory uncertainties** are those that arise due to random variations in influencing variables. In contrast, **epistemic uncertainties** arise when there is a lack of precise knowledge about the object under consideration or when the available knowledge is incomplete (Bedford and Cooke, 2001). Thunnissen (2003) adds **interactions** as effects in which system elements interact with each other unexpectedly. In Haberfellner et al. (2019), this phenomenon is also referred to as emergence. Agile methods appear to be particularly suitable for so-called VUCA conditions. VUCA stands for volatility, uncertainty, complexity, and ambiguity. A detailed analysis shows that these characteristics are not independent of each other but are hierarchically dependent (Penzik et al., 2023). Accordingly, complexity is seen as the primary source of uncertainty. The diversity and multiplicity of elements, which in turn are connected to each other by a diversity and multiplicity of relationships, results in a high degree of variety and connectivity (Patzak, 1982). High connectivity, in turn, can lead to emergence, i.e. to situations in which

the system does not react in the expected way. The variability of system elements and their dynamic behaviour over time is understood as a cause of volatility (Patzak, 1982). Uncertainties in the boundary conditions of development lead to ambiguities (Waller et al., 2019) that are either aleatory or epistemic in nature. They also require different strategies for dealing with them (Engelhardt, 2013). While aleatory uncertainties are primarily information problems, i.e., the data required for decision-making exists but is not available in the specific situation, epistemic uncertainties are more of a knowledge problem. Interdependencies are virtually unknown.

McManus and Hastings (2005) focus on the understanding of uncertainties with the aim of deriving implications for the design of complex technical systems. Based on the causal chain, “*uncertainties lead to risks/opportunities lead to results*” McManus and Hastings (2005, 485) identified causes of uncertainties:

**Lack of knowledge** describes that elements or facts are required but are not or not yet available at the given time to be able to make rational decisions. **Lack of specifications** or definitions describes that elements or facts are not yet specified but would be necessary for a valid decision. **Statistical uncertainties** describe elements that are known but whose characteristics fluctuate or are statistically characterized. **Known unknowns** refer to elements that are known to be considered but are not known. **Unknown unknowns** are elements that are not known per se and are therefore not considered.

Earl et al. (2005) have further developed this framework by arranging the identified uncertainties on two orthogonal axes. In this way, they create a framework for a further differentiation of uncertainties. The previously defined categories of information problem and knowledge problem can also be assigned here.

Walker et al. (2003) examine conceptual foundations for managing uncertainties for decision support. System models form the basis for classifying and assessing uncertainties. Models are always only representations of reality that make the complexity of the systems manageable by reducing it according to the purpose of the model (Bossel, 2018). Stachowiak (1973) defines three characteristics of models: Reduction Feature, Pragmatic Feature, and Mapping Feature, each of which addresses specific uncertainties and can therefore be taken into account. The basis for the uncertainty analysis is the categorisation of uncertainties into categories similar to those used by Thunnissen (2003), from which specific characteristics for uncertainties can be derived. Three dimensions are defined:

**Localization of uncertainty:** Uncertainties come from the object under consideration itself, the environment of the object under consideration or from the data with which the model of the object is fed.

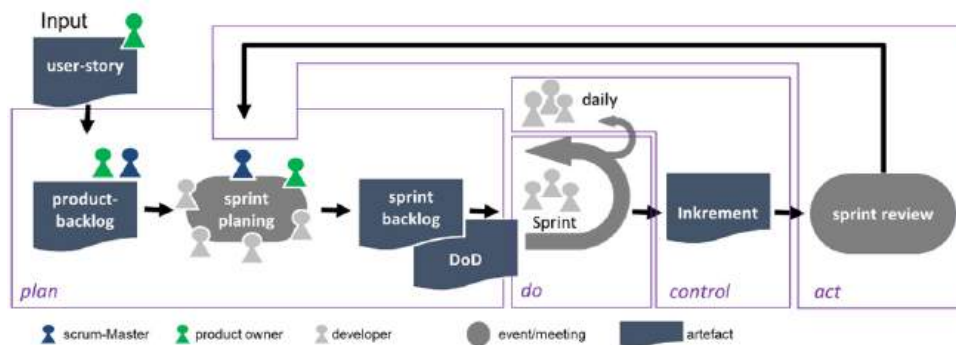
**Degree of uncertainty:** Addresses the gradual transition between determinism and complete ignorance. Although subcategories are named here (statistical uncertainty, scenario uncertainty, recognized uncertainty, and complete ignorance), these correspond to the categories identified in McManus and Hastings (2005). However, when applied to different locations or sources of uncertainty, there may well be differences in meaning.

**Type of uncertainty:** This in turn differentiates according to the effort required to gain knowledge to reduce uncertainties. Analogous to the aforementioned sources, a distinction is also made here between epistemic and aleatory uncertainties.

While the localization of uncertainties helps to identify them, the categories relating to the degree of uncertainty are helpful in analysing and specifying the reason for them. The type of uncertainty describes the effort required and the options for reducing it. In Walker et al. (2003), an uncertainty matrix is derived on the basis of this multidimensional categorization. This overview is designed to enhance awareness of the interdependencies of uncertainties, with the objective of enabling a more detailed analysis and control of their effects.

### Agile Approaches in Mechatronic Systems Development

Agile approaches are characterized by a high-frequency iterative, incremental approach. Development teams create increments at predefined intervals that can be validated by the customer (Schwaber and Sutherland, 2020). In addition to Scrum as one of the most popular methods (Digital.Ai, 2021), there are many other methods that essentially follow the same logic. Comprehensive descriptions of agile methods can be found in Dingsøyr et al. (2012).



**Figure 1:** Scrum process; according to Schwaber and Sutherland (2020).

The process model behind Scrum is summarized in Figure 1. The indicated cycle comprises both technical-physical development activities and technical management activities. The cycle begins with sprint planning based on user stories and the product backlog (Plan). During the sprint, the defined task is completed using product development methods (Do). Daily meetings serve as a control element, where the current work status is recorded, and problems and difficulties are discussed. At the end of a sprint, the functionalities implemented in prototypes (Increment; Control) are discussed with customers, product owners and other stakeholders in sprint review meetings (Act). This creates further potential to sharpen requirements, recognize problems and initiate changes. It enables the development team to remain agile, minimize waste in resources and ultimately deliver a

product tailored to the customer's needs. Sprint retrospectives, which take place after a sprint review but before the next sprint planning, give the team the opportunity to reflect on their own work as part of a learning process. An analysis of agile development methods and their elements reveals fundamental strategies for dealing with uncertainties:

- Elements that support team communication create transparency and help to solve information problems.
- Elements of customer integration help to clarify and concretize requirements and boundary conditions regarding customer needs and product use.
- Prototyping serves two primary purposes: it validates the results with the customer and it facilitates exploration by developing technical interrelationships. This approach helps to reduce the knowledge problem. The incremental and iterative approach in short cycles allows decisions to be postponed, thereby creating opportunities to complete the knowledge and information base, which in turn helps to reduce uncertainty.

However, in the development of complex mechatronic systems, there are also uncertainties that are determined by the product characteristics but are not explicitly addressed by elements of agile development:

- Development is based on a division of labour; the separate view can lead to a loss of integration into the overall context. Conflicts of objectives are difficult to identify, functional and organisational interfaces are not considered.
- The required regular prototyping after each increment is difficult to realize. This is not only due to the division of the development process into phases (planning, conceptualizing, designing, integrating (Bender and Gericke, 2021)), but also due to the resulting product itself. As a result, vertical prototypes are increasingly being used (Weiss et al., 2023), which only represent parts of the product. This means that the properties of the prototype do not correspond to those of the entire product, which leads to distortions in the interpretation of the system behaviour.

The previous points demonstrate that agile methods for the development of mechatronic systems require adaptations, additions, and novel elements to realize improved uncertainty management.

### **APPROACH FOR IDENTIFYING AND CLASSIFYING UNCERTAINTIES IN AGILE DEVELOPMENT PROCESSES**

The aforementioned classification schemes must be integrated to analyze uncertainties in the development process and adapted problem specific. The starting point is the differentiation and definition of types as proposed by McManus and Hastings (2005). Following Walker et al. (2003) the types should be differentiated according to the place of origin too. The systemic approach is useful here, as it also considers uncertainties in relation to the system description and the data and information used. The distinction between epistemic and aleatory uncertainties serves above all as an indication



**Table 1. Structuring matrix for uncertainties in agile product development.**

Context	Location	Level						
		Statistical uncertainties	Scenario uncertainties	Recognised ignorance				
Market/environment	Development task	Variations in relevant values	IP	Requirements uncertainties	IP	Novelties in technology	IPWP	
	Customer access	Variations in development tasks	IP			Additional unknown effects	WP	
		Variation in degree of novelty						
Company	Strategy	Degree of individuality	IP	Usage situation User behaviour Diversity of variants Freedom of choice in decision Scope for decision making	WP	Inadmissible assumptions about user behaviour	WP	
		Current order situation	IP			Priorities in development strategies	IP	
	Structural and process organisation	Uncertainties about specific contact persons	IP	Importance of roles for own task				
		Importance of people for own task						
Product architecture	IT-structure	Current versions, releases	IP	Authorisation for data access	IP			
		Accessibility of information carriers	IP	Which persons are information carriers in context	IP			
	Functional/ behavioural architecture	Variations in the function parameters	WP/PP	Uncertainties about interfaces unclear about secondary functions additional interfaces	WP/PP	Unknown cause-effect relationships		WP
		Variations in the structural characteristics	WP/PP	Uncertainties about interfaces	IPWP			
Task distribution	Organisational constraints	Importance of the task in the overall system	IP	Importance of one's own task in the overall system	IP			
		Uncertainty about the timeliness of the requirements	WP	Availability of requirements Responsibility for requirements	IP	Conflicting objectives in requirements	IPWP	
Task description	Data availability	Personal interpretation of the task Importance of the task in the overall system	IP	Unclear responsibilities	IP	Misunderstanding the requirements	WP	
		Actuality of the data includes availability, maturity, certainty, completeness, data security	IP	Data storage location Access to data sources	IP	Existence of the data	IP	
Requirements management	Information sources (communication)	Temporal unavailability of information carriers	IP	Uncertainties about communication channels Inhibitions to use information carriers	IP	Uncertainties about communication structures Sources of information are not known	IP	
		Out-of-date information						

IP = Information problem; WP = Knowledge Problem

of which strategies can provide support in dealing with them, as these are indicative of an information or knowledge problem. The identification of uncertainties is a consequence of the analysis of the development process, which must be considered in conjunction with the management processes that are employed to organise the development work. The findings were summarised in a matrix for identifying and describing uncertainties based on Walker et al. (2003) (see Table 1). The next step is to complete and refine the listed uncertainties. To this end, interviews and workshops will be conducted together with an industry partner, to check the approach for accuracy and plausibility, and to discuss and complete the contents of the matrix itself. This then serves as the basis for determining specific elements for the development process, with the objective of enabling the management of uncertainties.

## CONCLUSION

The paper presents a matrix for identifying uncertainties in the development process. Several uncertainties were also found to be linked through causal chains as the matrix was built. This is in line with the findings from the studies on VUCA, as described above. For taking this into account, it is necessary to extend the approach to consider the dependencies between uncertainties in the development process. Based on the matrix, elements for dealing with these uncertainties were developed and integrated into the development process with the intention of supporting agile product development.

## REFERENCES

- Beck, K./Beedle, M./van Bennekum, A./Cockburn, A./Cunningham, W./Fowler, M./Grenning, J./Highsmith, J./Hunt, A./Jeffries, R./Kern, J./Marick, B./Martin, R./Mellor, S./Schwaber, K./Sutherland, J./Thomas, D. (2001). Manifesto for Agile Software Development. Available online at [agilemanifesto.org](http://agilemanifesto.org) (accessed 4/16/2024).
- Bedford, Tim/Cooke, Roger (2001). What is uncertainty? In: Probabilistic Risk Analysis: Foundations and Methods, 17–38.
- Bender, Beate/Gericke, Kilian (2021). Pahl/Beitz Konstruktionslehre. Berlin, Heidelberg, Springer Berlin Heidelberg.
- Böhmer, Annette Isabel/Beckmann, Andreas/Lindemann, Udo (2015). Open Innovation Ecosystem - Makerspaces within an Agile Innovation Process. In: ISPIM Innovation Summit.
- Bossel, Hartmut (2018). Modeling and Simulation. A K Peters/CRC Press.
- Digital. Ai (2021). 15th Annual State of Agile Report. Available online at <https://digital.ai/resource-center/analyst-reports/state-of-agile-report/> (accessed 11/5/2023).
- Dingsøy, Torgeir/Moe, Nils Brede (2014). Towards Principles of Large-Scale Agile Development. In: Torgeir Dingsøy/Nils Brede Moe/Roberto Tonelli et al. (Eds.). Agile Methods. Large-Scale Development, Refactoring, Testing, and Estimation. Cham, Springer International Publishing, 1–8.
- Dingsøy, Torgeir/Nerur, Sridhar/Balijepally, VenuGopal/Moe, Nils Brede (2012). A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software* 85 (6), 1213–1221. <https://doi.org/10.1016/j.jss.2012.02.033>
- Earl, Chris/Johnson, Jeffrey/Eckert, Claudia (2005). Complexity. In: John Clarkson/Claudia Eckert (Eds.). Design process improvement. London, Springer London, 174–197.

- Ehrlenspiel, Klaus/Meerkamm, Harald (2013). *Integrierte Produktentwicklung. Denkabläufe, Methodeneinsatz, Zusammenarbeit*. 5th ed. München, Hanser.
- Engelhardt, Roland Alexander (2013). *Uncertainty mode and effects analysis - heuristische Methodik zur Analyse und Beurteilung von Unsicherheiten in technischen Systemen des Maschinenbaus*. Düsseldorf, VDI-Verlag.
- Freisleben, Dörte/Vajna, Sa'ndor (2002). *Dynamic Project Navigation: Modelling, Improving, and Review of Engineering Processes*. In: Volume 2: 28th Design Automation Conference, ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Montreal, Quebec, Canada, 29.09.2002 - 02.10.2002. ASMEDC, 919–925.
- Goldberg, Benjamin E./Everhart, K./Stevens, R./Babbitt III, N./Clemens, P./Stout, L. (1994). *System engineering toolbox for design-oriented engineers*.
- Haberfellner, Reinhard/Weck, Olivier de/Fricke, Ernst/Vössner, Siegfried (2019). *Systems Engineering*. Cham, Springer International Publishing.
- Heimicke, J./Chen, R./Albers, A. (2020). *Agile Meets Plan-Driven – Hybrid Approaches in Product Development: A Systematic Literature Review*. *Proceedings of the Design Society: Design Conference 1*, 577–586. <https://doi.org/10.1017/dsd.2020.259>
- McManus, Hugh/Hastings, Daniel (2005). *A Framework for Understanding Uncertainty and its Mitigation and Exploitation in Complex Systems*. *INCOSE International Symposium 15 (1)*, 484–503. <https://doi.org/10.1002/j.2334-5837.2005.tb00685.x>
- Ovesen, Nis (2012). *The Challenges of Becoming Agile*. PhD Thesis. Aalborg University.
- Patzak, G. (1982). *Systemtechnik, Planung komplexer innovativer Systeme. Grundlagen, Methoden, Techniken*. Berlin/New York, Springer-Verlag.
- Pendzik, Martin/Sembdner, Philipp/Paetzold, Kristin (2023). *Identification and Classification of Uncertainties as the Foundation of Agile Methods*. *Proceedings of the Design Society 3*, 2165–2174. <https://doi.org/10.1017/pds.2023.217>
- Rupp, Chris (2014). *Requirements-Engineering und -Management. Aus der Praxis von klassisch bis agil*. 6th ed. München, Hanser, Carl.
- Schrof, J./Paetzold, K. (2020). *Agile Produktentwicklung in einer zunehmend dynamischen Automobilwirtschaft: Potentiale und Grenzen/Agile Product Development in an Increasingly Dynamic Automotive Industry: Potentials and Limits*. In: Bernhard Geringer/Hans-Peter Lenz (Eds.). 41. Internationales Wiener Motorensymposium 22–24. April 2020. VDI Verlag, II-362-II-376.
- Schwaber, Ken/Sutherland, Jeff (2020). *The Scrum Guide*. Available online at <https://scrumguides.org/scrum-guide.html> (accessed 10/2/2023).
- Stachowiak, Herbert (1973). *Allgemeine Modelltheorie*. Wien [etc.], Springer.
- Sutherland, Jeff (2014). *Scrum: The Art of Doing Twice the Work in Half the Time*. New York, Crown.
- Thunnissen, Daniel P. (2003). *Uncertainty classification for the design and development of complex systems*. In: 3rd annual predictive methods conference.
- Walker, W. E./Harremoës, P./Rotmans, J./van der Sluijs, J. P./van Asselt, M. B. A./Janssen, P./Kramer von Krauss, M. P. (2003). *Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support*. *Integrated Assessment 4 (1)*, 5–17. <https://doi.org/10.1076/iaij.4.1.5.16466>

- 
- Waller, Robert E./Lemoine, Pamela A./Mense, Evan G./Garretson, Christopher J./Richardson, Michael D. (2019). Global Higher Education in a VUCA World: Concerns and Projections. *Journal of Education and Development* 3 (2), 73. <https://doi.org/10.20849/jed.v3i2.613>
- Weiss, Stefan/Paetzold-Byhain, Kristin/Michalides, Marvin/Pendzik, Martin/Scharold, Franziska/Stoiber, Lino (2023). *Agile Entwicklung physischer Produkte 2023*. Technische Universität Dresden.

# Interpolation and Depth Extraction: A Case Study of Shan Shui Artwork Generated by AI

**Man Lai-Man Tin**

School of Arts and Social Sciences, Hong Kong Metropolitan University, Ho Man Tin,  
Kowloon, Hong Kong

## ABSTRACT

Traditional Shan Shui artworks (Chinese landscape paintings) have been static representations of the beauty and tranquillity of landscapes, and they have a long history and significance in Chinese art. The advancement of artificial intelligence (AI) technologies brings new possibilities to artwork creation and innovation to tradition. This study proposes using AI technologies, specifically artificial neural networks and computer vision, to learn from traditional paintings, generate new landscapes and extract depths in Shan Shui paintings. The research aims to go beyond using AI technology solely to create new artwork. Instead, it explores the ability of AI to generate dynamic landscapes with perspectives, allowing more immersive and engaging experiences, and through analysis of the depths embedded in the AI-generated Shan Shui paintings, trying to gain insights into understanding interpolation of spatial and dimensional aspects in the work and address the limitation of 2-dimension in art. This research signifies the convergence of art and technology, explores novel ways of creating and viewing traditional Shan Shui paintings, and explores the possibilities of understanding the landscapes generated through the lens of AI and computer vision technology.

**Keywords:** Artificial intelligence, Creative AI, Shan Shui painting, Interpolation, Depth extraction, Computer vision, Machine learning art generation

## INTRODUCTION

Beginning in around the 4th century, Shan Shui is a traditional form of Chinese painting that depicts natural landscapes and scenery. The term ‘Shan Shui’ means ‘mountain and water,’ and such paintings always consist of key elements such as mountains, waterfalls, plants, bridges, houses, flowing water, and sometimes feature refined or official scholars. The Shan Shui paintings reflect the artist’s unique observations, travel experiences, artistic representations, imagination, and philosophical representation of the landscape. Gradually, Shan Shui has evolved into a distinct art form characterised by its portrayal of depth and perspective of nature. The composition and geographic elements of Shan Shui paintings represent changes in perspective, signifying the spatial and dimensional travel experience of the artist in nature.

With the advancement of technology, there have been many attempts to use new tools and technologies, including artificial intelligence, to create Shan Shui paintings, which paved a good foundation and exploration of how technology can influence traditional Shan Shui paintings. For example, Alice Xue in her research 'End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks' (2021) introduced the use of Sketch-And-Paint GAN (SAPGAN) to generate Chinese landscape paintings from end to end, without conditional input, demonstrates the capability of AI and deep learning technology to create traditional Shan Shui paintings. However, there has been limited research on applying AI technologies to address and understand the perspective, spatial, and depth aspects of generated Shan Shui paintings. This research proposes the use of artificial intelligence technologies to learn from and generate Shan Shui paintings, aiming to explore the possibilities of creating dynamic landscapes, which seeks to address the limitation of traditional 2-dimensional static artwork and to gain a better understanding of the transformation of perspectives. In the second part of this research, the depth within Shan Shui paintings is further investigated through AI and computer vision technologies, which allows an analysis of the AI-generated landscape correspondence to the composition and understanding of the traditional Shan Shui landscapes.

## **TEST OF SHAN SHUI PAINTING GENERATION**

To train AI, particularly machine learning, in the learning and generation of Shan Shui paintings, tests were conducted using Feng Shui (Chinese Geomancy, which means wind and water) landscape paintings with Shan Shui elements sourced from the internet as datasets. The datasets were used to generate new sets of paintings. 30 Feng Shui landscape paintings were employed in this attempt.

The Feng Shui landscape painting represents a variation of the traditional Shan Shui painting. It embodies some key principles of Shan Shui painting and incorporates the rules of Feng Shui. The structure of the painting and the geographical layout of Shan Shui and Feng Shui elements within the paintings have clear and standardised formats. Thus there is relatively less variation among different Feng Shui landscapes compared to traditional Chinese landscape paintings. This makes Feng Shui landscape paintings a suitable data set for the initial AI training and test to generate pictures with Shan Shui elements.

In this test, the key elements of landscape paintings, such as mountains, waterfalls, and flowing water were extracted and formed new combinations. The results demonstrate that the fundamental layout of Feng Shui landscape paintings is manifested in the generated pictures, indicating the ability of artificial intelligence to learn and generate landscape works, with the landscapes and Shan Shui elements visible in the results (see Figure 1).



**Figure 1:** Samples of Feng Shui landscape paintings generated by AI.

The next phase of this research aims to investigate the potential for creating dynamic, multi-perspective landscapes within Shan Shui paintings, in order to capture the spatial variations experienced across different artworks. The depth embedded in perspectives will be analyzed using AI technology in a subsequent stage.

Traditional Shan Shui paintings, due to the static nature, capture only a static moment of landscapes. To overcome this limitation, Shan Shui artists often employ multiple vanishing points to introduce multi-perspective views of landscape within the artwork. This study aims to explore the possibilities of generating dynamic work by transforming 2-dimensional static images and merging multi-perspective views from different Shan Shui paintings.

To achieve this, interpolation is employed to create interpolated frames between the generated images, creating transitional visual effects between the landscapes. The results demonstrate the ability of artificial intelligence to learn and generate dynamic landscapes (see Figure 2).



**Figure 2:** Samples of interpolated frames created between the generated pictures.

The learning, creation, and generation of paintings featuring Shan Shui elements, along with the presentation of multi-perspectives, are primarily accomplished using AI. The next phase is to learn and generate more complex

Shan Shui paintings of various styles and compositions, based on traditional Shan Shui paintings sourced from the internet.

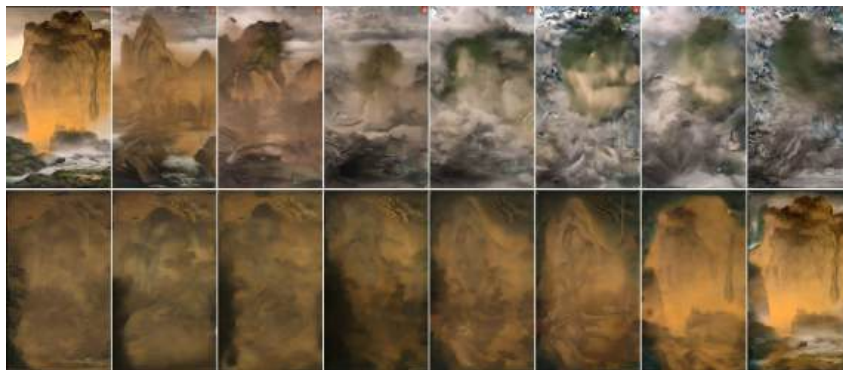
### AI-GENERATED DYNAMIC SHAN SHUI PAINTING

In traditional Shan Shui paintings, instead of directly depicting the landscape, the artists incorporate their own artistic system and expression into the paintings, making them a unique art form of landscape expression. For example, Christin Bolewski, in her paper “‘Shan-Shui-Hua’ – Traditional Chinese Landscape Painting Reinterpreted as Moving Digital Visualisation” (2008) discusses the characteristics of traditional Chinese landscape painting, unlike Western use of single-point perspective, traditional Chinese landscape painting employs multi-perspectives, creating ‘spatial illusions’ in the artwork.

Building upon the findings of the previous tests, this section aims to integrate machine learning with interpolation technology to create dynamic landscapes within Shan Shui paintings. The generated Shan Shui paintings incorporate some of the original compositions and traditional Shan Shui painting elements such as mountains, trees, bridges, and flowing water, as well as the traditional artistic systems of multiple vanishing points, indicating the capacity of AI to learn and generate dynamic Shan Shui paintings while incorporating the distinctive expression of multi-perspectives (see Figure 3 and Figure 4).



**Figure 3:** Samples of the generated traditional Shan Shui paintings with multi-perspectives.



**Figure 4:** Samples of interpolated frames created between the generated Shan Shui paintings.



The dynamic landscapes signify the changes in natural scenery across different landscapes and the travel experiences in nature with multi-spot views from different Shan Shui paintings. The interpolated frames connect the generated Shan Shui, forming an ongoing loop of landscapes. The work extends the spatial and temporal aspects of traditional Shan Shui painting, in which the travel experiences of different sceneries, times, and perspectives are expanded from the traditional static form of artwork. The variation of Shan Shui composition and dimension can guide the viewer's observation through landscapes in more immersive and engaging experiences, allowing for a deeper integration with nature, which echoes the spirit of traditional Shan Shui artistic imagination and philosophical representation of landscapes. For example, 'Daoists, in particular, had a great affinity for experiencing man in the natural environment and seeking a kind of harmonious unification between man and nature,' as highlighted in Sophia Suk-mun Law's research 'Being in Traditional Chinese Landscape Painting' (2011).





















### **DEPTH EXTRACTION THROUGH AI AND COMPUTER VISION**

Traditional Chinese Shan Shui painting features a unique representation of landscape composition and perspective. In addition to key landscape elements such as mountains, waterfalls, and plants, the depiction of spatial and depth aspects is also crucial for understanding how AI-generated works resonate with viewers. To better understand the depth embedded in AI-generated landscapes, AI and computer vision were used to extract the depth from dynamic Shan Shui paintings, attempting to analyse the interpolated spatial and dimensional visual information, which are particularly essential to examine the depth created in the AI landscape and how things can make sense to human in the age of machine intelligence.

Several attempts, such as deep learning and transformer models, which have been available on software platforms, were used to examine the depth in the artwork by estimating depth information from images and videos generated by AI. The results were compared to assess the ability of AI and computer vision to estimate depth accurately and determine whether the estimated depths are aligned (see Table 1).

The results indicate the ability of various AI and computer vision models to analyse spatial information and extract depth from AI-generated landscapes. The depths extracted in set 1 are the most pronounced compared to the other sets. The silhouette of the mountain can be clearly seen, and the levels of both near and far distances in the landscape can generally be differentiated. The deep learning transformer model also distinguishes the depth of the scenery, which can make sense to viewers. However, the depths extracted using the Depth Perception Transformer (DPT) in sets 2 to 5 are unclear. This is likely due to the unclear landscapes created in the interpolated frames, particularly noticeable in the original AI-generated landscapes in sets 3 to 5. Similar outcomes were observed in the use of the deep learning transformer model. As the landscapes become more difficult to interpret, the depth in the Shan Shui cannot be distinguished. For example, in set 5, the depth is essentially just a flat plane without any obvious depth of field or distance. Despite this, neural networks can analyse the dynamic landscape and capture the basic silhouette of the mountain as well as the distance and depth of field.

**Table 1.** Comparison results of depth estimation through AI and computer vision technologies.

Set	Original AI-generated landscapes	Depth extracted by Depth Perception Transformer from still landscapes	Depth extracted by deep learning transformer model from still landscapes	Depth extracted by neural networks from dynamic landscape
1				
2				
3				
4				
5				

With reference to the research by Chi-Min Hsieh and Hsiao-Ching Chou, 'Computational Three Distances: Exploring the Aesthetics of the Southern Song Dynasty and Its Adapted Simulation and Rendering' (2023), the difficulty in interpreting the landscape may also reflect the intersection between the landscape and fog or clouds. For example, more areas of the landscapes are covered by fog or clouds in sets 2, 3, and 4 compared with set 1, in which 'the view of distant mountains from the near bank of a wide body of water' is generated. Additionally, the fragmentary scenery in set 5 makes it difficult to identify the space and depth of the landscape. These align with the traditional theory of distances in Shan Shui painting, namely 'hidden distance,' 'board distance' and 'obscure distance' respectively.

The accuracy of the proposed methods still has room for improvement. Depth can be extracted by utilising AI and computer vision technologies, indicating a certain degree of comprehensibility in AI-generated content. This showcases the potential for AI technology to assist viewers in interpreting the content and therefore how the 'hidden' or 'potential' artistic spirits could be realised through computer program algorithms, as highlighted in 'Rendering and presentation of 3D digital ink landscape painting' Xunxiangby et al. (2023).

## CONCLUSION

This research demonstrates the convergence of art and technology, particularly, traditional Shan Shui paintings and artificial intelligence to learn and create landscape paintings. By employing artificial neural networks and computer vision technologies, the study not only tried to create new artwork, but also explored the ability of AI to generate dynamic landscapes with multi-perspectives, and extraction of depth in the landscapes to enhance the viewer's engagement and examine the comprehensibility of AI-generated content.

The interpolation employed in this research demonstrated the ability of image processing and computer vision to create interpolated frames between the generated landscape paintings. This allowed viewers to engage with the artwork with an experience expanded from traditional spatial and temporal dimensions, in particular, the dynamic perspectives, scenes, and spaces in the landscapes.

Depth extraction using AI and computer vision was another focus of this research. Various attempts were made to analyse the depths embedded in the generated Shan Shui paintings. Although the depths were not always clearly distinguished in some attempts, neural networks generally captured the field of depth when extracted from the dynamic landscape. This indicates the possibility of comprehending AI-generated content through AI and computer vision technologies. This research provides a foundation for future studies on the integration of AI in creative processes and the elements that affect the comprehension of traditional artwork, such as traditional Chinese ink paintings.

## REFERENCES

- Bolewski, C. (2008). 'Shan-Shui-Hua' – Traditional Chinese Landscape Painting Reinterpreted as Moving Digital Visualisation. *Electronic Workshops in Computing*. <https://doi.org/10.14236/ewic/eva2008.4>
- Hsieh, C.-M., & Chou, H.-C. (2023). Computational three distances: Exploring the aesthetics of the Southern Song Dynasty and its adapted simulation and rendering. *SIGGRAPH Asia 2023 Art Papers*. <https://doi.org/10.1145/3610591.3616426>
- Law, S. S. (2011). Being in Traditional Chinese Landscape Painting. *Journal of Intercultural Studies*, 32(4), 369–382. <https://doi.org/10.1080/07256868.2011.584615>
- Li, X., Huang, Y., Jiang, Z., Liu, Y., & Hou, Y. (2023). Rendering and presentation of 3d digital ink landscape painting. *Computer Animation and Virtual Worlds*, 35(1). <https://doi.org/10.1002/cav.2215>
- Xue, A. (2021). End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). <https://doi.org/10.1109/wacv48630.2021.00391>

---

# Enabling Factors in Complex Operations. Lessons From Jazz

**Pedro B. Água and Vítor Conceição**

CINAV, Escola Naval, Instituto Universitario Militar, 2810–001 Almada, Portugal

## ABSTRACT

Contemporary world affairs are facing a multitude of complex problems whose causes are often rooted in past solutions. Despite the ubiquity of systems in human activity, systems thinking and systems sciences are not widely incorporated into mainstream education curriculums. As a result, the lack of comprehensive understanding of systems and their rules has increased complexity resulting from political, economic, social, and technological factors. This has resulted in the emergence of new and complex issues that often prove to be insurmountable. Even sensitive domains such as security and defence have seen their operations being affected in the same complex endeavour. To address these growing problems, which are further exacerbated by the emergence of disruptive technologies, new ways of thinking are required, and non-traditional approaches are needed. One such approach is organisational design and management. Traditional management and organisational schools of thought are ill-equipped to address the rising complexity. Hence, alternative sources of knowledge should be sought. Jazz Organizations is one such field that provides answers to address complex problems. They have been studied by special operations forces and even delicate medical teams, who have derived solutions that can be adapted to their contexts. The Jazz Organization differs from most organisational forms, as it has evolved without much design thought. Like nature's systems, it emerged in a somewhat chaotic manner, yet it is pretty effective in harnessing complexity and delivering delicate operations, where improvisation and creativity are key. Jazz musicians do not abide by routine; in fact, they unlearn routines as a critical skill. They address problem-solving by jumping into uncharted territory, seemingly without fear. They accept "trial and error" as a normal *modus operandi*. Their structures are flexible, continuously adapting in the face of last-minute problems. They learn together within jam sessions, under no central leadership figure, and every musician has their own shining moment during any performance. Jazz musicians introduce disruptions on purpose so the entire ensemble can evolve. A pragmatic analysis is presented by following a logical thinking process, and an alternative approach to organisational design solutions suitable for managing the growing problems associated with complex operations is suggested.

**Keywords:** Complexity, Creativity, Innovation, Jazz, Leadership, Problem-solving, Systems

## INTRODUCTION

The present world paradigm is facing a rise in complexity across many domains, with wicked problems whose origin is rooted in the solutions of the past. Despite the ubiquity of systems in human activity, the

general lack of comprehensive understanding of systems and their rules has fostered increased complexity resulting from political, economic, social, and technological problems. To address such issues, the rise of complex operations has brought new forms of leadership and organisation. Any new major technological revolution (e.g., the industrial revolution or information revolution) has triggered new and innovative ways of leadership, where leaders must be comfortable with ambiguity. Whatever the individual's position, the ambiguity is directly related to the leadership required (DePree, 1992).

This context creates a paradigm where complex operations are the new normal. Depending on the context in which they occur, complex operations can be characterised by distinct features, among them the *interconnectedness* of components and their *interdependencies* within a specific system of interest or operations realm, with the output of some operations being the input into other stages, often with negative or positive feedback mechanisms at play. Such interconnectedness may require a *high degree of coordination* across sociotechnical systems involving people and technologies. This paradigm is commonly characterised by *high risk, uncertainty, or resource intensive*. It frequently requires *technology and systems integration* under a certain interoperability standard, requiring specialised knowledge, adaptation, and improvisation as part of the overall problem-solving approach – in sum, an overall *complexity* paradigm. This is true in the business world, government, and particular fields such as defence and security, where life and society are often at stake.

New leadership, management, and organisation methods are required to address this growing trend in complex operations, further accelerated by disruptive technologies. Traditional command-and-control organisational methods are unsuitable for such complexity paradigms (Rihani, 2002; Axelrod & Cohen, 1999; Dettmer, 2021; Snowden & Rancati, 2022). Moreover, many traditional organisational forms are equally ineffective (Mintzberg, 1979; Mintzberg, 1993).

Synergistic or “harmonic” operations, in which the whole system performs through effectiveness and efficiency as a result of a specific set of operations alignment, which mutually reinforce themselves towards the more significant “systems goal”, are critical for success (Valero y Vicente and Tomás, 1999, p. 97). Moreover, in his *Wisdom of Crowds* masterpiece, Surowiecki (2005) suggests that teams are crucial to success, something even more critical within complex operations, usually bringing together distinct human teams. However, for the desirable systemic performance to emerge within and across teams, an adequate professional coexistence is critical (Valero y Vicente and Tomás, 1999, p. 201), as teamwork demands cooperation, something that has been key along the evolution of societies (Axelrod, 1984), and for its side, cooperation also needs trust, a difficult factor to assure in any cooperative endeavour. Trust has been pointed out as the weakest link in any interoperability, multinational or just multi-organisational effort. It has been the cause of underperformance in cutting-edge projects, missions and complex operations.

After this introduction and the following section providing some relevant background, the third section briefly introduces the used methodology. The fourth section presents the results once the methodology was applied. In the fifth section, a discussion is provided, after which a concluding section ends this paper.

## RELEVANT BACKGROUND

With the rising complexity across all world affairs, businesses, industries, and governments, complex operations show the limits of traditional management and leadership approaches. A detailed analysis and diagnosis of the situation is the starting point for any design solution that is to be put in place. In designing such a solution, one must first understand the context of the particular operations and classify it according to one of five types (Snowden & Boone, 2007): *simple contexts*, *complicated contexts*, *complex contexts*, *chaotic contexts* or *disorder*. Afterwards, and with disregard for the simple and complicated contexts for which solutions are well known, one needs new forms of leadership, management, and organisation. One *sui generis* organisation that may answer complex problems, where adaptation and improvisation are key, is the Jazz Organization, which emerged in an apparently chaotic fashion without much design during the last 100 years. Like most natural systems, it is effective in harnessing complexity and delivering delicate operations, where improvisation and creativity are key (Barrett, 1998; Barrett & Peplowski, 1998; Bernstein & Barrett, 2011). The Jazz organisation leadership styles are not rooted in command and control or traditional forms of organisation (DePree, 1992). Such organisation has been regarded as suitable for innovative environments; it has been studied by special forces and even delicate medical teams, who are studying the jazz organisation in order to adapt it to their contexts (Tjan, 2010; Hulm, 2019; Groysberg & Masko, 2020). Organisations that prosper under complexity exhibit a certain degree of chaos, and their leaders will make some sense of it (DePree, 1992; Weick & Sutcliffe, 2015).

According to Barrett (1998, 2012), the jazz organisation has seven key characteristics that foster the emergence of desirable organisational performance and which may help lead and manage complex operations:

*Jazz musicians do not abide by routine*; in fact, they unlearn routines as a critical skill. In the same sense that “a ship is safe at a harbour, yet was not built for such purpose”, routines are comfort zones that prevent evolution and achievement of higher performance levels, especially in the context of critical operations with all the inner complexity. Jazz musicians intentionally introduce disruptions so the entire ensemble can evolve into a better unknown.

*Jazz musicians are comfortable jumping into the unknown*. They address problem-solving by jumping into uncharted territory, seemingly without fear. Solving complex challenges and problems, such as those found in leading and managing complex operations, requires improvisation and easiness with the unknown. Miles Davis is known for opposing routines and jumping into the unknown, fostering innovation by putting his jazz ensembles out of balance

in search for new songs and melodies. The *Kind of Blue*, Miles Davis's most famous record, which took just two short recording sessions in 1959, is a clear product of this way of acting (Austin & Stormer, 2008).

*Jazz musicians naturally engage in trial and error.* They accept "trial and error" as a normal modus operandi. For someone watching and listening to a jazz ensemble trial, it may, at first sight, appear chaotic and without any strategy, which is due to jazz musicians being comfortable experimenting with new avenues of trial and error to find new songs and melodies.

*Jazz ensembles use flexible structures.* Their structures are flexible, continuously adapting in the face of last-minute challenges. Without some flexibility, a system such as a jazz ensemble could not innovate or evolve. Some constraints must be present for evolution to take place. However, a certain level of flexibility is a prerequisite for achieving such evolution (Snowden & Rancati, 2022). The jazz organisation encourages experimentation and exploration, which enable jazz musicians' intuition to develop and create new solutions – aka new songs. The introduction of variation is one of the main levers to change a system (Garcia, 2004), and a systems thinking approach is critical for learning, as evidenced by distinct organisations, such as hospitals (Tucker & Edmondson, 2003).

*Jazz ensembles provide the proper context for learning and innovation.* They learn together in jam sessions, which provides a unique leadership lesson as there is no central leadership figure, and every musician has his own shining moment during any performance. Jazz sessions, being removed of traditional top-down command and control or other forms of centralised leadership, provide just the right environment for learning and innovation.

Leadership at its highest happens when someone follows someone else while he/she is under no obligation to do so. Therefore, most traditional references to leadership are ill-placed as they refer to paradigms where either command and control has been the rule, or people were (or are) part of organisations where an outrageous unbalance of power exists, so people are forced to accept their organisations' ruling. Such a "pseudo leadership" paradigm is not possible in jazz organisations. Therefore, the benefits of jazz organizations in terms of learning and innovation are barely matched by most current world forms of leadership, management, or organisation. Moreover, creative people must work with others of equal competence, which is not readily achievable within traditional organisational forms (Depree, 1992).

*Jazz musicians develop a superb ability to follow others.* To lead effectively, one has to first learn followership, which demands superb personal skills and abilities—self-mastery in the first place. Besides being among the best musicians in the world, jazz musicians develop humble and crucial followership abilities. Only then can they make a relevant contribution within the complex operations context of which a jazz ensemble really is an example.

*The ability to perceive an individual's or group's potential while it is not being displayed.* Even if there is no centralised leadership within the jazz organisation, with every jazz player having his or her leading moments at a time within the group, some jazz musicians develop the ability to perceive other jazz musicians' potential, encouraging them to go off limits in experimenting new avenues, hence creating new songs. Miles Davis is well



regarded as a distinct jazz musician, as he found several talented musicians under his ensemble sessions. Saxophonist John Coltrane is one such example (Austin & Stormer, 2008).

Having presented some of the main characteristics of the jazz organisation and surfaced some of the leadership features in a type of organization without centralized leadership or traditional command and control, a solution can be envisioned to adapt such knowledge to complex operations in diverse business contexts.

## METHODOLOGY

Departing from a Business Policy approach to organisations (Valero y Vicente and Tomás, 1999), the methodological approach takes a Business Policy approach to organizations and the Theory of Constraints (TOC) (Goldratt, 1990; Goldratt; 1992). The business policy body of knowledge also contributes to the management of complex operations, which is not a theory but a task, which starts with analysis and diagnosis, followed by a future paradigm selection (Valero y Vicente & Tomás, 1999, p. 81). For such Analysis and diagnosis, the TOC's four fundamental questions are in order (see Table 1).

**Table 1.** The TOC questions (Dettmer, 2003).

Key Questions	Objectives
Why change?	To pursue a desirable goal or new paradigm
What to Change?	What needs to change in order to eliminate undesirable effects
What to Change To?	What do changes look like to achieve desirable effects?
How to cause the change?	What enabling conditions or prerequisites are needed to reach the desirable effects?

Answering the first two questions ('Why change?' and 'What to change?') is related to the problem definition and the realisation of such a problem's consequences if it is not adequately addressed. The 'Why change?' question relates to the election of *the Goal* for the problem or *system of interest* (SoI). In the present case, the goal may be defined as *Adequate Leadership and management of Complex Operations Achievable*, which from its side implies management and leadership of complex operations. Answering the last two questions ('What to change to?' and 'How to cause the change?') prompts one to design a solution to address the defined problem(s) and how to deploy such a solution with the intended net benefits. The Logical Thinking Process, central to Goldratt's Theory of Constraints, is behind the methodological approach used in this paper (Goldratt, 1992; Dettmer, 2003; Dettmer, 2021).

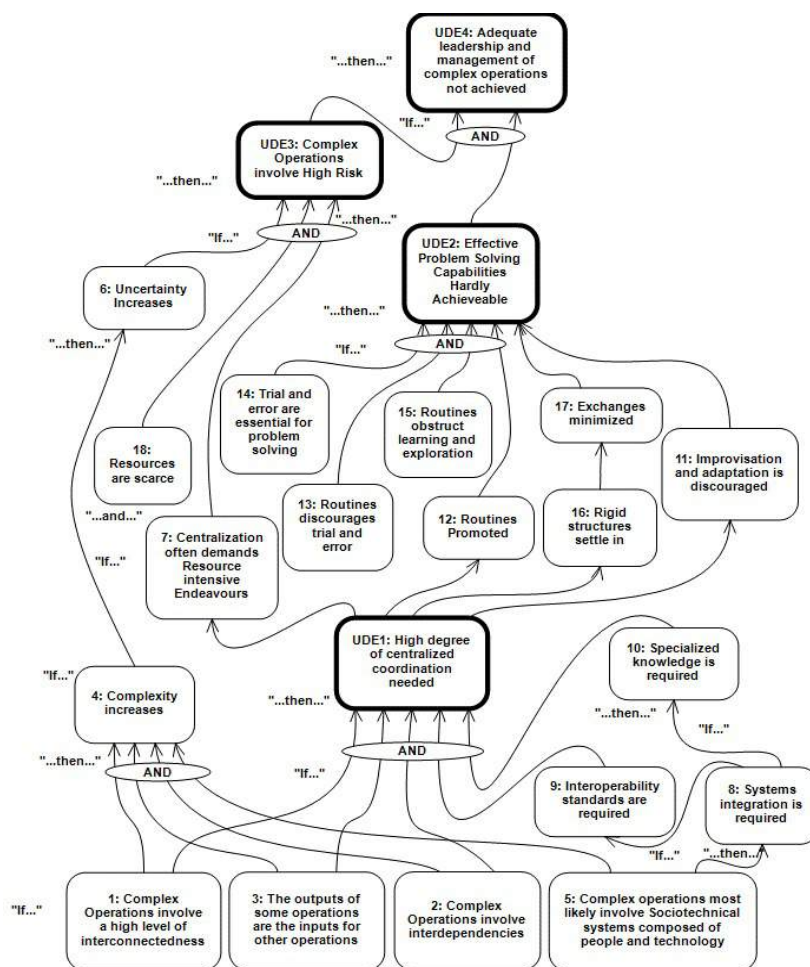
## RESULTS

The results from the methodological approach are shown as logic trees. The Current Reality Tree captures the analysis and diagnosis of the current paradigm. The solution is represented by another logical tree, the Future

Reality Three. Such logic trees were designed by following robust logical validation and using the *Categories of Legitimate Reservation* (Dettmer, 2003), a TOC tool to help validate logical relationships.

**Current Reality Tree, or Designing the Problem**

The Current Reality Tree (CRT) is a TOC logic tree which helps in the analysis and diagnosis of the problem at hand, identifying the UnDesirable Effects (UDE) within such tree and linking them down to the originating root causes. When considering the complex operations paradigm and the need to identify why it is so difficult to understand, manage and lead within such a paradigm, a CRT is quite helpful in showing the causal relationships among critical concepts (see Figure 1).



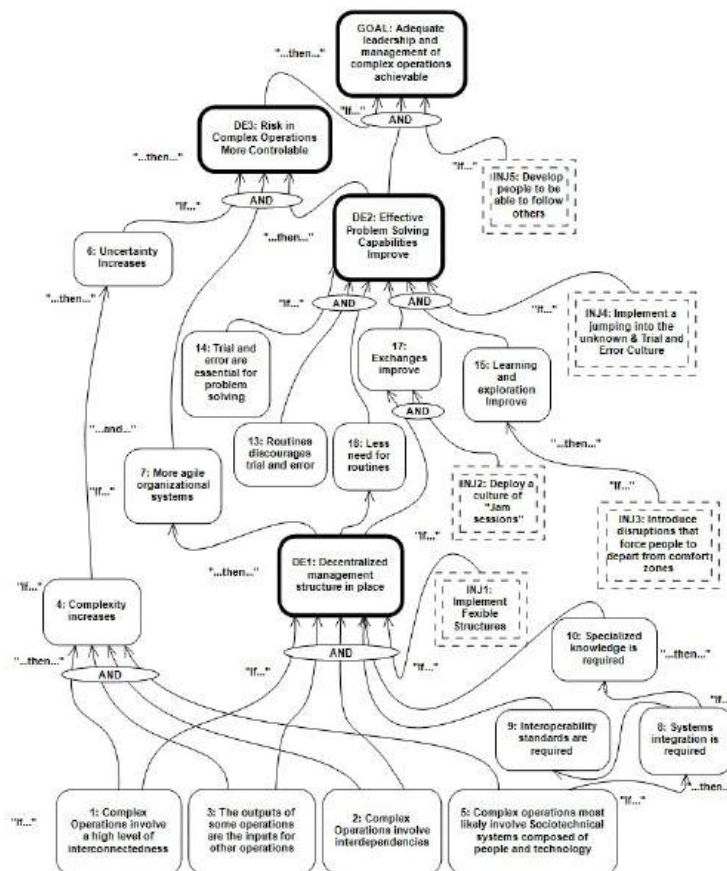
**Figure 1:** CRT - defining the problem links the UDEs to root causes (author’s).

The CRT is a TOC logic tree which shall be read from the bottom up as “if A and B ... then C”. From the CRT depicted, one can realise the undesirable effects resulting from causes such as (UDE1) *High degree of centralised coordination needed*, which promotes routines, and resources, rigid structures, while improvisation and adaptation are discouraged;

(UDE2) *Effective problem-solving capabilities hardly achievable*; (UDE3) *Complex operations involve high risk*, which prevents organisations from achieving an adequate level of leadership and management of complex operations; (UDE4) *Adequate leadership and management of complex operations not achieved*, which prevents achieving the organisation goal – the *adequate leadership and management of complex operations*. The chains of cause-and-effect down to the root causes are responsible for the undesirable effects (UDE 1-4) in the CRT and must be addressed by solution design in another logic tree.

### Future Reality Tree, or Defining the Solution

One can design workable solutions by injecting measures inspired by the jazz organisation principles presented in the background section before. The TOC suggests three ways to generate actions to be injected into the CRT, which will reverse the UDEs (UnDesirable Effects): (1) Use of *Evaporating Clouds*, a TOC tool; (2) creating an injection action that eliminates or minimises an undesirable root cause, and (3) creating an injection which satisfies a critical success factor (Dettmer, 2003). Such actions are named injections INJ# in the FRT (see Figure 2).



**Figure 2:** FRT – designing the solution through injections INJ 1–5 (author's).

As depicted in the FRT, five injections are now considered, which reverse the previous UDEs. INJ 1-5. INJ1 (*Implement Flexible Structures*) is needed to force an organisational design that incorporates just enough constraints and coordination to allow diversity and variation to succeed. With flexible structures, organizations can encourage exploration and experimentation, leading to adaptation, innovation and, ultimately, improvement in complex operations. INJ2 (*Deploy a culture of “Jam Sessions”*), in an analogy with the jazz organisation, suggests that other organisations can do the same by encouraging exchange through opportunistic interaction and sharing of ideas and adequate questioning. INJ3 (*Introduce disruptions that force people to depart from comfort zones*) suggests that for effective problem solving, people should often leave their comfort zones and try unfamiliar and new actions. INJ4 (*Implement a jumping into the unknown & Trial and Error Culture*) fosters a culture of solving complex challenges, while one has no assurance that his/her action will succeed. This demands a certain level of ‘psychological safety’. Finally INJ5 (*Develop people to be able to follow others*), deserves some explanation as it cannot be taken at face value. This ability to follow others is related to the paradigm in jazz organisations where the mastery attained by each player (aka system’s element) is such that they can follow others as these change course during a play. Only when people and teams are so skilled and developed to such a high level is this achievable – as such skills and ability to follow others become ‘second nature’ to such people and organisations. Due to the effect of these injections, not only do we achieve the elimination or mitigation of the previous UDEs, but we also foster the emergence of Desirable Effects DE1 to 3 (DE 1-3) and the achievement of the system’s overall goal - *Adequate leadership and management of complex operations achievable*. The TOC assumes, and as a more detailed FRT would show, that sometimes feedback branches develop, which will reinforce the overall designed solution. For instance, in the previous FRT, it is easy to expect that the DE2 (Desirable Effect 2) will further reinforce the acceptance of the injections INJ1-4.

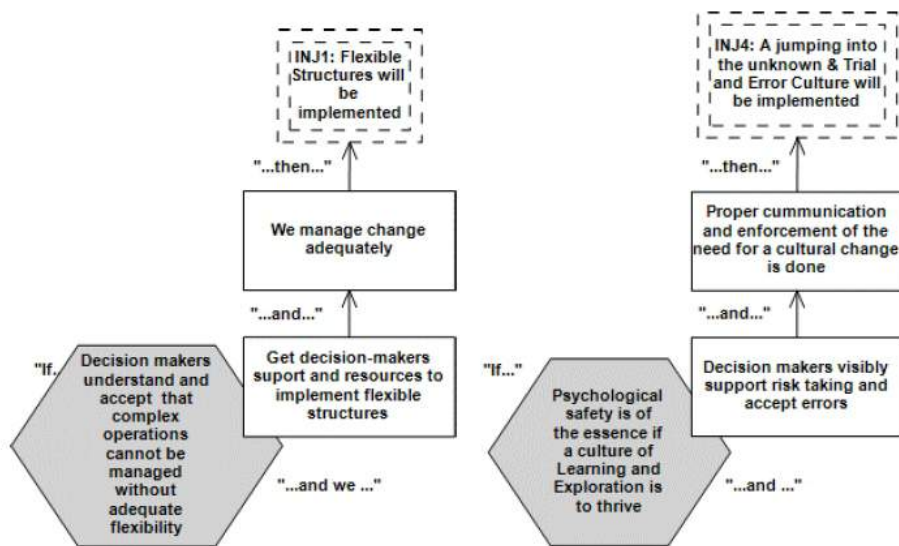
## DISCUSSION

This paper and the underlying study make use of logical validation by using the Categories of Legitimate Reservation (Dettmer, 2003; Dettmer, 2011), as only determinism, not correlation, can ensure adequate validation. The solution (FRT) is robust enough as a starting point for a more comprehensive research or particular research lines aimed at specific endeavours, industries or businesses. Once the analysis and diagnosis are complete using a clarifying CRT, and a solution is designed employing an FRT, implementation follows in order to deploy the solution while maximizing the chances of harnessing complex operations. Any deployment, however, needs some form of enabling conditions for its success.

### **Prerequisites Tree, or Ensuring the Critical Enablers Are Set in Place**

Too often, it is not the systems that fail but the ability to maintain them. If suitable organisational enablers are not set in place, no real and enduring

change will happen by just following the solutions depicted in the previous Future Reality Tree. The TOC provides an adequate type of tree for searching and clarifying the enabling factors that shall be in place for enduring change to happen – the *Prerequisites Tree* (PRT). The PRT is a set of branches made of chains of cause-and-effect influences that together ensure the system will perform as intended and sustainably. The PRT ensures that the designed injections will have the intended effects. As long as the “prerequisites” are in place, the system will perform as intended, maximising the performance within a complex operations paradigm (see Figure 3).



**Figure 3:** An example of a prerequisite tree to enable the solution in the FRT, exemplified with INJ1 and INJ4 (author's).

## CONCLUSION

Almost all domains are witnessing a rise in the complexity of their operations, and many are barely able to address it. Minimum acceptable performance ensures an organisation can survive. It cannot, however, excel; therefore, it is unlikely to be high-performing. The benefits of adopting some lessons from jazz are significant for businesses, governments, special services, and any organisation engaged in complex operations. Following a critical and logical thinking process, a pragmatic analysis of the complex operations paradigm was provided, and transformational actions were suggested. This approach to organisational design can help address the growing problems that traditional management and organisational schools of thought may be ill-equipped to solve. Jazz organisations derive strength from human bounds rather than human differences, thus a distinct type of organization where “leadership is granted by followers, not by organisational titles”. As suggested by DePree (1992, p. 106), “Just moving up in the hierarchy do not confer competence; so organisational power does not confer wisdom”. Leading complex operations

takes more than following static formulas from traditional management teaching – it needs sense-making, improvisation and adaptation, and the jazz organisation is perhaps the most promising one to help succeed in the realm of complex operations.

## ACKNOWLEDGMENT

The authors want to acknowledge CINAV and the Portuguese Naval Academy for their support.

## REFERENCES

- Austin, R. D., Stormer, C. (2008). Miles Davis: Kind of Blue. Harvard Business School Case 9-609-050.
- Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books, Inc., Publishers.
- Axelrod, R., Cohen, M. D. (1999). *Harnessing Complexity. Organisational Implications of a Scientific Frontier*. New York: The Free Press.
- Barrett, F. J. (1998) Coda—Creativity and Improvisation in Jazz and Organizations: Implications for Organizational Learning. *Organization Science* 9(5):605–622. <https://doi.org/10.1287/orsc.9.5.605>
- Barrett, F. J., & Peplowski, K. (1998). Minimal structures within a song: An analysis of “All of Me”. *Organization Science*, 9(5), pp. 558–560.
- Barrett, F. (2012). *Yes To The Mess. Surprising Leadership Lessons From Jazz*. Boston, Massachusetts: Harvard Business Review Press.
- Bernstein, E. S., and Barrett, F. J. (2011). Strategic change and the jazz mindset: Exploring practices that enhance dynamic capabilities for organizational improvisation. *Research in Organizational Change and Development*, 19, pp. 55–90.
- DePree, M. (1992). *Leadership Jazz*. New York: Dell Trade Paperback.
- Dettmer, H. W. (2003). *Strategic Navigation. A Systems Approach to Business Strategy*. Milwaukee, Wisconsin: ASQ Quality Press.
- Dettmer, H. W. (2021). *System Thinking*. Milton Keynes, UK: VBW Publishing.
- Garcia, J. M. (2004). *Sysware. La toma de decisiones empresariales en un mundo complejo*. Barcelona, Spain: Juan Martin Garcia Ed. (*in Spanish*).
- Goldratt, E. M. (1990). *Theory of Constraints*. Great Barrington, Massachusetts: North River Press.
- Goldratt, E. M., Cox, J. (1992). *The Goal*. New Haven, CT: North River Press.
- Groysberg, B., Masko, J. (2020). Lessons on Leading Through Chaos from U. S. Special Operations. *Harvard Business Review*. Nov.
- Hulm, V. (2019). Symphony or Jazz. *Mission-Planning Timelines. Military Review*. May-Jun. pp. 48–59.
- Mintzberg, H. (1979). *The Structuring of Organizations: A Synthesis of the Research*. United Kingdom: Prentice-Hall.
- Mintzberg, H. (1993). *Structures in Fives*. Englewood Cliffs, New Jersey: Prentice Hall, Inc.
- Rihani, S. (2002). *Complex Systems Theory and development Practice. Understanding Non-Linear Realities*. London, UK: ZED Books.
- Snowden, D., Boone, M. E. (2007). A Leader’s Framework for Decision Making. *Harvard Business Review*. November.

- Snowden, D., Rancati, A. (2022). *Managing complexity and chaos in times of crisis. A guide for decision makers inspired by the Cynefin framework*. Luxembourg: Publications Office of the European Union.
- Surowiecki, J. (2005). *The Wisdom of Crowds*. New York: Anchor Books.
- Tjan, A. K. (2010). Strategy as Jazz vs. Symphony. *Harvard Business Review*. Feb.
- Tucker, A. L., Edmondson, A. C. (2003). Why Hospitals Don't Learn from Failures: Organizational and psychological dynamics that inhibit system change. *California Management Review*. Winter 2003 | Vol. 45, No. 2.
- Valero y Vicente, A., Tomás, J. L. L. (1999). *Política de Empresa. El Gobierno de la Empresa de Negocios*. 4<sup>a</sup> Ed. Pamplona, Spain: EUNSA. (*In Spanish*).
- Weick, K. E., Sutcliffe, K. M. (2015). *Managing the Unexpected. Sustained Performance in a Complex World* (3<sup>rd</sup> Ed.). Hoboken, New Jersey: Wiley.

---

# Smart IoT soundproofing panels for enhanced environmental comfort

Luca Casarotto<sup>1</sup>, Lisa Battagliarin<sup>1</sup>, Filippo Carnovalini<sup>2</sup>, Marco Marigo<sup>3</sup>, Giulio Pitteri<sup>2</sup>, Mauro Longo<sup>4</sup>, Carlo Manara<sup>4</sup>, Jacopo Gonzato<sup>5</sup>, Andrea Padovan<sup>6</sup>, Antonino Di Bella<sup>3</sup>, and Antonio Rodà<sup>2</sup>

<sup>1</sup>Department of Architecture and Arts, University luav of Venice, Venice, ITALY

<sup>2</sup>Information Engineering Department, University of Padua, Padua, ITALY

<sup>3</sup>Department of Industrial Engineering, University of Padua, Padua, ITALY

<sup>4</sup>Ambiente 1985 & LvB Italia, Padua, ITALY

<sup>5</sup>Boccaccio432, Vicenza, ITALY

<sup>6</sup>BLR. System, Treviso, ITALY

## ABSTRACT

In the context of Ambient Assisted Living, the SISSI project aims to integrate soundproofing panels with IoT technologies, developing a new modular system to improve the comfort of people living and working in shared environments by taking advantage of the acoustic properties of the panels and the automatic monitoring of some relevant environmental variables. To reach a satisfying solution the need to experiment with different technologies arose, and thus the team needed to share observations, problems, and solutions right from the design phase. The first step was checking if perforating the company's current sound-absorbing panels and incorporating electronic components would affect their efficiency, measuring the material's performances in the laboratory to verify its sound absorption value.

Measurements were conducted in compliance with ISO 354 standards, assessing both types of sound-absorbing materials in various configurations. They show that arranging the material in a checkerboard pattern or in "thinned" lines is more advantageous in terms of absorption. Indeed, the laboratory tests have shown that a smaller amount of material results in better sound-absorbing characteristics. With this data, the final IoT panel will be more sustainable because less material can be used.

In the second step, new panels with integrated sensors to monitor temperature, humidity, CO<sub>2</sub>, brightness, and people's presence were tested. Based on the data collected, the team defined the functions of the panel, which will be able to monitor the level of air oxygenation and control room or desk brightness level and to switch on or off when the user is present, thus reducing energy consumption.

The experimental results changed the redesign of the panels, which now feature accessible electronic components and an interior that is not fully packed with material. Furthermore, the system is designed to integrate additional elements, such as lighting and electronic devices, for seamless interaction with the surrounding environment. Importantly, the sound-absorbing panel system will also provide clear user signals about the monitoring status, enhancing comfort.

**Keywords:** Soundproof panels, Smart furniture, Workplace comfort, Microcontroller programming, Psychoacoustics.

## INTRODUCTION



Ambient Assisted Living (Calvaresi et al., 2017) is a research field aimed at integrating new and existing technologies into people's life environments, supporting everyday living and promoting healthier and more sustainable lifestyles, and thus improving the overall quality of life of its users (Abtoy et al., 2020), with a particular attention to users with special needs (e.g. Zanolla et al., 2013; Mandanici et al., 2018).

In this context, the SISSI project aims to integrate soundproofing panels with IoT technologies, developing a new modular system to improve the comfort of people living and working in shared environments by taking advantage of the acoustic properties of the panels and the automatic monitoring of some relevant environmental variables.

The research project is being developed by a multidisciplinary team with several partners, including a manufacturing company, acoustic engineers, computer scientists and designers.

Making a system of sound-absorbing panels aimed at environmental monitoring led to the need to experiment with different technologies. Integrating electronic components for monitoring led to novel panel designs. These led to testing and experimenting with new functions and ways of using the sound-absorbing panels. The design phase followed multidisciplinary sharing dynamics based on digitized Industry 4.0 processes and on the principles of circular design. The panels are designed to be composed of parts, which can be easily disassembled, accessed, and repaired, recyclable or destined for correct recycling according to the Ellen MacArthur Foundation's butterfly diagram (Ellen MacArthur Foundation, 2020). Using the Jira Software platform by Atlassian, along with periodic meetings, enabled the team to share observations, problems, and solutions right from the design phase.

The team first tested if perforating the company's current sound-absorbing panels and incorporating electronic components would affect their efficiency. The sound-absorbing materials used were measured in the laboratory to verify the sound absorption value of the material.

Measurements were conducted in compliance with ISO 354 standards, assessing both types of sound-absorbing materials in various configurations. The laboratory results obtained were compared in terms of unit absorption change ( $\Delta A/n$ , m<sup>2</sup>). They show that arranging the material in a checkerboard pattern or in "thinned" lines is more advantageous in terms of absorption.

Thus, less sound-absorbing material inside the panel results in being more effective than a full arrangement of material. Indeed, the laboratory tests have shown that a smaller amount of material results in better sound-absorbing characteristics. With this data, the final IoT panel will be more sustainable because less material can be used. The data are interesting not only for the specific case but also because they open up new design possibilities. It will be possible to envision new solutions and assembly processes for panels that are more sound-absorbent.

In a second phase, new panels with integrated sensors to monitor temperature, humidity, CO<sub>2</sub>, brightness, and people's presence were tested. The interaction of

these data in the ThinkBoard platform enabled monitoring of the state of the environment and user well-being.

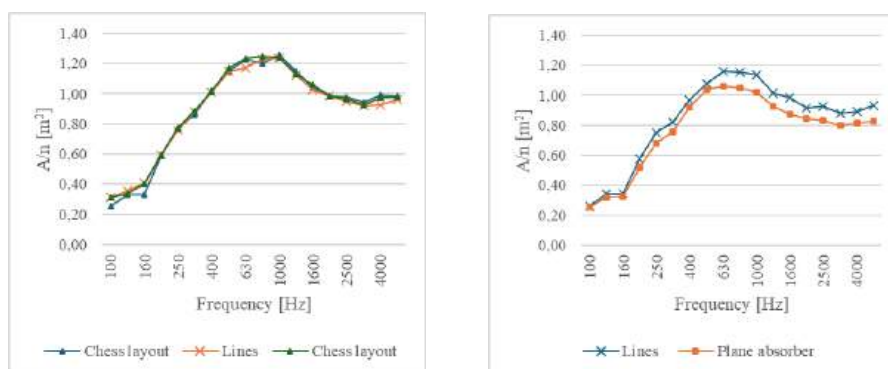
Based on the data collected, the team defined the functions of the panel, which will be able to monitor the level of air oxygenation and control room or desk brightness level and to switch on or off when the user is present, thus reducing energy consumption.

The experimental results changed the redesign of the panels, which now feature accessible electronic components and an interior that is not fully packed with material. Furthermore, the system is designed to integrate additional elements, such as lighting and electronic devices, for seamless interaction with the surrounding environment. Importantly, the sound-absorbing panel system will also provide clear user signals about the monitoring status, enhancing comfort.

## ACOUSTIC MEASUREMENTS

HSI experts contribute by ensuring that human capabilities and limitations are considered. It has become clear that treating the system as separate from the users results in poor performance and potential failure in the operational setting. Continued growth in technology has not delivered desired results. Systems engineers and others are beginning to understand the role humans play in technology systems. The core challenge is to balance successful hardware and software solutions with human friendly implementations. To define the requirements of humans as a fundamental system component, it is essential to understand the inherent capacity of user populations and their typical operational environment (Booher, 2003). A description of a population's capacity incorporates more than the basic anthropometrics or the cognitive capability of the average member of the user population (Chapanis, 1996).

### Evaluation of panel's material sound-absorbing capabilities



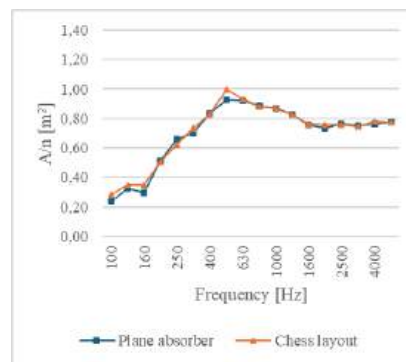
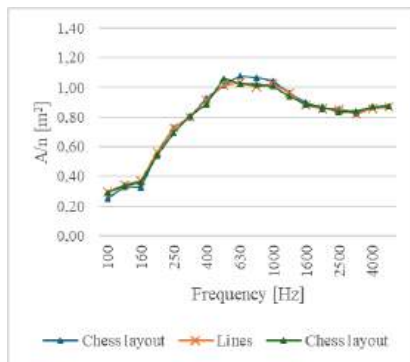
**Figure 1 (left):** Equivalent absorption area per panel  $A/n$  of elements under test with different arrangement modes (checkerboard layout and thinned lines) for 53.3% reverberation chamber occupancy compared to the maximum occupancy made.

**Figure 2 (right):** Equivalent absorption area per panel  $A/n$  of elements under test with different arrangement modes (thinned lines and plane absorber installation type "A") for 60% reverberation chamber occupancy compared to the maximum occupancy made.

The materials tested were polyester fiber panels, size 100x120 cm thickness 5 cm, nominal density 40 kg/m<sup>3</sup>.

The evaluation of their sound-absorbing properties involved measuring the equivalent sound absorption area according to different configurations of the material inside the reverberation chamber. Specifically, the materials were arranged according to the typical pattern of continuous plane absorbers (type "A," Appendix B of ISO 354), with checkerboard arrangement treating them as discrete objects and as "thinned" lines.

The materials were arranged in such a way as to create full and empty areas, so the treated surface is not completely covered by sound-absorbing elements but consists of a regular alternation of absorbing and reflecting elements. Figures 1-4 show the values for the equivalent absorption area divided by the number of panels tested (8, 9, 10, or 12), according to the arrangement as a continuous, checkerboard, and thinned-line surface based on the occupancy of the reverberation chamber relative to the maximum occupancy made.



**Figure 3 (left):** Equivalent absorption area per panel  $A/n$  of elements under test with different arrangement modes (checkerboard layout and thinned line) for 66.7% reverberation chamber occupancy compared to the maximum occupancy made.

**Figure 4 (right):** Equivalent absorption area per panel  $A/n$  of elements under test with different arrangement modes (matrix and plane absorber installation type "A") for 80% reverberation chamber occupancy compared to the maximum occupancy made.

## PANEL DESIGN

### Sustainable Panel Design

The design of the panels was focused on improving the sustainability of the products. To this end, the amount of materials used was reduced, as well as the processing and components employed. The project also aimed at reducing energy consumption in the application context and facilitated user-side monitoring and

improvement of application environments.

To optimize the design and production process in terms of sustainability, it was to transition from a linear to a circular design. It was therefore useful to refer to the Ellen MacArthur Foundation's butterfly diagram, which illustrates the continuous flow of materials in a circular process. This model divides the cycle into two parts: the first concerns materials and their technological transformation, while the second, on the right, directly involves design and production. Products and materials are kept in circulation through reuse, repair, regeneration, and recycling, with smaller cycles that allow preserving and acquiring maximum value from the product (Saidani, 2019).

To design a system of smart panels, it was necessary to ensure that the products were easily repairable, modular, with replaceable and/or upgradable components, and that these components were easily separable for recycling. Given the choice to use recycled polyester, one of the specific objectives was to verify whether reducing the amount of material used could still maintain optimal sound-absorbing quality (McDonough et al., 2002).

### **Sensor design and microcontroller choice**

In a first phase the research focused on the analysis and definition of the system architecture: specifically, attention was paid on the analysis of user's needs and comparison with similar existing systems; on the analysis of necessary technological requirements and the characteristics of materials to be used; and on the description of the logical level of the prototypes that are intended to be developed.

These initial analyses led to the definition of a more precise technology, chosen as a demonstrator of the project's goals: namely, desk panels. This type of sound-absorbing panels are furnishing accessories to be placed on office desks, as a divider in shared desks, or as a background element for desks placed against a wall.

In the initial phases of the project, through brainstorming activities that involved the entire research team, the project brief was defined. One driving factor for the design of these demonstrators was the technical requirement to be able to insert sensors necessary for environmental monitoring in the panel, and for potential access maintenance. For this goal, two approaches were explored. The first involved the creation of aluminum extrusions which would cover two or three sides of the panel and in which the fabric can be inserted. The missing sides would be then attached with a zip or a Velcro strip.

By doing this, if it were necessary to remove or access the internal components, it would not be a problem considering the design of the product made it reversible. Depending on the design of the extruded aluminum lateral closure and on how the fabric would be attached in the specific inlets, the panel would take different shape aesthetics, while maintaining the characteristics necessary to access the electronic components.

Once the structure was defined, the first prototype sound-absorbing panel was equipped with acoustic transducers (speakers), temperature sensor, humidity

sensor, light sensor, presence/motion sensor i.e. infrared (PIR), ultrasonic distance sensor, LCD character screen. This panel used a ESP32 processor, capable of internet connection.

The sensors were installed by creating square holes (10x10cm) in the panel, which were filled with cardboard structures to support the electronic components (see image 5). Everything was then enclosed in fabric to check the system's overall functionalities.



**Figure 5:** The panel used for the prototype for the first solution. On the left it's shown in full view with the housings for the electronic components and two example cardboard "modules". On the right, it's shown housed behind a desk where the functionalities of the various sensors were evaluated. During the tests, the facade was then covered with fabric to test the sensors' response in this condition.

It quickly emerged that some sensors, when covered by fabric, lose their functionality or become noticeably less precise. We considered keeping some sensors outside the fabric, or perforating the fabric over the sensors. Various ways to be able to transfer these solutions on large scale production were considered, but we finally considered these solutions to be impractical, and therefore the main structure of the prototype was changed.

From the initial solutions, which envisaged the integration of all parts, the design evolved towards a panel with the electronic components separated from the sound-absorbing part. A configuration was then chosen that positioned the technology on top, allowing the sound-absorbing element to be integrated with a lateral structure. This configuration made it possible to join or separate the panels, which already had a tested production cycle, and to concentrate the technological parts in an aluminum profile, facilitating access and disassembly of the parts (Braungart et al., 2007).

The panel therefore ended up very similar to those already made by the partner of SISSI research "Ambiente 1985", but sustained by a lateral structure that holds the components, keeping them on an "arch" above the sound-absorbing part. This arch will be an extrusion in aluminum made specifically to contain the electronic components (sensors and actuators), holding a continuous band of LEDs (along the entire length of the panel), housing the appropriate covers and obtaining specific useful lanes for inserting further accessories.



**Figure 6:** Prototype with sensors inserted into the ALL-LED extruded profile, with the addition of a LED strip for lighting.

In the prototyping phase, to reduce prototyping costs, an already existing commercially available extrusion was used and therefore not customized. In this new iteration of the prototype, a LED strip has been added to the sensors allowing desk-mounted panels to also provide lighting. Compared to the first prototype, a CO<sub>2</sub> sensor and a microphone used for ambient noise detection were also added. The resulting prototype is shown in image 6.

The following table summarizes the main electronic components chosen for the prototypes, excluding resistors, switches, connection cables, and other minor components.

**Table 1.** Sample human systems integration test parameters (Folds et al. 2008)

Electronic Component	Category	Purpose
ESP32-WROOM-32E	Microcontroller	Control of sensors, WiFi/Internet communication, actuator management
DHT11	Sensor	Temperature and humidity
SCD41	Sensor	CO <sub>2</sub>
MSM26S4030H0	Microphone	Environmental noise control
HC-SR04	Sensor	Distance
AS312	Sensor	Infrared presence (PIR)
VCNL4200	Sensor	Luminosity
MAX98357A	Amplifier	Sound emission through speakers

Subsequently, three more panels with the same characteristics were produced, to be used for demonstrating purposes and to carry out environmental comfort experiments.

## RESULTS

To carry on environmental comfort experiments the four built panels were placed inside a testing laboratory room equipped with ambient sensors, together with the capability to change and control light, temperature, humidity and CO<sub>2</sub>. The room setup was done as follows: four desks were placed one in front of the other with two desks per side, and each desk was equipped with a workstation and the panel sitting behind it at the bottom edge of the desk. The laboratory room with the four panels in place is shown in figure 7.

**Figure 7:** Photo of the panels placed inside the environmental testing laboratory room.

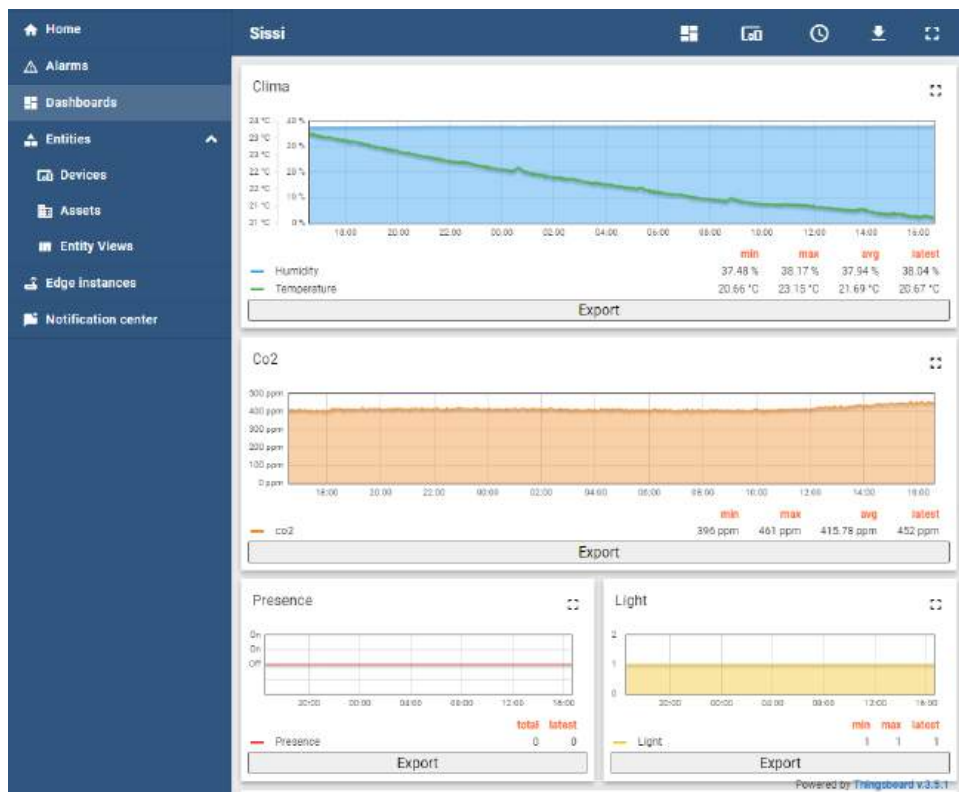


The panels have been placed and turned on for a week and they sent their data continuously to the ThinkBoard platform (see figure 8) and their measurements were compared with the one gathered from the laboratory room instruments. Overall a total of about 50 Mb of data were collected without noticeable data loss. Regarding the microcontroller resources usage it is worth noting that with the ESP32 chip the code nearly uses all the memory available, in particular considering the FFT code used to evaluate the environmental noise was already heavily optimized both regarding memory usage and computational time. Moreover, the masquerading noise was encoded using a plain PCM scheme, thus avoiding computing-heavy encodings such as mp3, considering the limited cpu frequency available to the microcontroller.

## CONCLUSIONS

The study of the equivalent sound absorption area of the materials tested showed that, with the same materials used, the achievable differences depend on the way the materials are installed. It was observed that checkerboard or in thinned line arrangements, i.e., installations that have fewer panels than the total area that can be occupied by a continuous arrangement, are more advantageous. It can therefore be effective to use a small number of "isolated" panels compared to an extended arrangement in this way it is possible to reduce the amount of sound-absorbing material to have equal sound-absorbing quality. This is also an excellent result from the standpoint of product sustainability which, from a design perspective, can thus reduce its environmental impact while maintaining technical performance. Tests conducted in the laboratory have shown that the panels are capable of correctly measuring and sending to the dashboard the main environmental telemetry values.

One consideration regarding possible future improvements may arise regarding the capability of the panels to communicate between themselves without requiring an existing WiFi infrastructure (router), to which extent an already existing technology such as "Data over Sound" may be employed (Pitteri et al., 2021).



**Figure 8:** Example screenshot of the ThinkBoard dashboard used to collect sensor data from the panels.



Another aspect may be the extension of the panels functionalities to include work stress related aspects, in a way that the user may be alerted regarding the need for a pause, or even helping the user's stress decompression by means of a mobile app linked to the panel's facade buttons, utilizing a gamification approach to increase the stress avoidance efficiency (Pitteri et al., 2019).

## ACKNOWLEDGMENT

This research was carried out within the project "Sistemi modulari Inclusivi per la gestione del comfort ambientale attraverso lo Sviluppo di pannelli fonoassorbenti Sensorizzati e Interattivi (SISSI)" (Modular inclusive systems for the management of environmental comfort through the design of interactive sensorized sound-absorbing panels). funded by Regione Veneto on the POC 2014-2020 program. Project ID: 10449164.

## REFERENCES

- Abtoy, A., Touhafi, A., & Tahiri, A. (2020). Ambient Assisted living system's models and architectures: A survey of the state of the art. *Journal of King Saud University-Computer and Information Sciences*, 32(1), 1-10.
- Braungart, M., McDonough, W., & Bollinger, A. (2007). Cradle-to-Cradle Design: Creating Healthy Emissions – A Strategy for Eco-Effective Product and System Design. *Journal of Cleaner Production*, 15(13-14), 1337-1348.
- Calvaresi, D., Cesarini, D., Sernani, P., Marinoni, M., Dragoni, A. F., & Sturm, A. (2017). Exploring the ambient assisted living domain: a systematic review. *Journal of Ambient Intelligence and Humanized Computing*, 8, 239-257.
- Ellen MacArthur Foundation. (2020). Circular Economy Diagram: The Butterfly Diagram. Ellen MacArthur Foundation.
- ISO. Acoustics — Measurement of sound absorption in a reverberation room, Standard 354:2003, International Organization for Standardization, 2003.
- Mandanici, M., Altieri, F., Rodà, A., & Canazza, S. (2018). Inclusive sound and music serious games in a large-scale responsive environment. *British Journal of Educational Technology*, 49(4), 620-635.
- McDonough, W., & Braungart, M. (2002). *Cradle to Cradle: Remaking the Way We Make Things*. North Point Press.
- Pitteri, G., Micheloni, E., Fantozzi, C., & Orio, N. (2021). Listen by Looking: A framework to support the development of serious games for live music. *Entertainment Computing*, 37, 100394.
- Pitteri, G., Micheloni, E., Rodà, A., Fantozzi, C., & Orio, N. (2019, September). Listen By Looking: a mobile application for augmented fruition of live music and interactive learning. In *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good* (pp. 136-141).
- Saidani, M., Yannou, B., Leroy, Y., Cluzel, F., Kendall, F. (2019). A taxonomy of circular economy indicators. In *Journal of Cleaner Production*, vol. 207 (pp. 542-559).
- Zanolla, S., Canazza, S., Rodà, A., Camurri, A., & Volpe, G. (2013). Entertaining listening by means of the stanza logo-motoria: an interactive multimodal environment. *Entertainment Computing*, 4(3), 213-220.

# Document Sharing Without Internet Connectivity During Study Abroad Programs

Saumick Pradhan, Nathan Wick, and Cedrick A. K. Kwuimy

University of Cincinnati, Cincinnati, OH 45221-0030, U.S.A.

## ABSTRACT

With the increasing demand for study abroad trips, students and faculty require a streamlined platform to manage assignments and surveys during the trip. It is common for students studying abroad to lack a reliable internet connection, making it challenging to share and access documents efficiently. This is especially true in countries with underdeveloped areas with poor connectivity infrastructure. Our mobile application, UC Transform (originally “Woforo”, linked below) addresses this issue by offering a one-stop-shop platform for accessing all assignments without requiring an active internet connection. Teachers can build courses, create assignments, and track progress using visual graphs. Students simply join with a code and retrieve their course materials online. Once connected to the internet, they can download the assignments and store them locally on their device. Throughout the trip, students can access the materials and provide their responses. Upon regaining a stable internet connection, students can upload their answers for review by professors. We utilized the Flutter and Dart framework for the User Interface deployed on iOS, Android, and Web Platforms, while leveraging Google Firebase for authentication and data storage. In this paper, we introduce the challenges faced by the students and educators, our methodology/ process of designing and developing UC Transform Application as a solution and sample case study exploring the scenario.

**Keywords:** Study abroad, Application, Students, Assignments, Software, Internet connectivity, Educators, Document sharing, Offline access

## INTRODUCTION

In the context of study abroad programs, students and faculty often face significant challenges in managing and sharing academic materials due to unreliable internet connectivity. Our focus is on locations which are underdeveloped in the infrastructure for connectivity. Students and faculty travelling to these remote places for research, field work, social welfare, and educational programs. They face challenges in document sharing.

According to a Michigan Public Policy Survey (Rubin, 2021), “these places with disparities between urban and rural areas based on broadband access, which can lead to a “rural penalty,” a term to describe the increased economic and social burden that rural communities face as a result of their distance from seats of power and centres of commerce.” These challenges are many and impact both students and educators:

1. **Connectivity Issues:** Talking especially about remote locations, Africa, where the internet penetration rate is 43%). These study-abroad locations frequently have inconsistent or limited internet access, making it difficult for students to download, upload, or even view necessary documents and assignments (Sabzaliev, 2021).
2. **Inefficient Communication:** Traditional methods of sharing documents via email or cloud storage require a stable internet connection, which is often unavailable or unreliable during international travel. “For example, in Tajikistan, only 3% of the population have a home broadband connection and most students are reliant on mobile data, which costs more than in any other ex-Soviet country”, according to University World News (Sabzaliev, 2021).
3. **Course Management Difficulties:** Educators struggle to efficiently distribute assignments, collect responses, and track student progress without a centralized and reliable system that functions offline.
4. **Student Accessibility:** Students need a dependable way to access course materials and complete assignments without relying on continuous internet access. Interruptions in connectivity can hinder their ability to meet academic deadlines and maintain consistent communication with instructors (Workman, 2013).
5. **Data Security and Synchronization:** Ensuring that students’ work is securely stored and synchronized when an internet connection becomes available is crucial. Without a reliable system, there is a risk of data loss or duplication, leading to confusion and additional workload for both students and educators.

UC Transform was conceived to address these issues by providing a robust platform that ensures seamless document sharing and course management, even in the absence of an active internet connection.

## METHODOLOGY AND RESULTS

### A) Requirements Analysis

- Conducted interviews with faculty and students to identify key challenges faced during study abroad trips.
- Analysed existing platforms to determine essential features and user preferences.

### B) Design

- Created wireframes and prototypes to visualize the user interface and user experience.
- Ensured designs were intuitive and catered to the needs of both students and faculty.

### C) Implementation

- Utilized the Flutter framework (Flutter, 2017) to develop a cross-platform application compatible with iOS, Android, and Web

platforms. We used the Dart Programming Language (Bracha, Gilad, 2016).

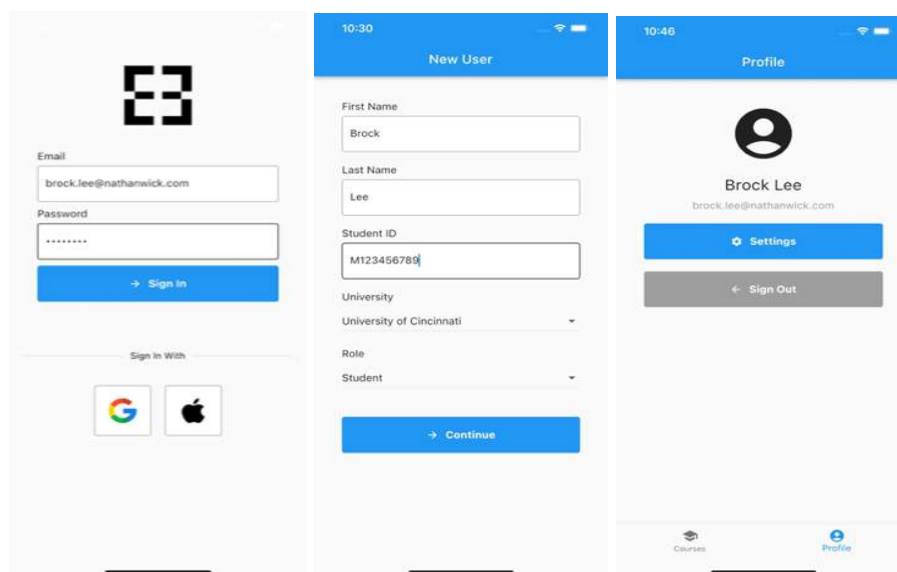
- Implemented features such as course creation, assignment management, and progress tracking for teachers.
- Developed functionalities allowing students to join courses using code, download assignments, and store them locally on their devices.

#### D) Backend Integration

- Integrated Google Firebase for authentication and real-time data storage (Google, 2011).
- Implemented offline data storage capabilities, allowing students to access downloaded materials without an internet connection.  
Ensured synchronization of data upon regaining internet connectivity, enabling students to upload their responses for review.

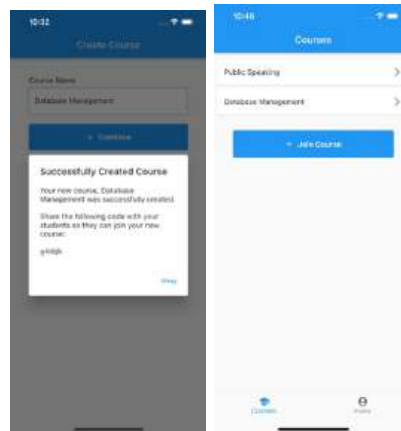
#### E) Testing

- Conducted unit tests to ensure individual components functioned correctly.
- Performed integration testing to verify the interaction between different modules.
- Executed user acceptance testing with a small group of faculties and students to gather feedback and make necessary adjustments.



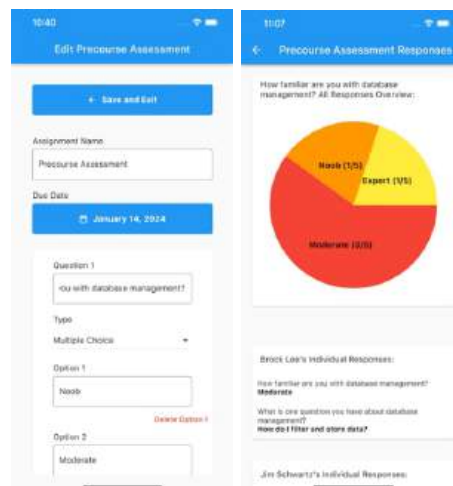
**Figure 1:** User authentication interface.

Users can sign in with their email and password, Google, or Apple account. New Users are prompted to enter essential information such as their name (see Figure 1).



**Figure 2:** Course creation interface.

Faculty may create a course to share with students. Students can view the courses that they're part of or join a course by entering the course code (see Figure 2).



**Figure 3:** Assignments and graphs interface.

Faculty may create assignments for the students in a course, then view the assignment responses with graphs (see Figure 3). Students can download assignments for a course, complete them while offline, and automatically sync their responses when they're back online (Flchart.dev, 2024).

The Flutter app uses Google Firebase's persistence feature (Firebase, 2022) to automatically cache information such as courses, assignments, etc., from a central database to the user's local device's storage (see Figure 4). As the user uses the app offline, their changes are recorded to their local device's storage. Once a user goes back online, their changes to their local storage are synced with a central database, which the faculty uses to view the student's responses.

```
if (kIsWeb) {
  await FirebaseFirestore.instance
    .enablePersistence(const PersistenceSettings(synchronizeTabs: true));
} else if (Platform.isAndroid || Platform.isIOS) {
  const databaseSettings = Settings(
    persistenceEnabled: true,
    cacheSizeBytes: Settings.CACHE_SIZE_UNLIMITED,
  );
  FirebaseFirestore.instance.settings = databaseSettings;
}
```

Figure 4: Cache settings.

## SAMPLE CASE STUDY AND USER STORY

### 1. Sample Background

The University of Cincinnati offers a study abroad program to Ghana for a group of 30 students. The course, titled “Cultural Studies in Ghana,” requires students to complete various assignments and surveys throughout the trip. Historically, managing these tasks has been challenging due to unreliable internet connectivity in some locations. To address this, we decided to implement UC Transform, an app designed to facilitate offline document sharing and course management, providing a seamless academic experience regardless of internet availability.

### 2. Pre-Trip Preparation

Before departing, the course instructor introduces UC Transform to the students. The instructor explains that the app will streamline their coursework by allowing offline access to assignments and enabling efficient communication.

#### a. Account Creation and Initial Setup:

- (i) **Signing In:** Students download UC Transform and are prompted to sign in using their Google accounts. This provides a quick and secure way to authenticate users.
- (ii) **Profile Setup:** On first sign-in, students enter their first and last names, select their role as ‘Student,’ choose ‘University of Cincinnati’ from a dropdown menu, and input their student ID numbers.

#### b. Course Enrollment:

- (i) **Course Code Entry:** The instructor creates the course “Cultural Studies in Ghana” on UC Transform, generating a unique course code. This code is then shared with the students.
- (ii) **Joining the Course:** Students use the course code to enroll in the class within the app, ensuring that only those registered for the study abroad program can access the course materials.

### 3. Assignment Distribution and Completion

- a. **Creating Forms:** During the trip, the instructor creates various forms and assignments through UC Transform. These include

multiple-choice quizzes, short essay questions, and file upload tasks (images of cultural landmarks, video diaries, and documents summarizing their experiences).

- b. **Offline Access and Data Caching:** Students see a list of all their incomplete forms ordered by due date. They download these forms to complete offline, leveraging the app's ability to cache data locally on their devices. This is crucial given the spotty internet connectivity in some areas they visit.
- c. **Submission of Work:** Upon reconnecting to the internet at their hotel or a café, students upload their completed assignments to UC Transform for the instructor to review. The app seamlessly synchronizes cached data with the cloud, ensuring that no work is lost, and all progress is tracked accurately.

#### 4. Data Visualization and Feedback

- a. **Tracking Progress:** The instructor uses UC Transform's data visualization tools to monitor student submissions and performance. They view summary graphs (histograms, pie charts, etc.) that highlight how students are engaging with the course materials. These visualizations are crucial for quickly assessing overall class performance and identifying trends or areas needing attention.
- b. **Providing Feedback:** The instructor downloads tables of student responses for more detailed analysis and provides feedback directly through the app.

#### 5. Additional Features and Security

- a. **Multiple Authentication Methods:** To ensure flexibility, UC Transform supports sign-in via Apple and Microsoft accounts, catering to students with different device preferences.
- b. **Secure Data Storage:** All data is securely stored in Firebase, ensuring that student information and submitted assignments are protected both locally (cached) and in the cloud.

#### 6. Potential Benefits

##### a. Increased Accessibility and Efficiency

UC Transform significantly improves the efficiency of managing assignments during the study abroad trip. Students can easily access and complete their coursework offline, avoiding disruptions caused by unreliable internet connections. This functionality, supported by local data caching, leads to timely submissions and better overall academic performance.

##### b. Enhanced Student Experience

Students appreciate the reliability of UC Transform. The ability to download and work on assignments offline meant they could focus on their studies without constant concern about finding

an internet connection. This leads to higher engagement and satisfaction with the course.

c. **Faculty Feedback and Engagement**

The instructors find the app's data visualization tools invaluable for tracking student progress and identifying areas where additional support is needed. The ease of creating and managing assignments allows the instructor to focus more on teaching and less on administrative tasks, enhancing the overall educational experience.

7. **Conclusion for the Case Study**

UC Transform proved to be a transformative tool for the University of Cincinnati's study abroad program in Ghana. By addressing the challenges of unreliable internet connectivity and providing robust features for course management, offline access, data caching, and data visualization, the app significantly enhanced the academic experience for both students and faculty. The success of this implementation suggests that other institutions with similar challenges could benefit from adopting UC Transform.

## **FUTURE WORK**

This research and product development exercise successfully addressed the significant challenge of managing academic materials during study abroad trips where reliable internet connectivity is often unavailable. By developing the UC Transform mobile application, we have created a robust and streamlined platform that enhances both teaching and learning experiences for faculty and students participating in study abroad programs. Key accomplishments of this project include Offline Access and Data Caching, Real-time Synchronization, User-friendly Interface, and Enhanced Course Management. We also showed a sample case study to explain the working of the product and its impact.

### **Some Areas for Future Development and Improvement**

1. **Scalability and Performance Optimization:** Enhance the app's scalability to support larger numbers of users and more extensive data storage requirements
2. **Advanced Data Analytics:** More advanced data analytics tools for educators to gain deeper insights into student performance and engagement, including predictive analytics.
3. **Integration with Educational Platforms:** Establish integration with other popular educational platforms and learning management systems (LMS) to provide a more comprehensive academic ecosystem.

The Application can be tested here: <https://woforo-app-reu.web.app/#/>.



## AUTHOR CONTRIBUTIONS

Conceptualization, C.K, S.P and N.W.; Solution Design, C.K, S.P and N.W.; Application Development, N.W and S.P.; Application testing, S.P and N.W.; Data Visualization, N.T.; Results, S.P. and N.T.; Writing-Original Draft Preparation, S.P, C.K and N.W.; All authors have read and agreed to the published version of the manuscript.

## ACKNOWLEDGEMENT

This work is supported by a National Science Foundation Award NSF EEC-2106229-001.

## REFERENCES

- Bracha, Gilad (2016), "The Dart Programming Language." Addison-Wesley.
- Firebase (2022), "Authentication State Persistence | Firebase.", Google, Website: [firebase.google.com/docs/auth/web/auth-state-persistence](https://firebase.google.com/docs/auth/web/auth-state-persistence).
- Flchart.dev (2024), "Fl\_chart: Flutter Package." *Dart Packages*, Website: [pub.dev/packages/fl\\_chart](https://pub.dev/packages/fl_chart).
- Flutter (2017), "Build Apps for Any Screen.", Website: [flutter.dev/](https://flutter.dev/).
- Google (2011), Google, Website: [firebase.google.com/](https://firebase.google.com/).
- Rubin, Julie (2021), "Internet Access and Connectivity across the Urban-Rural Spectrum." *Center for Local, State, and Urban Policy*.
- Sabzaliev, Emma, et al. (2021), "The Impact of the Digital Divide on Student Mobility." *University World News*.
- Workman, Mike (2013), "Technology Issues for Research in Remote Areas/Developing Regions." Global Health Training Centre.

# Integrating Space Syntax Methods in Building Environmental Simulations and Urban Studies Pedagogy to Enhance Designers' Critical Reflective Practice

Deborah A. Middleton

Department of Architecture, Xi'an Jiaotong-Liverpool University, Suzhou  
215000, China

## ABSTRACT

The aim of the study assesses how architecture students reflectively evolve their procedural (tacit) knowledge, iterate innovative design strategies, and link simulation modeling to develop an architectural understanding synthesis. A model of Reflective Synthesis Design Cognition (RSDC) is developed in order to extend Donald Schön's seeing-moving-seeing model of reflection in action with material of a design situation. Two design pedagogy approaches were trialed to test the validity of the RSDC model. Course assignments required students to pilot and test design strategies while integrating different types of simulation modeling tools commonly used in building environmental performance and urban studies. The first trial required young designers to adapt Space Syntax methods with a variety of simulation modeling approaches to discover and transform the environmental performance of their architecture studio project. A second trial required students to investigate the urban spatial morphology surrounding their project site, through qualitative analysis of spatial representations that also included Space Syntax simulation models. Space syntax methods were an unfamiliar methodology that required students to 'see-move-see' reflectively in the design situation. Research methods use in this study include, content analysis of the simulation modeling outcomes, associative linking of formative application of simulation modeling to design strategy, and use of a digital survey questionnaire to explore user experience, and self-identification of psychometric processes. Findings highlight the need for more descriptive studies examining referential reflection, qualitative assessment reflection, and summative versus synthesis constructivist knowledge formation.

**Keywords:** Human design cognition, Environmental building systems modeling, Space syntax, Qualitative reflection

## INTRODUCTION

Design simulations result from the interaction of variables, from spatial form, configuration and orientation, to tactile materials, and variables that provide contextual framing such as environmental, temporal, behavioural, social organization. In contemporary design practice, digital simulations combine and form interactive relationships from single to multiple parameters, that are

descriptive as 2D and 3D digital models, that may also contain information as output relational data (i.e. quantity of solar irradiance exposure on a building façade). Using computer aided simulation techniques, designers engage in mental processes that frame design tasks to solve a design problem, establish logical sequences of design moves - as steps in interactive informational processing with a simulation.

To advance the simulation phase into concrete design phase, designers must engage in reference and reflection with the designed object, in order to achieve understanding. This process engages enhancing students ability to distinguish and isolates variables through their per-formative interaction and apply reflective qualitative judgment to link simulation outcomes to design goals.

This paper provides evidence of how designers engage in relational and reflective conversations with new design visualizations to generate a summative synthesis, as knowledge and understanding. In order to gain an understanding of different cognitive process, research trials introduced students to a previously unknown design simulation methodology, that generates new visual representation. This study illustrates how designers define design problems, and apply a combination of simulation types to model and adjust a design strategy and proposal.

In the next section, I presents a brief literature review on previous insights into design cognition and iteration. Donald Schön's model of seeing-moving-seeing is then adapted as a new cognitive schema, that engages constructs of reference, analysis, and reflection. The research approach and design output samples from two different trials is presented in Section discusses the results and presents highlights of the interaction between information processing and design strategy formulation.

Insights from an on-line survey deepen understanding of the cognitive processes related to reflection, reference and summative synthesis. The conclusion poses new questions for design cognition research and possible ways to investigate these further.

## **LITERATURE REVIEW**

### **Design Cognition Research**

Early research into design cognition focused on the structure of the design task from the perspective of protocol analysis, to examine problem-solving behaviour and mental information processing (Ericson & Simon, 1984). Design task analysis has also categorized cognitive behaviour as seeking patterns, encoding meaning and developing classification and ordering systems (Atkins, 1978; Eastman, 1969). Studies of design work flow have identified the inherent challenge facing a designer to scope the design context, characterize an ill-defined design problem, define a strategic approach, and concept, and apply the designers experience and tacit knowledge, to organize different sub-tasks (Goel et al., 1996). Dorst (1997) found that designers rely heavily on the external representations involved in order to develop an awareness of what the designer is doing. Schön's model of knowing-in-action with materials of a design situation, suggests that a designers' process is

revealed in and by doing actual designing, when they are in transaction with a design situation which in turn they create (Schön, 1992). Schön further suggests that through engagement with design processes, a designer interacts with intermediate design representations, that may change their view of the current design as a result of them generating and interpreting representations of the design (Schön, 1992). This research seeks to gain further insights into the design psychometric process.

### **The Role of Iteration in Design Cognitive Processes**

The design learning process is also iterative. Designers reflectively study design outcomes, engage in constructing referential systems of meaning, respond, react, plan to modify, and adjust, a designed object or its spatial relationships, in a cyclical way. Iteration may occur at multiple phases in the design process, from innovative concept development, to the framing and decomposition of complex problems, or in the need to integrate design processes (Eppinger et al., 1994). Feedback loops, as a stage-based design development process, are used by designers to drive innovation within a project, and also generate learning across consecutive or concurrent projects (Pahl & Beitz, 1996; Kline, 1985; Eppinger et al., 1994). Exploration and convergence processes, also apply iteration, to enable alternatives to be repeatedly generated, analysed, and evaluated (Wynn & Clarkson, 2005). Schön & Wiggins (1992) viewed this to involve a form of sequential development interplay in design problem and concept formation. The co-evolution of the problem and solution spaces (Dorst and Cross, 2001) may be defined in relation to time duration, scale, context and scope of the design problem, and aspects of value-added versus corrective action, (Wynn and Clarkson, 2016) align to the reframing of the design problem (Kolodner and Wills, 1996). In these cases, qualitative reference and reflective processes may be engaged by the designer.

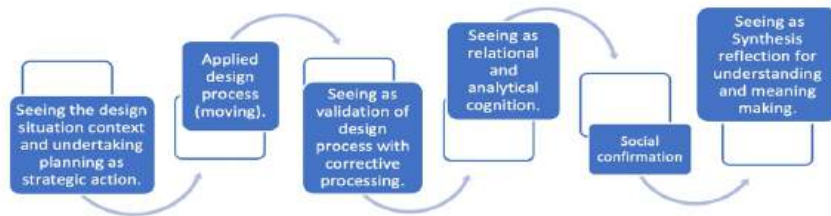
### **Design Appropriate Learning Models and Theories**

Piaget's cognitive development model, and Kolb's learning styles, are two learning theories suited to the pedagogy of architecture and design in their consideration of learning as an internal mental process. Teaching from this perspective engages students to develop cognitive abilities to problem-solve, organize information, and develop purposeful self-directed learning practices. Students apply their knowledge and focus on their internal learning as cognitive process. Using reflective observation learners transition their experiential learning into knowledge, which then can be flexibly applied to a range of new situations (Kolb, 1984). Piaget in his theory of cognitive development posits that 'through the construction, expansion, and integration of mental structures, the individual not only acts (mentally and behaviorally) following the principles of logic, but also becomes more and more able to accurately predict future events (Piaget, 1971 in Kazi & Galanaki, 2020).

## REFLECTIVE SYTHESIS DESIGN COGNTION MODEL

In this paper I draw attention to the above processes, and present a new model with a focus on information processing, reflection, reference and synthesis. The RSDC framework, expands on Schön's reflection in action model, by placing emphasis on psychometric processes that aid the development of design strategies. Additionally the approach taken in this study, integrates many of the above constructs to examine how student designers engage in experimental prototyping, order and classification in information processing, and engage in iterative and variable processes of analytical reasoning, reflection and reference.

Specifying the sequential process steps and linkage between cognitive phases is not addressed in this study, rather the research focus is on classifying user reference and reflection psychometric processes.



**Figure 1:** Design cognitive reflective synthesis learning model adapted from Schön seeing-moving-seeing (1992) framework of design cognition.

## METHOD APPROACH

In order to examine the appropriateness of this model to understand design cognition and learning processes, two undergraduate architecture modules, ARC201 Environmental Performance of Buildings, and ARC206 Urban Studies, were selected for an intervention that introduces new analytical simulation modeling approaches. The pedagogy applied in the two trials, places emphasis on reflection with simulation data representations, in order to test and develop sustainable design strategies for the site and building.

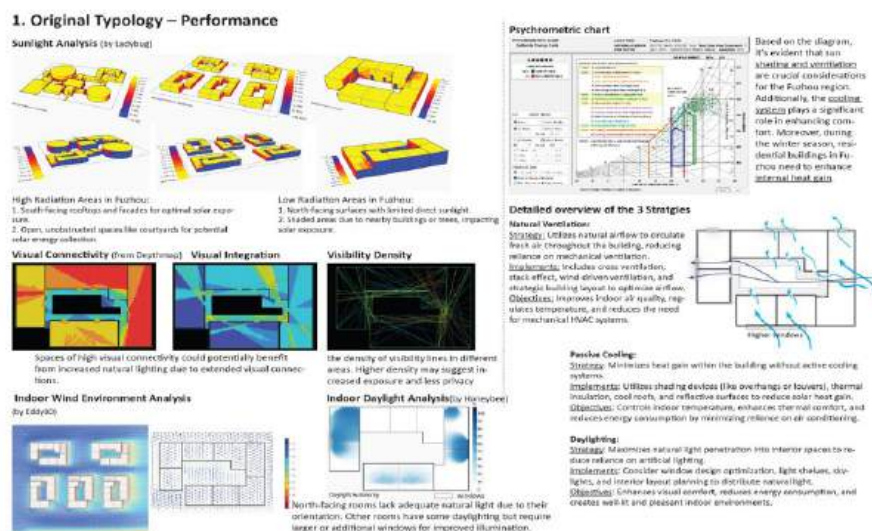
To minimize the influence of previous knowledge feedback loops in the cognitive process, a new previously unknown digital simulation tool depthmapX and Space Syntax analytical methods were used to augment environmental and urban studies analytical approaches. Students did not have prior experience, or knowledge of how to use space syntax theory or depthmapX, and were challenged to find an appropriate way to integrate this methodology into their workflow to pilot a design problem and environmental performance simulation strategy. The assignment structure sought to integrate psychometric processes from contextual assessment (seeing) strategy formation, memory, (moving) information processing sequencing of steps (moving), analytical and qualitative reasoning to modify simulation output and methodology (seeing), generate new, or engage in

iterative testing of design strategy (moving), and engage in reflection (seeing) with aspects of qualitative value judgment.

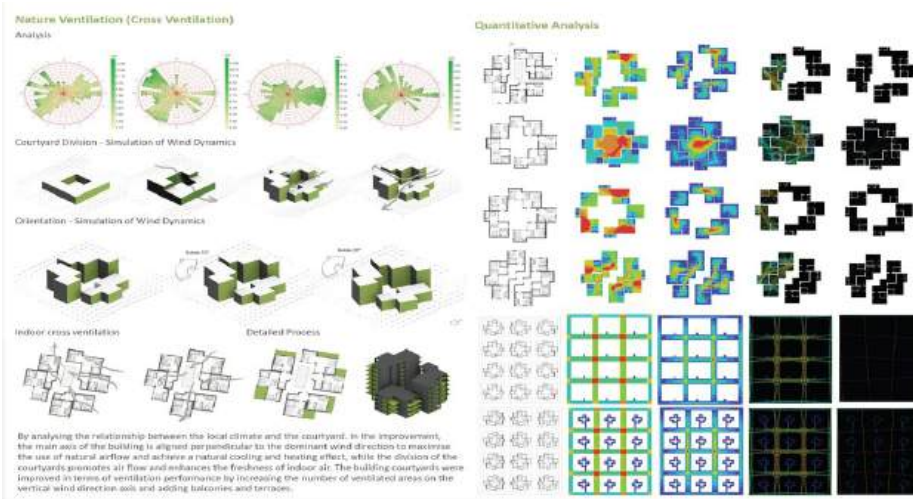
Reflection in action in the context of the two trials was framed to be holistic in nature, distinguished as summative - presented as a type of analogue structuring of transactions, or synthesis - illustrated through a back and forth formation of knowing, as goal, methodology, strategy validation, for design implementation.

In trial one, student were tasked to devise a quantitative methodology for building and site environmental performance simulation to solve a design problem. The assignment required students to produce a summative understanding of environmental performance simulation approaches, aligned to a comprehensive design strategy that would address one or more problems that the student was to define. Simulation tools included depthmapX, ladybug & honeybee environmental plug-ins for Rhino, computational fluid dynamic modeling to achieve thermal comfort through, mean wind speed ratio (MWSR), and comfort wind zone ratio (CWZR) analysis. Within each type of simulation software there are a number of tools for focused analysis. It was anticipated that students would address design problems such as mitigating urban heat island, site building orientation paired to timed solar irradiance exposure, site design relationships to climate data with site wind dynamics, and spatial configuration that would enhance user comfort through daylighting, ventilation flows and visual experience.

The research approach in this trial undertook content analysis of student individual synoptic submissions to categorize the design problem, applied simulation methodology, applied design strategy, summative or synthesis format, partial or comprehensive design approach.

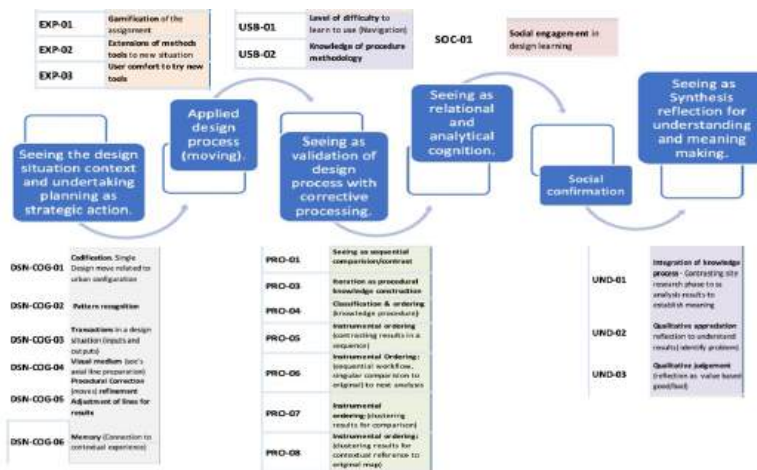


**Figure 2:** Trial one sample examining interior courtyard environmental performance simulations applying space syntax approach.



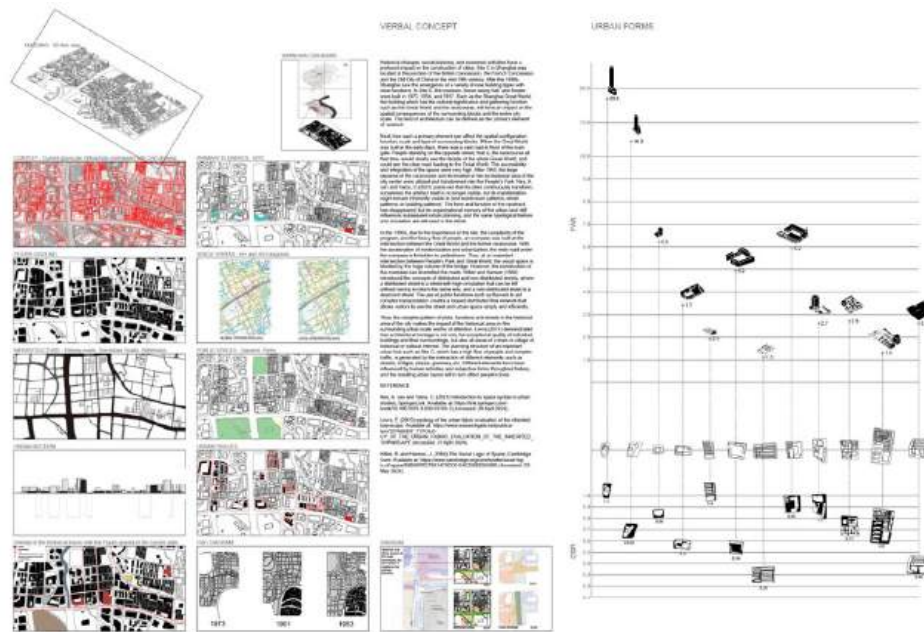
**Figure 3:** Trial one sample of iterative testing of design variations of spatial configuration applying space syntax approach and environmental performance simulations for ventilation.

In Trial two, the urban area surrounding the site of the architecture project was analyzed to understand spatial morphology transformation over time. Students were required to examine the impact of additional variables on qualitative aspects of the site. Information approaches used by students include, physical site surveys, content mapping of spatial function location, user behavior in the urban environment, figure-ground diagramming with space syntax depthmapX to be integrated into the student course work. Students generated coded representations of variables at different resolutions and scales and selectively formatted the presentation as 2D and 3D maps. The assignment required a formative synthesis to demonstrate the students ability to engage in qualitative value judgment, and present summative or synthesize insight regarding the urban spatial structures transformations over time.



**Figure 4:** Framework for the questionnaire survey for Trial two aligned to the RSDC model.

The research method applied in this trial presented students with an online survey requiring them to reflectively identify their applied mental process. The aim of the survey was to distinguish dominant cognitive processes, framed as gamification and exploration, information processing, pattern recognition, instrumental ordering and analysis typologies, memory, social validation, qualitative appreciation and qualitative judgement (see Figure 4). A content analysis of the course work submission was not undertaken in trial 2.



**Figure 5:** Trial two, sample synopsis panel with syntax analysis of street configuration contrasted to building typology and urban spatial morphology.

## FINDINGS

### Information Processing, Simulations and Strategy Formation

In Trial one, 26 course work submissions directly applied the space syntax methodology out of 141 total submissions. Results of the content analysis of submitted synoptic presentations, reveal that mean number of applied simulations to be 4.8, with minimum of 2 and maximum of 12 simulations applied. The mean number of design strategies applied was 7, with minimum of 2 and maximum of 12.

Ranking of applied typology of methods and strategies is shown in Figure 6 and 7. The design strategy developed required a robust number of design approaches be applied to solve sustainable problems and develop a comprehensive design proposal. To this end the majority of student submissions developed fine grain strategies that were surprisingly creative in applying space syntax approaches, and aligned to environmental simulation modeling data outputs. This was an unexpected outcome for the trial, requiring further investigation.



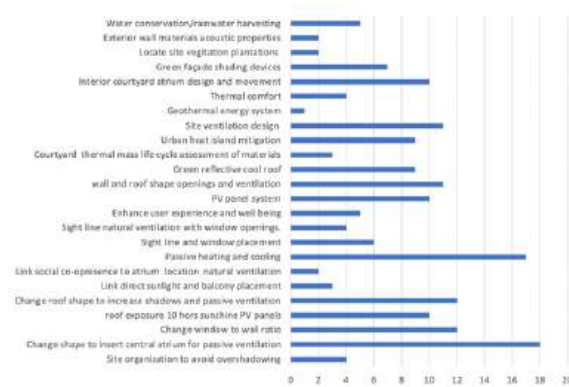


Figure 6: Trial 1 applied design strategies.

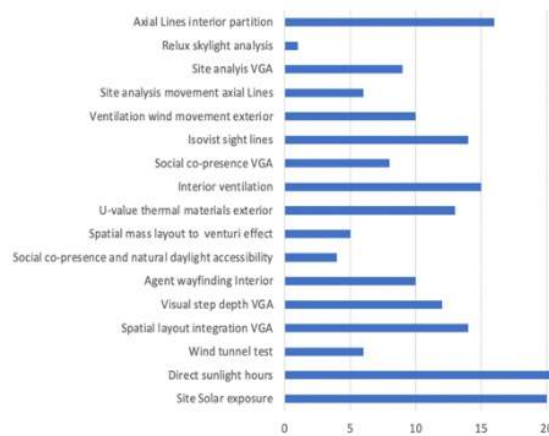


Figure 7: Trial 1 applied simulations.

### Reflection Synthesis Psychometric Processes

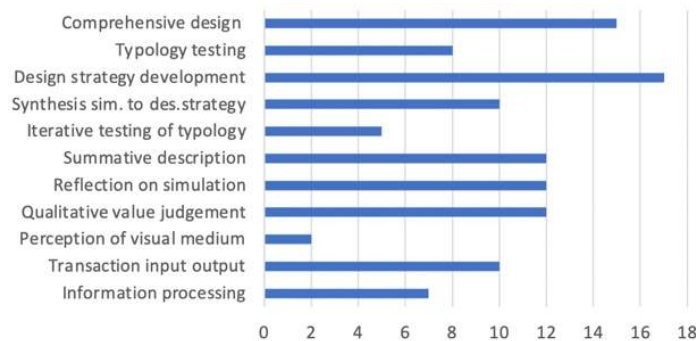
The cognitive processes listed below reach levels of significance of 4.20 and above, to best illustrate student agreement to typologies of psychometric activity (see Figure 8). The survey results highlight a number of interesting findings regarding reference and reflective mental processes for students working with analogue and digital modelling. Pattern recognition processes received the lowest agreement response with instrumental ordering, contrasting results in a sequence receiving the highest agreement.

Significant processes distinguished in the survey relate to seeing as a sequential comparison or contrast activity, instrumental ordering as contrasting results in a sequence, followed by instrumental ordering of sequential workflow with singular comparison, and instrumental ordering/clustering results for comparison with instrumental ordering, clustering results for contextual reference. Mental tasks of memory (connection to contextual experience) and transactions in a design situation (inputs - outputs), are also seen to play a role in formative understanding.



**Figure 8:** Survey results on designers reflective awareness of mental processes Trial 2.

In Trial two, formative understanding is distinguished by mental processes such as ‘integrate knowledge process - contrasting site research phase’ and ‘qualitative appreciative reflection to understand results’. This is contrasted to Trial one, where design strategy and comprehensive design reach highest agreement as the significant pedagogical outcome. Student samples highlight summative description, reflection on simulation, and qualitative value judgment to be significant, followed by mental processes that primarily engage transactions of input output and synthesis linking simulation methods to design strategy.



**Figure 9:** Types of psychometric processes used by students in Trial 1.

**CONCLUSION**

This paper reports on findings from two trials framed by a model of reflection synthesis design cognition (RSDC) that extends Schön’s seeing-moving-seeing model, applied as reflection in action with material of a design situation. It is noted that further clarification mapping into reference vs. reflection in

action with materials from simulation modelling is required, to identify sub-classification types of design cognition. Further investigation into formative knowledge construction is also needed to fine tune the distinction between summative and synthesis processes in reflection, in order to advance cognitive learning models for design pedagogy. A second interesting outcome of this study is the experimentation richness that has resulted from introducing space syntax into environmental performance simulations. With limited guidance, students were successful in the formative construction of simulation models that informed unique, creative and comprehensive design proposals. While abstract thinking processes are often highlighted in conceptual design generation, more concrete outcome oriented simulations are required now to advance sustainability to the comprehensive design stage in architectural study at the junior pedagogy level.

## REFERENCES

- Atkin O How do architects design? An Artificial Intelligence and Pattern Recognition in Computer-Aided Design, Latombe J. C. Ed. North Holland Press,. 1978. New York, pp. 65–104.
- Baya V. and Leifer, L. 1996. Understanding Information Management in Conceptual Design, in Cross N., Christians H., Dorst K., Eds. *Analysing Design Activity*, John Wiley and Sons, Chichester. pp. 151–168.
- Dorst K. 1997. *Describing Design: A comparison of paradigms*, Ph. D. thesis, TU Delft.
- Dorst & Cross. 1998. Co-Evolution of problem and solution spaces in creative design; observations from an empirical study, Computational models of creative design, IV J. Gero and M. L. Maher (ed) University of Sydney USW Australia.
- Eastman, C. 1969 Cognitive processes and ill-defined problems: A case study from design. *Proceedings Joint International Conference on Artificial Intelligence*. Walker, D., and Norton L., eds. ACM, Washington D. C.
- Ericson K. and Simon H. 1984. *Protocol Analysis: Verbal reports as data*. MIT Press, Cambridge MA.
- Eppinger S. D. 1999. Model-based approaches to managing concurrent engineering. *Journal of Engineering Design*. 2(4): 283–290.
- Goel G. A., Grue N., Murdock W., Recher M., Govindaraj T. 1996. Design explanations in interactive design environments, in Gero J. Eds. *Proceedings Fourth International Conference on AI in Design*, Palo Alto.
- Kazi S., Galanaki E. 2020. Piagetian theory of cognitive development. Entry in the *Encyclopedia of Child and Adolescent Development*. Hupp S. & Jewell J. D. Eds. Wiley Publishers, New York.
- Kline, S. J. 1985. Innovation is not a linear process. *Re. Mang* 28(4):36–45.
- Kolb D. A. 1984. *Experiential learning: Experience as the source of learning and development (Vol. 1)* Englewood cliffs, NJ. Prentice-Hall.
- Kolodner J. L. & Wills L. M. 1996. Powers of observation in creative design. *Design Studies* 17(4): 385–416.
- Middleton D. A. 2024. *Integrating Space Syntax into the architecture design studio: Assessing adoption of new software tools to enhance critical reflective practice*. Unpublished report EDU PGCERT 402 Xi'an Jiaotong-Liverpool University une 12, 2024.
- Pahl, G. & Beitz, W. 1996. *Engineering Design: A systematic approach*, 2<sup>nd</sup> Edition Springer.

- Schon D., 1992. Designing as Reflective Conversation with the Materials of a Design Situation, *Research in Engineering Design*, 3. pp. 131–147.
- Schon D. & Wiggins, G. 1992. Kinds of seeing and their function in designing. *Design Studies*, 13:2, pp. 135–156.
- Wynn, D. & Eckert, C. M. 2016. Perspectives on iteration in design and development. In *Res. Eng. Design*. Springer-Verlag, London.

# AdTech's AI Appetite: A Case Study in Advertisers' Perceptions and Concerns of AI Integration

Rachael Boyle, Ruslana Pledger, and Hans-Frederick Brown

LiveRamp, San Francisco, CA 94104, USA

## ABSTRACT

This case study explores the challenges and considerations of adopting artificial intelligence (AI) at LiveRamp, a business-to-business software-as-a-service (B2B SaaS) advertising company. By surveying both customers and employees, we garnered insights into AI perceptions (i.e. concerns, barriers, ideas) and performed correlation analysis to investigate the factors that influence these perceptions. While the 635 survey responses revealed a positive attitude towards the adoption of AI within the business, it also showed that many respondents are conflicted on their AI stance as well. Top concerns vary between privacy for employees and accuracy for customers. Factors like prior AI tool usage were shown to influence and correlate with AI perceptions. We believe this work can be valuable for other B2B SaaS businesses navigating the complexities of AI adoption, especially those within heavily regulated industries like advertising.

**Keywords:** Human-computer interaction (HCI), Artificial intelligence (AI), Advertising, AI perceptions, AI adoption, User surveys, Business-to-business, Software-as-a-service

## INTRODUCTION

Over the past decade, there has been unprecedented advancement and commercialization of artificial intelligence (AI) technologies. These technologies have far exceeded computational expectations, transforming the experience between the user and the machine while prompting pressing questions about ethics, privacy, and security. With the creation of embeddable AI tools, it is theoretically easier than ever for companies to integrate AI into their products. But balancing consumer demands while navigating an undefined regulatory landscape can be incredibly challenging. How does an organization determine when and how to take the leap of investing in AI over other competing priorities?

Answering these questions is difficult, but especially so for B2B SaaS organizations given their distance from the ultimate end-user who would best guide these decisions, and especially so if they exist within highly regulated industries like advertising. AI has been used in this space, prior to this most recent AI boom, for actions like predictive modeling, targeting, and ad personalization. These capabilities led to a level of precision and

persuasiveness that, for better or worse, had not been seen before. Thus the question is, what does this new era of AI mean for advertising?

As researchers at LiveRamp, a B2B SaaS company that operates in this space, we sought to find these answers and inform our company's own AI strategy through the perspective of our customers and employees. We found these two perspectives (i.e. employees as the adopters and integrators of AI, B2B customers as the users of the AI tools) to be missing in existing literature today since much of it focuses on the end-consumer (i.e. the ones seeing or engaging with the output of these AI tools). While these insights are invaluable, it is just one piece of the puzzle for B2B companies whose customers are not individuals but rather businesses. How our customers think about AI should inform AI strategies because their experience in our products determines the success of our business. Meaning, they are experts and the ultimate influencers of what our company should and should not pursue. How our employees think about AI should inform AI strategies because they are the most knowledgeable in our offerings, technical resources, and privacy / data ethics restrictions. Meaning, they are the influencers and ultimate experts of what we can and can not do.

With this in mind, we constructed a 23 question survey and captured responses from 469 of our customers and 166 of our employees at our company, LiveRamp, to help understand the perceptions, hesitations, risks, barriers, and ideas that surround implementing AI into our product suite. Through the insights uncovered in this survey, we help guide and ground our company's AI strategy with the opinions of those most important to enabling and ensuring its success. Through sharing our insights with other B2B SaaS businesses, we hope to serve as a launchpad for their own AI strategization. Through sharing our insights with the broader HCI community, we hope to add key perspectives to the discussion of AI adoption in highly regulated industries like advertising, unveil key barriers / issues to this adoption, and start to create the connective tissue for this AI conversation between various stakeholders (i.e. for advertising, between consumers, marketers, ad technology providers).

## **LITERATURE REVIEW**

While we are entering a new age of AI, it is by no means a new concept in the field of advertising. AI has been discussed in this industry as early as the 1950s (Huh et al., 2023), with the discussions evolving as the technology evolved, the use cases and wealth of data growing increasingly complex in turn. It grew clear that this complexity could be aided through the use of AI to derive better consumer insights (Kietzmann et al., 2018). Many papers dive into potential use cases which, prior to large-language models (LLMs), included things like consumer insight discovery (Li, 2019), ad creation and impact evaluation (Li, 2019), predictive analysis modeling (Enache, 2020), media planning and buying (Chen et al., 2019), personalization (Rafeian & Yoganarasimhan, 2023), and even synthetic ad creation (Campbell et al., 2021).

With the most recent advancements, AI tools will undoubtedly introduce “dramatic changes in how advertisements are conceived, produced, edited, and targeted” (Colin et al., 2022). While AI can allow for better advertising control (Shah et al., 2020), research claims the new challenge marketers face is ensuring the use of these tools does not elicit negative reactions from their end-consumers, like feelings of eeriness (Wu & Wen, 2021) or concerns over their data privacy. While unlocking these use cases could help businesses be more effective or gain a competitive advantage, the potential harms are plentiful as well. These include: job security risks; revenue model impact; organizational structure changes; lack of transparency, deemed “critical to developing ethical and socially responsible AI technology” given the trend towards technological opaqueness and “consumers’ limited ability to understand AI technology” (Huh, Jisu et al., 2023); and reputational or legal risks.

Much of the existing literature is focused on these marketer benefits or consumer concerns. Little to no research focuses on businesses, specifically B2B businesses, and the adoption phase such a business must undergo before deciding to implement AI. “While AI holds many possibilities for marketers, achieving its potential is not easy” (Campbell et al., 2019). While some are successfully considering or embedding AI in their business strategy decisions (Baker, 2017; Wolska, 2017), many businesses “are struggling to see the forest for the trees” and “navigate AI adoption” (Campbell et al., 2019). Businesses within heavily regulated industries especially need this sort of guidance. Beyond a complicated regulatory landscape, there is a steep learning curve, unclear return on investment, and large reputational risks if anything goes wrong (Rodgers, 2021).

However, AI commercialization and market demands ensure that more and more businesses will need to go through this adoption and strategizing phase in order to stay competitive (Enache, 2020). As Campbell et al. (2019) states, “AI should be a consideration for all marketing managers as it represents the highest growth of any technology in marketing (Salesforce, 2017), is expected to increase in use (Columbus, 2018), and is predicted to have a \$40 billion effect on marketing by 2025 (Reavie, 2018).” We argue that B2B companies’ AI adoption journey (from AI consideration, “AI Foundation”, to “AI Orientation”) is an important area of study since it produces many downstream effects that warrant its own research (i.e. effects on the consumer, societal impact).

To guide our own company’s adoption journey, we surveyed our employees and our customers (i.e. marketers) to understand their perceptions and concerns of integrating AI into our product offering as well as pinpoint where in our customers’ journey we should or should not be introducing it. While existing research focuses on the consumer or the marketer, this case study highlights the businesses that connect the two together. These businesses typically (though not solely) implement the AI technologies that marketers leverage and whose output consumers engage with. Thus, they are important to include in the framework of perspectives we evaluate. We hope our work begins to answer some of the calls to action made in existing literature to introduce frameworks that guide “the future of AI advertising”

(Huh et al., 2023) and investigate the use / impact of AI as it is integrated into existing advertising processing (Li, 2019).

## METHODOLOGY

### Survey Design

To ensure our questions would be tactically impactful to our company's roadmap, we gathered the key user tasks within our products, or "Jobs To Be Done" (JTBD), and framed questions from the context of those jobs (Christensen & Raynor, 2003). The survey went through multiple rounds of review with relevant stakeholders (i.e. product management, engineering, data ethics). Our final survey had a total of 23 questions with a variety of matrix, multiple choice, multi-select, and free-response questions. We only made pivotal questions mandatory to provide optionality for our respondents and to ensure data quality. Our questions can mainly be grouped into one of the following categories:

- *Background* (4 questions) to compare cross-sections of our responses by respondent type (i.e. employee vs customer, product usage, job function)
- *AI in the Customer Journey* (6 questions) to understand what tasks or jobs-to-be-done (JTBD) users are already leveraging AI for, would like to leverage AI for but aren't, or have no intentions of using AI for
- *AI Concerns and Barriers to Use* (9 questions) to understand concerns barriers that prevent users from already leveraging existing AI solutions
- *AI Tradeoffs* (6 questions) to understand viewpoints on tradeoffs between risks vs. opportunities and transparency vs. automation for AI integration
- *Importance of AI* (2 questions) to understand how important incorporating AI is and what information is most critical for it to succeed
- *AI Suggestions Questions* (2 questions) to understand users free-form thoughts on AI and how LiveRamp should incorporate it

### Data Collection

The survey, created on Qualtrics, was distributed internally and externally via email to gather feedback from both employees and customers. Customers were incentivized with a raffle drawing, while employees received no incentive. A total of 7471 customers and 1755 employees were contacted, resulting in 635 responses. Data analysis utilized techniques such as data visualization, cross tabulation, and correlation analysis. The questionnaire, analysis, and data (anonymized for privacy) are accessible on GitHub for future scholarly reference.

### Hypothesis Testing

Before conducting our survey, we formulated hypotheses regarding the potential relationship between AI experience, technical proficiency, and employee status on AI perceptions. We established a null hypothesis (H0) stating that there is no significant correlation between these factors, and



alternative hypotheses positing correlations between AI experience (H1), technical proficiency (H2), employee status (H3), and AI perception. To test these hypotheses, we used survey questions on job function, product usage, and AI tool usage as proxies for these variables. The categorical responses were then converted into numerical scores based on our expertise. Additionally, survey responses to questions about AI perception were used as variables and already measured on a numerical scale.

## RESULTS

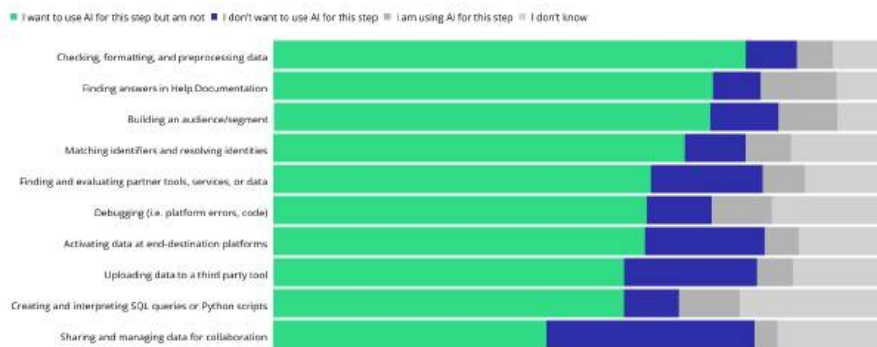
Out of 9226 individuals surveyed, we received 635 responses, yielding a response rate of 6.88%. Among the 7471 customers surveyed, we obtained 469 responses (6.27%), while out of the 1755 employees surveyed, we received 166 responses (9.46%).

### Background

The majority (74%) of respondents were external customers, with 26% being employees. The most common job function among respondents was Marketing (27%), followed by Operations (13%) and Data Science (10%). Majority (66%) of our customer respondents use one or more of our flagship products which are typically more UI-based, requiring less technical knowledge. Among customer respondents, a large majority had prior experience with AI tools, with ChatGPT being the most popular (82%). Other notable tools included DALL-E (21%), Bard (16%), and Midjourney (11%).

### User Journey

Thinking about the marketing and advertising related tasks you perform (whether using LiveRamp or not), please sort each task into the following categories:



**Figure 1:** AI use preference for 325 of our customer respondents across all JTBD. The JTBD are sorted in order of highest number of “want to use” selections (in green) to least.

The survey included user journey matrix questions focusing on customers’ main Jobs To Be Done (JTBD) regarding product usage. The responses

indicated a strong desire to use AI for various tasks, with the majority expressing interest in leveraging AI but currently not doing so. Notably, JTBDs like checking formatting and preprocessing, building an audience, and finding answers in help documentation stood out as areas where customers expressed a desire to use AI but were not currently doing so. On the other hand, customers reported using AI for tasks such as finding answers in help documentation, debugging, and creating/interpreting SQL queries. Majority (64%) of those already using AI for certain JTBD were satisfied with their experience. Those dissatisfied listed scalability, lagging adoption, unreliability / unexpected behaviors, rate limits, accuracy, and the addition of manual steps as reasons for their dissatisfaction. While there is a strong desire to use AI for various tasks, some JTBD customers showed reluctance to use AI for, such as data sharing, uploading data, activating data, and evaluating partner tools.

The survey also assessed attitudes towards the level of automation or human oversight if AI was implemented across JTBD. The responses generally leaned towards wanting some oversight on AI outcomes, indicating a preference for semi-automated processes. Tasks like building segments and finding/evaluating partner tools were highlighted as areas where respondents wanted complete oversight on AI. However, there were differences between employee and customer responses in terms of desired oversight, suggesting nuanced perspectives based on user roles. Additionally, the JTBD of managing data with partners for collaboration and analysis received a notably higher number of responses advocating against AI use, indicating a wariness of AI being involved in data collaboration and a higher scrutiny of the type of role it plays.

### **Concerns and Barriers**

When surveyed about their concerns regarding the use of AI in LiveRamp's products, respondents showed a higher level of concern when asked specifically about potential biases in AI algorithms compared to their overall concerns about AI usage. Specifically, 48% expressed concern about bias, while 23% were concerned about AI usage overall. The top overall AI concerns included privacy and data security (79%), accuracy of AI-driven insights (66%), transparency in how an AI decision was reached (55%), ethical data usage (53%), control/configurability over AI process (36%), and the inability of AI system to capture human complexity (36%). While the top two concerns were consistent across employees and customers, there were some notable differences.

Moreover, customers' concerns regarding AI varied depending on whether they were asked tactically or theoretically about it. When asked tactically, accuracy was the top concern, followed by privacy. However, when asked theoretically, privacy emerged as the primary concern. Barriers to leveraging AI for desired tasks included technical resource limitations, potential inaccuracies, privacy/security concerns, work prioritization, lack of AI options at work, lack of transparency/explainability, inability of AI systems to capture human complexity, lack of configurability/control, ethical data usage concerns, and cost. Interestingly, privacy concerns seemed to be

overshadowed by other technical constraints when respondents were asked in a tactical manner.

### **AI Tradeoffs**

A majority of customers (56%) agreed that the potential benefits of AI outweigh the risks, with 30% remaining neutral and 14% disagreeing. To validate these findings, a reversed question was posed later in the survey, yielding similar trends: 20% agreed, 27% were neutral, and 52% disagreed. This suggests a general inclination towards favoring AI's benefits over risks, alongside a notable level of uncertainty or neutrality among respondents. Regarding comfort levels, most respondents were comfortable (68%) with AI providing suggestions. However more were neutral (19%), with a slight skew towards discomfort (46%), with AI making decisions for them. This indicates a greater uncertainty among respondents regarding AI decision-making compared to AI offering suggestions.

### **AI Importance**

Respondents were divided on the importance of AI integration into LiveRamp's products, with 37% being neutral, 42% considering it important (17% very important), and 21% deeming it relatively unimportant (8% not important). While there's a slight skew towards importance, a significant portion remains undecided. Subsequently, respondents ranked the importance of AI details, with information on how privacy is maintained topping the list (mean ranking: 2.29), followed by understanding how data powers AI (mean ranking: 2.45), how AI features work (mean ranking: 2.54), and knowing what data is used to power AI (mean ranking: 2.72). While the order was consistent across employee and customer respondents, slight differences in alignment and average rank were observed.

### **Suggestions and Ideas**

From free response questions, the following themes emerged on the use cases AI might be most helpful with: data recommendation / augmentation, training and customer support (i.e. "if LiveRamp has 24/7 technical support, that would be great"), natural language searches and queries (i.e. "provide python [and] sql template scripts for jumpstart"), data onboarding (i.e. "if an uploaded file isn't formatted correctly, AI could be used to automatically correct the uploaded file"), data analysis (i.e. "having AI to apply filters, create visualization or [give] information on how to interpret numbers would help a lot"), and enhanced identity resolution and/or backend automation.

Employees also had ideas on how AI could help internal processes within LiveRamp, specifically seeing it useful for: onboarding new hires, engineering processes (i.e. LLMs for code base), customer support via internal and external chatbots (i.e. "tying together historical support case threads with slack messages, email threads, and conversations related to times where similar issues were resolved or similar questions were asked by clients"), sales processes, and competitive analysis. Staunch AI proponents highlighted concerns around taking an "excessively cautious" stance on AI given the

concrete ways “we could meaningfully enhance our processes and products today.” While some proponents concede that the state of AI does not seem “super useful” quite yet, they still call for LiveRamp to closely examine how it can make us a “better company.”

While the majority of comments were generally positive about integrating AI into our products, there was also a passionate minority concerned with this path. These comments ranged from mild to serious concern, specifically around themes of data reliability (i.e. “if AI had any malfunction and provided the wrong data to an important client that could hinder future business and budget costs”), bias and regulatory risk (i.e. “it’s clearly important that AI doesn’t end up accidentally causing a whole set of biases that create risk for marketers, given the complex regulations in place that govern how and why marketers can use data to make advertising decisions”), and broader AI cynicism (i.e. “there’s nothing you can do to convince me AI will work in this type of application”).

### Differences in AI Perception

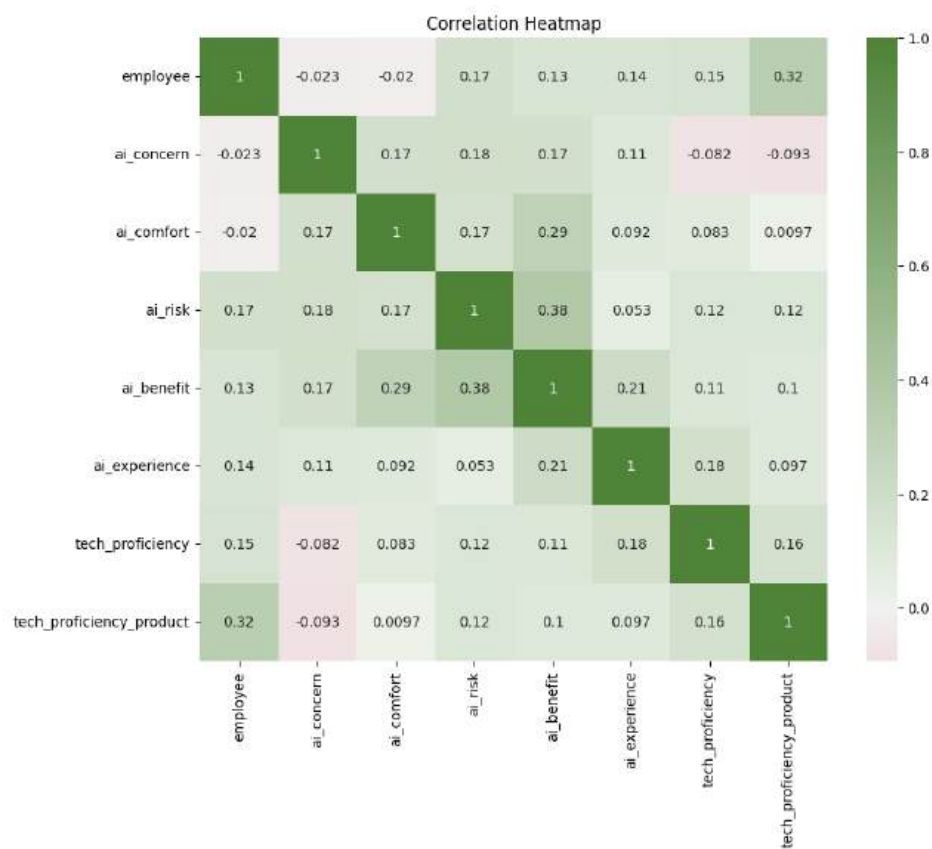
Customer respondents were mostly from marketing backgrounds, while engineering was predominant among employee respondents, indicating a difference in technical proficiency that could influence response nuances. There was also variation in AI tool usage, with Bard being notably more popular among employees compared to customers. In terms of AI stance, customers appeared more conflicted and neutral compared to employees. Customer responses showed more neutrality regarding the importance of integrating AI into LiveRamp’s platform and the balance between AI benefits and risks. Similarly, customers were more undecided about their comfort level with AI making decisions within products. Conversely, employees showed clearer opinions, with a majority considering AI integration important and expressing greater comfort or greater concern with AI decision-making. Differences also emerged in preferences for oversight in the user journey with AI introduction. Customers prioritized tasks related to audience building and partner tool evaluation, while employees focused more on SQL queries and data management, reflecting concerns about privacy. Moreover, top concerns differed between employees and customers, with employees emphasizing privacy and ethical considerations more than customers, who were more concerned about accuracy. These differences likely stem from each group’s role, responsibilities, and relationship to LiveRamp.

**Table 1.** Statistically significant correlation between variables of AI perception and variables for prior AI experience, technical expertise, and employee status.

Variables	Spearman’s $\rho$	p-value
<i>AI_EXPERIENCE</i> and <i>AI_CONCERN</i>	0.1072	0.0273
<i>AI_EXPERIENCE</i> and <i>AI_BENEFIT</i>	0.2076	< 0.001
<i>TECH_PROFICIENCY</i> and <i>AI_BENEFIT</i>	0.1119	0.0288
<i>EMPLOYEE</i> and <i>AI_RISK</i>	0.1731	0.0085
<i>EMPLOYEE</i> and <i>AI_BENEFIT</i>	0.1339	0.0062

## Correlation Analysis

To test our hypotheses, we first created our variables by identifying relevant questions from the survey and properly transforming the data. For AI perception, we took questions on the user's level of AI concern, on whether AI's benefits outweigh AI risks, on whether AI's risks outweigh AI benefits, and comfortability with AI making decisions. These questions were each on a 5-point agreement scale and re-coded to 1–5 representing increasing acceptance of AI. For technical proficiency, questions on job function and LiveRamp product usage were used as proxies and mapped to a technical proficiency score ranging from 0 (less technical) to 2 (highly technical) based on the technical requirements of the role or product. The same was done for prior AI experience where 0 represented no experience and 2 denoted more experience with AI tools. Responses indicating the use of the most mainstream AI tool, ChatGPT, were given a score of 1, while use of less popular tools were scored as 2.



**Figure 2:** A heatmap of the calculated Spearman's rank correlation Coefficients between all variables, regardless of statistical significance. A dark green indicates a strong, positive coefficient equal to 1. A dark red indicates a strong, negative coefficient equal to -1.

We then employed Spearman's rank correlation coefficient to assess the strength and direction of monotonic relationships between AI attitude variables and the variables we hypothesized might influence that attitude. We used Spearman's correlation since it is robust to non-normality and works better with ordinal data. The results of this analysis allowed us to reject the null hypothesis and showed support of H1, H2, and H3. However, all statistically significant correlations between variables show only a weak or very weak correlation. The strongest correlation (Spearman's  $\rho = 0.2076$ ) was shown to be between prior AI experience and the belief that AI's benefits outweigh its risks (i.e. those with more experience using AI are more likely to have a positive perception of AI).

### LIMITATIONS AND RECOMMENDATIONS

The study acknowledges several limitations despite providing insights into AI adoption dynamics. Firstly, there's an uneven distribution of employee and customer respondents, potentially introducing bias. Survey limitations include the inability to provide context or probe deeper into responses, as well as the challenge of conditional questions affecting response rates. The study's abstract nature due to our early AI adoption stage also limits specificity in questions, possibly leading to ambiguity or confusion for respondents filling out the survey. For our analysis, the assignment of AI experience and technical expertise scores is subjective and may lack reliability. Suggestions for future research include conducting similar studies across different contexts and industries to compare AI appetites, executing longitudinal studies to track perception changes over time, and investigating factors influencing AI implementation success or failure.

### CONCLUSION

This study explores the challenges and considerations of AI adoption for B2B SaaS companies within regulated industries through surveying 469 customers and 166 employees at LiveRamp, a B2B SaaS company in advertising. Results showed overall positive attitudes towards AI adoption but revealed nuanced insights. We hope these findings provide guidance for refining AI strategies in this space and contribute to the discourse on AI adoption within businesses.

### REFERENCES

- Campbell, C., Planger, K., Sands, S., and Kietzmann, J. (2021). Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising. *Journal of Advertising*, 51(1), pp. 22–38.
- Campbell, C., Planger, K., Sands, S., Kietzmann, J., and Bates, K. (2022). How Deepfakes and Artificial Intelligence Could Reshape the Advertising Industry: The Coming Reality of AI Fakes and Their Potential Impact on Consumer Behavior. *Journal of Advertising Research*, 62(3), pp. 241–251.
- Campbell, C., Sands, S., Ferraro, C., Tsao, H. Y., and Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons*, 63, pp. 227–243.

- Chen, G., Xie, P., Dong, J., and Wang, T. (2019). Understanding Programmatic Creative: The Role of AI. *Journal of Advertising*, 48(4), pp. 347–355.
- Christensen, C., and Raynor, M. (2013). The innovator's solution: Creating and sustaining successful growth. *Harvard Business Review Press*.
- Enache, M. C. (2020). AI for Advertising. Annals of Dunarea de Jos University of Galati. Fascicle I. *Economics and Applied Informatics*, 26(1), pp. 28–32.
- Ford, J., Jain, V., Wadhvani, K., and Gupta, D. (2023). AI advertising: An overview and guidelines. *Journal of Business Research*, 166, p. 114124.
- Horzyk, A. (2023). Data Protection and Privacy: Risks and Solutions in the Contentious Era of AI-Driven Ad Tech. *International Conference on Neural Information Processing* (pp. 352–363).
- Huh, J., Nelson, M., and Russell, C. (2023). ChatGPT, AI Advertising, and Advertising Research and Education. *Journal of Advertising*, 52(4), pp. 477–482.
- Kietzmann, J., Paschen, J., and Treen, E. (2018). Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence Along the Consumer Journey. *Journal of Advertising Research*, 58(3), pp. 263–267.
- Li, H. (2019). Special Section Introduction: Artificial Intelligence and Advertising. *Journal of Advertising*, 48(4), pp. 333–337.
- Rodgers, S. (2021). Themed Issue Introduction: Promises and Perils of Artificial Intelligence and Advertising. *Journal of Advertising*, 50(1), pp. 1–10.
- Shah, N., Engineer, S., Bhagat, N., Chauhan, H., and Shah, M. (2020). Research Trends on the Usage of Machine Learning and Artificial Intelligence in Advertising. *Augmented Human Research*, 5(1).
- Wu, L., and Wen, T. (2021). Understanding AI Advertising From the Consumer Perspective: What Factors Determine Consumer Appreciation of AI-Created Advertisements? *Journal of Advertising Research*, 61(2), pp. 133–146.

# Privacy Policy Analysis and Evaluation of Mobile Psychological Consultation Services in Saudi Arabia

**Abdulmajeed Alqhatani**

Department of Information Systems, College of Computer Science and Information Systems, Najran University, Najran, 61441, Saudi Arabia

## ABSTRACT

Psychological consultation apps have been increasingly used in the last few years. These services collect a variety of sensitive personal information. Typically, a privacy policy is the main way to reduce users' concerns about sharing personal health information. However, psychological consultation apps are considered an emerging service and their privacy practices have not been fully explored. This study analyzes and evaluates the privacy policies and Terms of Service (ToS) agreements of four Saudi psychological consultation apps, focusing on seven key privacy practices: types of collected information, the purpose of data collection, data sharing, data ownership, data retention, data storage and protection, and notifications about policy and ToS updates. Overall, the findings indicate that the privacy policies of these services must be improved to better inform users, particularly regarding the purpose of collecting their data, with whom the data are shared, and what data are archived. This paper also provides a set of implications to improve the existing privacy policies of psychological health apps.

**Keywords:** Psychological health apps, Mental health apps, Privacy, Privacy policy, mHealth

## INTRODUCTION

Mobile Health (mHealth) has become widely used, especially after the global COVID-19 pandemic. In Saudi Arabia, the use of mHealth services has increased significantly (Alanzi, 2022), and this increase can also be attributed to the National Transformation Program which encourages rapid digitalization of healthcare services. Psychological consultation apps are an mHealth service that people find helpful. Such apps provide users with therapeutic and consultation services, including live sessions and instant messaging with psychotherapists, symptom tracking, and social support. Studies indicate that these health apps can positively impact users (Marques et al., 2021).

Given the nature of such services, users may need to disclose personal information to the application and the consultants, and some of this information can be considered sensitive. Typically, privacy policies and ToS agreements are the main mechanism to communicate to users how their personal information is collected, shared, and protected. Yet in terms of



mHealth, these services are still new, and their privacy practices have been relatively unexplored.

This study analyzes and compares the privacy policy of four popular Saudi apps that are used for psychological consultation, focusing on seven practices: types of information collected, the purpose of data collection, data sharing, data ownership, data retention, data storage & protection, and policy updates. This paper makes the following contributions:

- Analysis of the privacy policies of four popular Saudi Psychological health apps that collect sensitive personal information.
- Evaluation that reveals the privacy policies of psychological health apps lack sufficient details regarding key practices about which users should be informed.
- Identification of a set of design implications and future research opportunities to improve the privacy policies of this important type of mHealth services.

## **BACKGROUND**

“The authors would like to indicate that the terms psychological health, psychological support, and mental health are used interchangeably throughout the remainder of the paper.”

### **Psychological Health Applications**

Mobile Health (mHealth) apps present new and convenient methods for personal health management. These digital technologies are perceived by health professionals to promote the wellbeing of users. Among mHealth apps that have gained attention are those used for tracking mental health issues, such as depression and anxiety (Linardon et al., 2024; Atallah et al., 2017). During and after the global coronavirus (COVID-19), The use of apps for tracking psychological health has especially become prevalent (Alanzi, 2022). Marques et al. (2021) studied the impact of Covid-19 on the mental health of Spanish university students, and the findings show that many of the participants were willing to use mobile apps to manage their mental health.

However, there are some issues that might hinder the use and engagement with health apps, such as usability and ease of use, lack of trust, and privacy issues (Alqahtani and Orji, 2020; Qingchuan Li, 2020). The study by Alqahtani and Orji (2020) analyzed user experience with mental health apps, and they found several factors that impact people to abandon the use of mental health applications; those factors include mainly poor usability security, privacy, and trust concerns. Privacy issues cannot be ignored given the sensitivity of some of the information collected about users of psychological health apps. Li (2020) indicates that privacy concerns may have a significant impact on the use of health applications. The next part of this section discusses privacy concerns related to medical apps, including psychological health apps.

## Privacy Concerns

Medical applications collect a variety of personal data ranging from contact (e.g., phone number and email) to demographic (e.g., age and gender) and health-related information (e.g., biometrics). This collection can impose several privacy issues (Sharma et al., 2021; Lustgarten et al., 2020). For example, data breaches targeting digital healthcare systems can negatively impact its users, including embarrassment and reputation damage (Iwaya et al., 2020). There is also a concern that users' data inside those health apps might be shared with third parties without users' awareness (Kitkowska et al., 2023).

Researchers have further explored mHealth apps to understand users' behaviors and potential solutions to privacy problems. Building upon the privacy paradox, Zhu et al. (2021) indicate that users of mHealth are intended to disclose their personal information if they perceive benefits. However, people sharing preferences in health applications may also be affected by other factors, such as trust and previous security and privacy incidents (Kitkowska et al., 2023). Implementing social cue factors within health apps can decrease privacy concerns when users are interacting with both physicians and the application itself (Zhang et al., 2022).

To protect the sharing and use of personal data in digital systems, several governments have initiated regulations, such as the European Union's General Data Protection Regulation (GDPR) and the Saudi's Personal Data Protection Law (PDPL). Given the current advancements in digital systems, those regulations may not fully address the issue (Nurgalieva et al., 2020). A key privacy protection method is to improve users' awareness about service providers' data practices through privacy policies, which will be discussed next.

## Privacy Policy

Privacy policy is a legal document presented by a service provider to explain the different practices in relation to users' data. For example, it specifies what data will be collected from a user and for what reasons, with whom it will be shared, and how it will be protected. However, psychological support apps can be considered new in the context of the mobile app industry, and privacy practices related to those apps have not been fully addressed.

Parker et al. (2019) analyzed 61 mental health apps from a dominant-English speaking market, and they found that nearly half of the examined apps did not provide a privacy policy for users. In a similar study, Robillard et al. (2019) reviewed 100 mental health apps from popular digital stores, indicating that most of the apps did not provide a privacy policy and a ToS agreement. Rowan & Dehlinger (2014) conducted a comparison study of the privacy policies of health and fitness tracking applications. Their study shows that specific permissions in those apps are insufficiently addressed by the privacy policy developers.

If people are to provide their mental health information, it is important to evaluate the transparency of the service's data practices. O'Loughin et al. (2019) evaluated the transparency of data processing in mobile apps used

for depression. Out of 116 apps examined in their study, only 4% of apps were assigned an “acceptable” score for transparency. Huckvale et al. (2019) evaluated the content of privacy policies of apps used for depression and smoking, and only 64% of those policies presented descriptions about secondary uses of the data. It is essential to inform users about sharing their personal information with external parties, as previous studies have demonstrated that mHealth data can be aggregated from multiple sources and re-identified (Galvin & Demuro, 2020). This current study differs from previous research in that it focuses on 7 key practices that a privacy policy should address to enhance users’ awareness of sharing their mental health information.

## METHODOLOGY

This research seeks to understand the data practices of mental health app companies by analyzing these apps’ privacy policies. Figure 1 below shows the methodology applied for that analysis.



**Figure 1:** Study procedure.

The first phase involved identifying the selection criteria for choosing the apps to be examined. To be included in the study, each app had to satisfy two conditions: (a) it must be approved by the Saudi Ministry of Health, (b) and must have a release date in or prior to 2020. Using those inclusion criteria, the author searched the Apple App store for mental health apps using the key word “psychological consultation.” Apps that satisfied the conditions were then selected based on the “top charts” feature in the Apple store. Four mental health-related apps were downloaded for analysis that were ranked among the top 100 “medical” and “health and fitness” apps in the Apple store, namely, Labayh, Estenarh, Tetaman, and Famcare.

In the second phase, the author signed up and used each service for a while in order to understand the possible situations and scenarios that require

sharing different personal information in each service. For example, some apps optionally request the disclosure of certain health information, such as asking about the mental health history of a user and the user's family or whether a customer has ever been harassed or bullied for the purpose of evaluating a case.

Next, three evaluators independently read and take notes of each app's privacy policy and ToS agreement, focusing on the seven primary data practices (previously mentioned in the introduction) that privacy policies normally address.

In the last phase, each service policy was summarized and a table was created to compare the commonalities and differences between these policies regarding of the seven evaluation criteria. The evaluators then discussed each policy together and an overall score was assigned to each service representing the extent to which the privacy policy fulfills each of those privacy aspects.

## EVALUATION

In this section, each service privacy policy and ToS agreements are described in the light of the seven identified criteria for privacy. Then, a table that summarizes and rates each policy is presented to assess how well each of the privacy aspects is implemented in each service.

### Analysis of Privacy Policies

**Types of collected information.** Companies, typically, collect different personal information from users to enable the provision of services or shape future improvements. Yet service providers must inform users about what personal information is being collected about them. Overall, each service seems to collect a variety of personal data, including gender, birth date, marital status, mobile number, email, medical history, and payment information. Labayh, Tetaman, and Famcare seemed to thoroughly identify the types of information they collect about users by classifying information into the following categories: basic account information (name, password, mobile number, payment information, information regarding customers' complaints, evaluations of consultants by customers, and information sent by users to company email addresses. However, Famcare privacy policy also states that they may collect users' information from other sources and aggregate it for analysis.

**Purpose of information collection.** One important aspect that impacts users' willingness to share their mental health information is whether they are informed about how their information will be used. Estenarh and Famcare provide detailed descriptions about the purpose of collecting personal data, such as using cookies and log data to provide technical services and use content for advertisements. However, in its policy, Estenarh mentioned that it may not require a user's consent to use some of the recorded data in the app. The other two apps did not provide an explicit statement in their policy describing why they collect users' information.

**Data sharing.** Users' comfort levels depend on who will have access to their personal information via data sharing (Galvin and DeMuro, 2020). Labayh

and Tetaman assert that users and consultants data will not be shared with any party other than what is permitted according to the law of Saudi Arabia, and that the data will only be accessible to the service team only. Famcare indicates that users' data will not be shared with external parties without users' consent. Estenarh mentioned that the recorded data may be shared with several parties, including the judicial authority, research parties, and future owners of the service.

**Ownership of data.** Use and sharing decisions can significantly be influenced by the level of control a user has over personal information. Estenarh states that users can be somewhat in control over their data through the account settings, but some data will still be kept even if users opt-out of the service. The other services did not clearly explain how users can be in control of their own personal data.

**Data retention.** One major concern of users when using a service, especially health-related services, is whether the service provider keeps the data after users stop using the app and, if so, for how long. Labyah and Tetaman policies include vague statements, indicating that users' data will be retained as long as the user continues to use the service and as long as the account remains registered in the service. Those statements are each followed by a statement indicating that some personal data will be retained for lawful and regulation reasons. Famcare mentioned in its policy that some basic user's information will be deleted from their servers if the user requests so, but other information will be saved as long as the Famcare needs it. Estenarh clearly identified a list of information that will be archived for commercial, technical, and lawful reasons.

**Data storage & protection.** The examined services provide about the same level of data security. Their policy states that users' data, such as passwords, will be encrypted. Estenarh privacy policy additionally states that any documents shared by a user will be encrypted and that the connection between the client and the consultant will completely be confidential.

**Policy changes.** All the services state that they will post any modifications to their privacy policies directly in the application. The ToS agreement of both Labyah and Tetaman and the privacy policy of Famcare state that users may also be notified via email about any amendment to the service and user conditions.

### Comparison Table

Table 1 below summarizes the previous investigation of the privacy policy and ToS agreement of each app in terms of the seven privacy criteria. For each service, a score is assigned that shows how well each privacy aspect has been described in each policy (where: 0 = not described, 1 = partially described, and 2 = fully described). An overall score is also calculated for the privacy policy of each service.

**Table 1.** Evaluation of the Four Privacy Policies.

	Labyah	Estenarh	Tetaman	Famcare
Types of collected information	2	1	2	2
Purpose of information collection	0	1	0	2
Data sharing	1	2	1	2
Ownership of data	1	2	1	1
Data retention	1	2	1	1
Data storage & protection	1	2	1	1
Policy updates	2	1	2	2
<b>Total score (out of 14)</b>	<b>8</b>	<b>11</b>	<b>8</b>	<b>11</b>

## IMPLICATIONS

**Improving transparency** - Studies indicate that users are willing to disclose their personal information for the sake of their wellbeing and mental health (Zhu et al., 2021). However, if users do not feel that they are in control of their personal information, they might give up using the service. Thus, privacy policymakers should provide clear and easily understandable descriptions of how users' information is handled by the app. When data is shared with external parties, users should have a choice to share their information in part or in full, and they should be allowed to opt-in and out at any time. Therefore, users will feel more comfortable about sharing their data and so will be able to build trust with the app.

**Anonymization** - Given the nature of psychological consultation apps, users are likely to disclose sensitive personal information to the consultant during treatment sessions or to the app itself. The disclosed information may contain personally identifiable information and is typically stored in the servers of the app's company. Previous incidents have shown that health information is valuable to hackers. In this study, only one service mentioned that users' information will be anonymized. If data breaches occur on a company's server, customers' information will be compromised. Therefore, mHealth apps should implement anonymization techniques, such as data masking and generalization to protect the confidentiality of customers' information.

**Visual cues** - One issue with a privacy policy is that it usually comes lengthy and full of jargon, which enforces users to skip reading it. Rather, visual cues can engage a user to pay attention to specific information in the privacy policy. Such cues can be as simple as different text colors to represent sensitive personal information collected by the app. It can also be designed as comics to summarize key information in the policy (Tabassum et al., 2018).

**Compliance with regulations** - The increasing advancements in digital technologies mean that regulations, such as the GDPR need to be reviewed and updated. This also means that mHealth companies need to ensure that their privacy policies align with existing regulations. Conducting regular assessments of a privacy policy and a ToS is a good practice to ensure compliance with existing laws and regulations.

## LIMITATIONS

The present study has two limitations. First, it was limited to analyze the privacy policies of psychological consultation apps in the Saudi market. The reason for this decision is because Saudi Arabia has witnessed a significant growth in the development of mHealth apps, and we wanted to explore the current privacy practices by mobile app companies in this market. Second, the study investigates the policies of the top five mental health apps in the Apple App store only, which represents a small portion of apps for mental health.

## CONCLUSION

Mental health apps are still emerging in the context of medical apps, and the privacy policies of these services have not been fully studied. With the goal of assessing the current-state of transparency in the privacy practices of psychological consultation apps, this study analyzes four services that are approved by the Ministry of Health in Saudi Arabia. The overarching findings of this study indicate that the privacy policies lack transparency regarding why certain data are collected, with whom data are shared, and what data are retained after a user stops using the service. Further, while all services claim to protect users' data through encryption, only one service indicates that users' information will be anonymized if or when the app shares data with other parties. In the future, we plan to examine users' awareness and expectations about mental health app services and to what extent those expectations align with the existing privacy policies.

## ACKNOWLEDGMENT

The author would like to thank Aseel Reshan and Maryam Jammah for helping in the analysis process.

## REFERENCES

- Alanzi, T. M., 2022. Users' satisfaction levels about mHealth applications in post-Covid-19 times in Saudi Arabia. *PloS one*, 17(5), p. e0267002.
- Alqhatani, F. and Orji, R., 2020. Insights from user reviews to improve mental health apps. *Health informatics journal*, 26(3), pp. 2042–2066.
- Atallah, N., Khalifa, M., El Metwally, A. and Househ, M., 2018. The prevalence and usage of mobile health applications among mental health patients in Saudi Arabia. *Computer methods and programs in biomedicine*, 156, pp. 163–168.
- Galvin, H. K. and DeMuro, P. R., 2020. Developments in privacy and data ownership in mobile health technologies, 2016–2019. *Yearbook of medical informatics*, 29(01), pp. 032–043.
- Huckvale, K., Torous, J. and Larsen, M. E., 2019. Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA network open*, 2(4), p. e192542.
- Iwaya, L. H., Ahmad, A. and Babar, M. A., 2020. Security and privacy for mHealth and uHealth systems: A systematic mapping study. *IEEE Access*, 8, pp. 150081–150112.

- Kitkowska, A., Karegar, F. and Wästlund, E., 2023, September. Share or Protect: Understanding the Interplay of Trust, Privacy Concerns, and Data Sharing Purposes in Health and Well-Being Apps. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter* (pp. 1–14).
- Li, Q., 2020. Healthcare at your fingertips: The acceptance and adoption of mobile medical treatment services among Chinese users. *International journal of environmental research and public health*, 17(18), p. 6895.
- Linardon, J., Torous, J., Firth, J., Cuijpers, P., Messer, M. and Fuller-Tyszkiewicz, M., 2024. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World Psychiatry*, 23(1), pp. 139–149.
- Lustgarten, S. D., Garrison, Y. L., Sinnard, M. T. and Flynn, A. W., 2020. Digital privacy in mental healthcare: Current issues and recommendations for technology use. *Current opinion in psychology*, 36, pp. 25–31.
- Marques, G., Drissi, N., de la Torre Díez, I., de Abajo, B. S. and Ouhbi, S., 2021. Impact of COVID-19 on the psychological health of university students in Spain and their attitudes toward Mobile mental health solutions. *International Journal of Medical Informatics*, 147, p. 104369.
- Nurgalieva, L., O’Callaghan, D. and Doherty, G., 2020. Security and privacy of mHealth applications: a scoping review. *IEEE Access*, 8, pp. 104247–104268.
- O’Loughlin, K., Neary, M., Adkins, E. C. and Schueller, S. M., 2019. Reviewing the data security and privacy policies of mobile apps for depression. *Internet interventions*, 15, pp. 110–115.
- Parker, L., Halter, V., Karliychuk, T. and Grundy, Q., 2019. How private is your mental health app data? An empirical study of mental health app privacy policies and practices. *International journal of law and psychiatry*, 64, pp. 198–204.
- Robillard, J. M., Feng, T. L., Sporn, A. B., Lai, J. A., Lo, C., Ta, M. and Nadler, R., 2019. Availability, readability, and content of privacy policies and terms of agreements of mental health apps. *Internet interventions*, 17, p. 100243.
- Rowan, M. and Dehlinger, J., 2014. A privacy policy comparison of health and fitness related mobile applications. *Procedia Computer Science*, 37, pp. 348–355.
- Sharma, T., Islam, M. M., Das, A., Haque, S. T. and Ahmed, S. I., 2021, June. Privacy during pandemic: A global view of privacy practices around COVID-19 apps. In *Proceedings of the 4th ACM SIGCAS Conference on Computing and Sustainable Societies* (pp. 215–229).
- Tabassum, M., Alqhatani, A., Aldossari, M. and Richter Lipford, H., 2018, April. Increasing user attention with a comic-based policy. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–6).
- Zhang, J., Luximon, Y. and Li, Q., 2022. Seeking medical advice in mobile applications: How social cue design and privacy concerns influence trust and behavioral intention in impersonal patient–physician interactions. *Computers in human behavior*, 130, p. 107178.
- Zhu, M., Wu, C., Huang, S., Zheng, K., Young, S. D., Yan, X. and Yuan, Q., 2021. Privacy paradox in mHealth applications: An integrated elaboration likelihood model incorporating privacy calculus and privacy fatigue. *Telematics and Informatics*, 61, p. 101601.



# Human Detection Method by 3D-LiDAR With Low Calculation Costs

Haruki Mochizuki<sup>1</sup> and Ryoza Kiyoharara<sup>2</sup>

<sup>1</sup>Graduate School of Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

<sup>2</sup>Kanagawa Institute of Technology, Atsugi, Kanagawa, Japan

## ABSTRACT

Robot technologies are developing and it is becoming possible to perform not only simple tasks but also tasks that involve interaction with humans. In particular, robots are expected to be introduced in patrol security, which requires a large number of personnel, as much of the work involves confirming that there are no problems. In Japan, for example, robots that serve food at restaurants move around in the flow of people. In terms of security, if it cannot be confirmed by robots that there is no problem, a security guard is required to go and check. In order to reduce such cases, improving the accuracy of confirmation is an absolute requirement. In addition, it is necessary to reduce the cost of introducing and operating the robot as much as possible. Technology for recognizing people using cameras mounted on robots and 3D-LiDAR (Light Detection And Raging) has already been established and is in use. However, no technology has been established for use in poorly lit areas or by low-cost 2D-LiDAR recognition. In this paper, we propose a new method for reducing calculation cost and evaluate our method.

**Keywords:** Human detection, 3D-LiDAR, Security robots

## INTRODUCTION

In many developed countries, aging populations and slow immigration are predicted to result in labor shortages. Labor shortages are already a problem in Japan. In the security industry, robots are being introduced for simple patrol tasks. Patrol task is primarily crime deterrent. If suspicious humans are found, its information have to be reported. Robot carts that can freely move outdoors at night are equipped with 3D-LiDAR (Light Detection And Raging) in many cases. Moreover, they move while estimating its own position using SLAM (Simultaneous Localization and Mapping) technologies.

SLAM with 3D-LiDAR requires many calculation costs. Therefore, human detection processing should have as little computational cost as possible. Therefore, we propose a human recognition method that uses only part of the pointcloud data obtained from the on-board 3D-LiDAR as data obtained from multiple 2D-LiDARs. If 3D-LiDAR collects pointcloud data from 16 different heights, using only one can reduce the number of pointcloud data to 1/16. Furthermore, the computational cost of machine learning evaluation is also reduced. Therefore, we evaluate the validity of the proposed method by

evaluating the person recognition rate using multiple 2D-LiDARs. Moreover, by evaluating the learning required for the human orientation and distance from LiDAR, we evaluate whether it is possible to recognize people in any orientation or distance.

## LIDAR

LiDAR emits laser light and measures distance based on the phase difference of the reflected light. By continuing to rotate the irradiation, pointcloud data can be collected. LiDAR is installed on the Robot Cart as shown in Figure 1. Robot Cart estimate the location by SLAM with Pointcloud data to move.



**Figure 1:** Robot cart with 3D-LiDAR as a security guard.

## RELATED STUDIES

Many studies were conducted regarding human recognition. Human recognition using image processing is one of the most commonly used methods. However, Robot as a security guard moves in midnight. Therefore, there are not enough light. There are some studies on human recognition with 3D-LiDAR. Pedestrian recognition using 3D-LiDAR for autonomous vehicles (Nagai, 2022) uses a support vector machine (SVM) for human recognition. Other approaches use 3D LiDARs as multiple 2D- LiDARs (Chung, 2012). This study gathered pointcloud data from many different heights and achieved a classification rate of up to 96%. An example of using reduced pointcloud data is dividing 3D pointcloud data in the target into parts, that is, the trunk and legs, to achieve robust in situ pedestrian recognition (Asad, 2020).

All of these studies use 3D pointcloud data, which requires a lot of CPU resources. 3D-LiDAR can obtain a pointcloud of the entire person. Therefore, it can identify the person regardless of the shape of the clothes or other factors. However, there is no need to have multiple person shapes that can be obtained from similar height locations. Therefore, in this study, we propose a person recognition method that uses multiple 2D-LiDARs to reduce computational costs.

In our previous study (Nagai, 2022), we proposed a person detection method based on low-resolution 2D-LiDAR and a PointNet-based

autoencoder (Qi, 2017). We used 2D-LiDAR to collect data from both the front and side from a height of 90 cm, which is considered the typical waist height of a human. The proposed method successfully detected people within 5 m of the LiDAR.

In a previous study (Watanabe, 2023), we proposed the human recognition method by multiple 2D-LiDARs. However, both methods require to recognize from exact front. Moreover in a previous study (Mochizuki, 2024), we showed possibility of our method. In this paper, we propose a method for recognizing people regardless of the sensing angle and distance.

## PROPOSED METHOD

Based on the case at Kanagawa Institute of Technology, human to be detected by patrol security is an elderly human who enters the campus without malicious intent. By focusing on human between 135 cm and 175 cm in Japan, which is the average height of elderly people by age, it is possible to detect most of the elderly human who intrude without malicious intent. Human characteristics are broadly classified into head, torso, arms, and legs. The head has a smaller proportion on the body than the torso, arms, and legs. Moreover the height used to obtain pointcloud data about the head tends to change depending on the height. Therefore, we obtain pointcloud data about human torso, arms, and legs. The pointcloud data is acquired by setting the range in which the torso and arms can be acquired in the height range of 135 cm to 175 cm, with the torso and arms being 80 cm to 110 cm in height, and the legs being 20 cm to 55 cm in height.

In this paper, limited number of lasers irradiated by 3D-LiDAR can reduce the amount of pointcloud data to obtain for reducing calculation cost. The laser is selected by distance from obstacles and installed height. Therefore, the distance should be less than approximately 25 m to select more than three rays which are for arms, torso and legs. Moreover, if the distance is more than approximately 37 m, it is impossible to get pointcloud data for legs. Therefore, we assume that the robot only recognizes people within a distance of 25m (see Table 1).

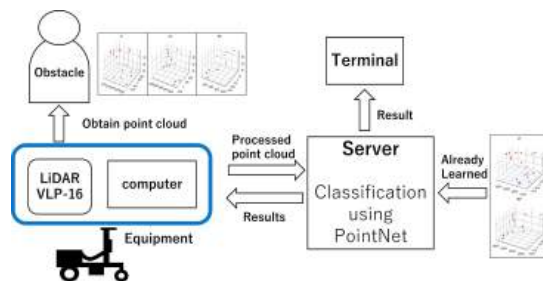
**Table 1.** Distance and height relations for 3D LiDAR( height : cm).

Laser No.	2 m	5 m	15 m	25 m	37 m
7	89.5	126.3	249.1	378.1	525.4
5	82.5	108.7	196.2	288.1	393.0
3	75.4	91.2	143.6	198.6	261.5
1	68.4	73.7	91.1	109.5	130.4
-1	61.5	56.27	38.8	20.4	0.00
-3	54.5	38.8	0.0	0.0	0.0
-5	47.5	21.2	0.0	0.0	0.0

When a human is detected, the security robot notifies the security guard. Figure 2 shows the recognition process used in patrol security to notify security guards when a person is detected.

First, a robot equipped with LiDAR and a computer patrols and collects pointcloud data of the surrounding area. Background subtraction processing is performed on the computer equipped with the acquired pointcloud data.

Next, the extracted pointcloud data of obstacles such as people and objects are sent to the server through the wireless network, and the server side uses PointNet to identify human or obstacle. If the target is recognized as a human, a security guard is notified, the security guard heads to the robot's location, and the robot tracks the human.



**Figure 2:** Notification flow for security guard when human is detected.

On the server, in order to perform human recognition using the acquired pointcloud data, they are classified using machine learning technologies. To build a system for classification, use “Classify pointcloud using PointNet,” which is published in the MATLAB library. This is a method that uses PointNet to compress and restore pointcloud data and extract only the features for classification.

Classification using PointNet is supervised learning, so it is necessary to learn pointcloud data to be classified in advance. Therefore, it is learned pointcloud data about humans and trees, and classify the acquired pointcloud data into humans and trees. Since the number and shape of pointcloud data differ depending on the size of the target object, learning is performed by acquiring point cloud information of various types of humans and trees. Furthermore, the number and shape of pointcloud data that can be acquired differs depending on the orientation of the person. When humans learn one direction, it is difficult to recognize directions that have not been learned.

In the case of front facing, pointcloud data of two human legs are obtained. Pointcloud data for the two legs are obtained in the same way even in orientations other than landscape. Therefore, by learning the front facing direction, it is possible to recognize human with a certain degree of accuracy in situations other than those facing sideways. Moreover, since the learned orientation can be determined in the case of landscape orientation, it can be determined by learning the landscape orientation.

Therefore, we think that by learning two directions, frontal and sideways, the range of human orientations that can be determined will be increased, making it possible to identify human even in a variety of orientations.

## EVALUATION

Table 2 shows the specification of the LiDAR for our evaluation.

**Table 2.** Specification of the LiDAR for our evaluation.

Name	VLP-16 ( by Velodyne)
Range	100m
View Angle (Vertical)	30°(±15°)
View Angle (Horizontal)	360°
Rotational Speed	5Hz-20Hz

First, the number of pointcloud data have been compared between 16 lasers and 2 lasers for standing human at various distance. Moreover, we also have been compared the number of pointcloud data acquired when using two 3D-LiDAR lasers and when using two 2D-LiDAR for standing human at various distance. We have been able to confirm to reduce the amount of pointcloud data.

Next, accuracy of human recognition is most important. Then, we have confirmed the accuracy of human recognition by two lasers of 3D-LiDAR. The distance range of pointcloud data to be acquired has been set from 2 m to 5 m, and training data and test data have been acquired separately. In the training data, pointcloud data for humans and trees have been used to prepare pointcloud data for each class classification. For the human pointcloud data in the training data, we have prepared 20 pieces of tree pointcloud data, 10 pieces each of frontal and sideways pointcloud data as the two orientations to be learned. Moreover, pointcloud data of humans and trees acquired from a distance of 3 m from LiDAR has been used as training data. The number of epoch has been 150 for each. As test data, we have prepared 16 pieces of pointcloud data of humans facing forward (0°), diagonally (45°), sideways (90°), and trees, acquired at distances of 1 m each from 2 m to 5 m. Table 3 shows the F value of the result of learning the training data in advance and classifying it.

**Table 3.** F Value by limited 3D LiDAR for our evaluation.

	0°	45°	90°
2m	1.00	1.00	1.00
3m	1.00	1.00	1.00
4m	1.00	1.00	1.00
5m	1.00	1.00	1.00

From this result, sufficient accuracy has been able to be obtained when acquired using 2 lasers in 3D-LiDAR.

We have confirmed whether 3D-LiDAR has been able to measure human torso, arms, and legs, and has been able to recognize human in various orientations at a maximum measurement distance of 25 m. The distance range of the pointcloud data to be acquired has been set from 5 m to

25 m, and training data and test data have been acquired separately. For training data, we have prepared pointcloud data for humans and trees for class classification. For the training data, we have prepared 10 pieces of human pointcloud data for each of the two orientations to be learned, frontal and sideways, and 20 pieces of tree point cloud data. Moreover, pointcloud data of humans and trees acquired from a distance of 3 m from LiDAR has been used as training data. The number of epoch have been 150. For the test data, we have prepared 16 points each of pointcloud data of a human facing forward, diagonally, sideways, and a tree, taken at 5 m intervals from a distance of 5 m to 25 m. Table 4 shows the F value of the result of learning the training data in advance and classifying them.

**Table 4.** F Value by distance for our evaluation.

	0°	45°	90°
5m	1.00	1.00	1.00
10m	0.91	0.91	0.00
15m	1.00	0.67	0.75
20m	0.91	0.67	0.91

According to the recognition results, for diagonal orientations, the accuracy decreased to 0.67 over 15 m. Moreover, although it has been possible to distinguish between humans and trees at other distances when facing sideways, it has not been possible to distinguish between humans and trees at 10 m. We think that the cause has been the low rotation speed and the small number of acquired pointcloud data, which led to a decrease in accuracy. In other words, the pointcloud data used for training data has been too little and it has not been possible to distinguish between landscape orientation and tree features. Moreover, because the number of test data has been too little, it is thought that the shape shown by the pointcloud data are biased. Therefore, it necessary to find an appropriate amount of data.

## CONCLUSION

In order to relieve the burden of patrolling security at universities due to the labor shortage problem in the security industry, we have verified the human recognition rate in various orientations using a security robot equipped with 3D-LiDAR. We have proposed a method that have been able to reduce the calculation cost and recognize human with high accuracy by acquiring fewer pointcloud data than the original number for person recognition using 3D-LiDAR. We have showed that by selecting two lasers, it was possible to sufficiently recognize human even in situations where the amount of pointcloud data is little. However, the accuracy have not been stable when the distance was large, so it has been necessary to acquire more point cloud data and perform classification.

In the future work, the influence of clothing and human posture on recognition accuracy should be verified. In particular, jackets and skirts can worsen recognition accuracy.

## REFERENCES

- Asad, L. Ehsan, N. Luis F. M-M. (2020) “Development and evaluation of a real-time pedestrian counting system for high-volume conditions based on 2D LiDAR,” *Transportation Research Part C: Emerging Technologies*, Vol. 114, pp. 20–35.
- Chung, W. Kim, H. Yoo, Y. Moon, C. -B. and Park, J. (2012) “The Detection and Following of Human Legs Through Inductive Approaches for a Mobile Robot With a Single Laser Range Finder,” *IEEE Trans. on Industrial Electronics*, vol. 59, no. 8, pp. 3156–3166.
- Mochizuki, H. Watanabe, T. Horii, M. and Kiyohara R. (2024), “Stationary Human Detection Method by 2D-LiDARs,” *IEEE International Conference on Information Networking (ICOIN)*.
- Nagai, Y. Sawano, Y. Terashima, Y. Suzuki, T. and Kiyohara, R. (2022). “A Method for Detecting Human by 2D-LiDAR,” *IEEE International Conference on Consumer Electronics (ICCE)*.
- Qi, R. C. Su, H. Mo, K. Guibas, J. L. (2017). “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Watanabe, T. Terashima Y. and Kiyohara R. (2023), “Standing Human Detection Method Using 2D-LiDARs,” *IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1260–1265.

# Sexting, Age and Digital Vulnerabilities

Ingvar Tjostheim<sup>1,2</sup>, Chris Wales<sup>1</sup>, and John A. Waterworth<sup>3</sup>

<sup>1</sup>Hauge School of Management, NLA, 0130 Oslo, Norway

<sup>2</sup>Norway & Norwegian Computing Center, Norway

<sup>3</sup>Umeå University, Umeå, Sweden

## ABSTRACT

Despite popular misconception, it is not only young people that are sharing nude pictures and videos. There are a number of studies targeted towards the younger generation about their sexting, but few on older age-groups. In general, younger people take more risk than older people, and some seem to care less about possible negative consequences. For this study we commissioned a market research company to collect data from a national population, with a representative sample from 16 to 69 years old, in total 1071 citizens. We used binary logistic regression for the analysis of responses, a method that can be used to predict a categorical dependent variable – in our case whether a person has been sexting the last 12 months or not. In the study we included the following independent variables: gender, education, self-efficacy; the cognitive reflection test (CRT) to distinguish between a intuitive versus analytical decision style; Machiavellianism, to distinguish a personality trait characterized by manipulateness and deceitfulness; willingness to share personal data, and finally whether the citizens had experience of ID-theft or credit-card misuse within the previous 12 months. Our results show that the ID-theft/credit-card variable was a significant predictor of sexting for the age-groups 16-29, 30-39 and 50-69 years old. For youngest group, the manipulateness and deceitfulness trait is also a predictor, whereas for the oldest group, the intuitive decision style and a high willingness to share personal data are also significant predictors.

**Keywords:** Sexting, Digital vulnerability, Personal traits, The cognitive reflection test, Willingness to share personal data

## INTRODUCTION

Sexting involves the sharing of nude images of others, or of oneself alone, or together with others (Reyns 2014). Barrense-Dias et al. (2017) distinguish between passive and active sexting, but there is no clear definition of the two, whilst Dodaj and Sesar (2020) identified four forms of sexting: relational, reactive, forced and violent. Dir et al.'s (2013) results showed that higher frequency of sexting was positively related to impulsive behaviour (sensation seeking and negative urgency). Machimbarrena et al. (2018) have shown many potential negative consequences. Although sexting can be a means of flirting or enhancing a sexual relationship it can highlight potential vulnerabilities to victimisation (Cooper, 2016). March and Wagstaff (2017) included the “dark triad” of Machiavellianism, narcissism, and psychopathy in their study of sexting.



While many studies of sexting are only about social and psychological vulnerabilities, our study includes two personality traits, and the willingness to share personal data and financial vulnerability such as ID-theft. Individuals that have experienced ID-theft, debit or credit-card misuse might share some similarities with those who engage in sexting. A question that intrigues us relates to whether a person who has recently been sexting is more likely to experience ID-theft or credit-card misuse or vice versa. We have used a survey to investigate this phenomenon. While some respondents will give socially desirable answers, others are not concerned by the information they reveal. Studies on the survey's method (Krosnick, 1991; MacKenzie and Podsakoff, 2012) show that we cannot always assume honest and correct answers. For the questions about ID-theft and credit-card misuse, we asked the respondents to tell us about what happened to them. Most of the respondents did give us the additional information of the incidents. We therefore believe that the answers to this question is reasonably accurate. For the other questions, the personal trait and the sexting vs has not-been-sexting questions, we do not have additional information that can be used to evaluate level of honesty. In general, this type of survey is used for statistical purposes, such as those carried out by Eurostat (European Union, 2017).

For this paper we formulated two research questions: to what extent can sexting be predicted by intuitiveness, the willingness to share personal data, experience of ID-theft or credit-card misuse, or demographical variables? Secondly, which of these variables are the best predictors, if any, of sexting?

## **A NATIONAL SURVEY ABOUT ONLINE BEHAVIOUR**

Professional market research companies recruit individuals that can participate in surveys. Sometimes market research companies use the term 'panels' to indicate that the company has a database with contact information. Our participants were members of a large national panel of Norwegian citizens. For each survey, panel-members with demographic characteristics similar to the actual population are invited to participate. Often, for Norway, the size of the sample is set to approximately 1000 to get a representative sample. It does not mean that the results are accurate or represent a good prediction every time, but over time with this sample size and method of recruiting respondents, reliable results can be assumed.

We commissioned the market research company Norstar AS to collect data from the Norwegian national population, with a representative sample aged from 16 to 69 years old, in total 1071 citizens. To characterise our participants, we present their demographic profile in Table 1 and Table 2 for the four age-groups that we use in our analysis. The demographic profile our participants are:

**Table 1.** Gender and the age profile of the participants.

	16-29 years N = 201	30-39 years N = 281	40-49 years N = 200	50-69 years N = 389
Male (N = 514)	40%	48%	48%	52%
Female (N = 557)	60%	52%	52%	48%

**Table 2.** Educational profile of the participants.

	Primary or secondary education	Bachelor degree (or equivalent)	Higher degree (Master or higher)
19-29 years, N = 201			
Male	80%	16%	4%
Female	70%	18%	12%
30-39 years, N = 281			
Male	47%	25%	28%
Female	27%	32%	42%
40-49 years, N = 200			
Male	47%	26%	27%
Female	30%	31%	39%
50-69 years, N = 389			
Male	47%	27%	27%
Female	39%	32%	29%

The survey was filled in by the respondent on either a PC, a tablet or a mobile phone. For this study our response variable is whether the citizen has shared nude pictures or videos of themselves or of others, or both them and others. This method is a self-reported measurement, therefore we cannot test whether or not the respondents are actually telling the truth. In general, for large scale survey by market research companies we assume that data is reliable. For some, and we speculate that it could be the case for individuals in the oldest age-groups, this type of behaviour is not socially acceptable and, in many cases, criminal behaviour.

## THE MEASUREMENTS IN THE STUDY

The respondents answered questions about whether they personally had experienced ID-theft or credit-card misuse within the last 12 months. Individual Machiavellianism was measured with the trimmed MACH-IV (Rauthmann, 2013; Christie & Geis, 1970) a 5-item Likert scale where participants respond to statements such as “Anyone who completely trusts anyone is asking for trouble”. For self-efficacy, we used 3 items from the Psychological Capital Questionnaire (Luthans, Youssef et al., 2007). For ID-theft and credit-card misuse, the questions were identical to the questions used in surveys by the national bureau of statistics in Norway and the Eurostat-survey on European attitudes toward cyber security.

The Cognitive Reflection Test (Frederick, 2005) has IQ-type questions with either a correct or an incorrect answer. It can be used to assess

participants thinking styles, described as intuitive versus analytical. In the survey, the respondents were asked to answer four numerical CRT-questions, three drawn from Frederick (2005) and a question about the number of pupils in a class (Toplak, 2014). Furthermore, we include a consent-test (Tjostheim and Waterworth, 2020), which is a behavioural measure concerning disclosure of personal data and demographics. In the consent test, we asked the participants to give us access to all their personal data that the market research company already had in their database, as such access is not otherwise allowed under the contract agreed between the panel members and the market research company. Although some respondents gave consent, we did not actually receive the data from the marker research company, but a yes-answer indicates a high willingness to share personal data.

**Table 3.** Descriptive statistics of the variables and measurements (N = 1071).

	Min	Max	Mean	SD	Skewness	Kurtosis
Gender, Male-female	1	2	1.52	0.500	-0.08	-1.997
Education	1	3	1.82	0.835	0.373	-1.468
Data-disclosure (binary)	1	2	1.67	0.470	-0.735	-1.462
Sharing nudes (binary)	1	2	1.12	0.327	2.322	3.398
ID-theft, Credit-card misuse (binary)	1	2	1.15	0.361	1.919	1.686
CRT	1	4	1.48	1.053	0.198	-1.003
Self-efficacy	1	3	2.11	0.842	-0.209	-1.559
Machiavellianism	1	5	2.37	1.051	0.714	0.095

Table 3 shows descriptive statistics of all 8 variables. Kurtosis indicates the extent to which a distribution of scores is relatively flat or relatively peaked. Skewness indicates the extent to which scores have a tendency toward the upper or lower end of a distribution. There is a skewness problem if the result is greater than  $\pm 2.0$ , which is the case for the variable 'sharing nudes', where both skewness and kurtosis are peaked. In total, across all age-groups, 12% reported that they had sexted in one form or other. This means that it a small group, and in our analysis it is used as the dependent variable, and not together with other independent variables. The following table shows the percentages.

**Table 4.** Percentages that have shared nude pictures or videos - the four age-groups.

	16-29 years	30-39 years	40-49 years	50-69 years
Male	24%	15%	16%	6%
Female	23%	12%	14%	2%
All	23%	14%	15%	4%

Table 4 shows that there are differences between three of the four age-groups. For the intervals 30–39 and 40–49, there is no difference. An argument for analysing age-groups separately and not all age-groups together, is differences in sharing-frequency between the groups.

### The Analysis With Binary Logistics

Binary logistic regression is a form of regression analysis. The dependent variable is a dichotomy variable coded as 0 or 1. The independent variables can be of any type, for instance continuous and categorical variables. The four age-group intervals are analysed as separate models.

We first report the four models' statistical summaries that include the Hosmer-Lemeshow test (Archer et al., 2007) with the threshold criteria of  $> 0.05$ . The Hosmer-Lemeshow test is often used as a goodness of fit test.

**Table 5.** Model summaries, the final model for each of the four age-groups.

Age-group 16–29 years - Model 1 summary			
-2 Log likelihood		Cox and Snell R square	Nagelkerke R square
Step 5	208.210	0.51	0.076
Hosmer and Lemeshow Test			
Chi-square		df	Sig.
Step 5	2.990	4	0.560
Age-group 30–39 years - Model 2 summary			
-2 Log likelihood		Cox and Snell R square	Nagelkerke R square
Step 6	213.327	0.22	0.041
Hosmer and Lemeshow Test			
Chi-square		df	Sig.
Step 6	0.926	4	0.921
Age-group 40–49 years - Model 3 summary			
-2 Log likelihood		Cox and Snell R square	Nagelkerke R square
Step 7	158.856	0.033	0.059
Hosmer and Lemeshow Test			
Chi-square		df	Sig.
Step 7	0.618	1	0.432
Age-group 50–69 years - Model 4 summary			
-2 Log likelihood		Cox and Snell R square	Nagelkerke R square
Step 5	110.464	0.057	0.198
Hosmer and Lemeshow Test			
Chi-square		df	Sig.
Step 5	4.724	7	0.694

The Wald statistic (Hosmer and Lemeshow, 1989) is used to identify the significant variables in each of the four models. This is the square of the t-statistic and gives equivalent results for a single parameter. It can be used to test the significance of particular predictors in a statistical model. We chose

backward Wald for selecting how independent variables are entered into the analysis. With backward Wald, all the predictor variables chosen are added into the model, and those variables that do not (significantly) predict anything on the dependent measure are removed, one by one, from the model.

In Table 6 we present step 1 in the binary logistic regression analyses that includes all independent variables.

**Table 6.** Descriptive statistics of the variables and measurements in step 1.

Variable code:	Beta est.	SE	Wald	df	Sign.	Exp (B)
Gender: Female = 0, Male = 1						
Age 16–29	–0.035	0.369	0.009	1	0.924	0.966
Age 30–39	0.059	0.384	0.492	1	0.483	0.857
Age 40–49	0.077	0.440	0.031	1	0.861	1.080
Age 50–69	1.066	0.679	2.468	1	0.116	2.904
Education: three levels (primary, bachelor, master)						
Age 16–29	–0.046	0.310	0.022	1	0.881	0.955
Age 30–39	–0.154	0.220	0.492	1	0.483	0.857
Age 40–49	–0.897	0.311	8.326	1	0.004	0.408
Age 50–69	0.219	0.351	0.390	1	0.532	1.245
Data-disclosure (consent): No = 0, Yes = 1						
Age 16–29	0.124	0.359	0.120	1	0.729	1.132
Age 30–39	0.321	0.380	0.715	1	0.086	0.273
Age 40–49	0.653	0.443	2.174	1	0.140	1.921
Age 50–69	1.179	0.559	4.459	1	0.035	3.252
ID-theft and/or Credit-card misuse: No = 0, Yes = 1						
Age 16–29	0.837	0.417	4.032	1	0.045	2.309
Age 30–39	–1.300	0.758	2.941	1	0.086	0.273
Age 40–49	0.094	0.530	0.032	1	0.859	1.099
Age 50–69	1.894	0.588	10.379	1	0.001	6.644
CRT: all wrong, one correct, two correct, three correct, all four correct						
Age 16–29	0.276	0.157	3.093	1	0.79	1.318
Age 30–39	–0.108	0.185	0.339	1	0.560	0.898
Age 40–49	0.128	0.221	0.338	1	0.561	1.137
Age 50–69	0.694	0.295	5.526	1	0.019	2.002
Self-efficacy, low, medium, high						
Age 16–29	–0.200	0.210	0.913	1	0.339	0.819
Age 30–39	0.263	0.223	1.391	1	0.238	1.300
Age 40–49	0.191	0.254	0.563	1	0.453	1.210
Age 50–69	–0.181	0.323	0.312	1	0.587	0.835
Machiavellianism, from low to high, 1 to 5						
Age 16–29	0.276	0.157	3.093	1	0.079	1.318
Age 30–39	0.198	0.150	1.732	1	0.188	1.219
Age 40–49	–0.266	0.172	2.407	1	0.121	0.766
Age 50–69	–0.101	0.295	0.116	1	0.733	0.904

In reviewing the numbers in step 1, a high Wald-score is an indication of the final model, of what can be expected. In the final step, only the significant predictors (<0.1) remain.

**Table 7.** Descriptive statistics of the variables and measurements in step 1.

Variable code:	Beta est.	SE	Wald	df	Sign.	Exp (B)
<b>Age 16–29 years</b>						
ID-theft and/or Credit-card misuse: No = 0, Yes = 1	0.857	0.407	4.438	1	0.035	2.356
Machiavellianism, 1 to 5	0.327	0.151	4.681	1	0.031	1.387
Constant	-2.224	0.445	24.959	1	<0.001	1.080
<b>Age 30–39</b>						
ID-theft and/or Credit-card misuse: No = 0, Yes = 1	-1.236	0.750	2.715	1	0.099	0.291
Machiavellianism, 1 to 5	0.236	0.141	2.795	1	0.095	1.266
Constant	-2.426	0.465	27.208	1	<0.001	0.088
<b>Age 40–49 years</b>						
Education: three levels (primary, bachelor, master)	-0.650	0.263	6.108	1	0.013	0.522
Constant	-0.608	0.477	1.626	1	0.202	0.544
<b>Age 50–69 years</b>						
ID-theft and/or Credit-card misuse: No = 0, Yes = 1	1.968	0.574	11.757	1	<0.001	7.154
Data-disclosure (consent): No = 0, Yes = 1	1.233	0.550	5.027	1	0.025	3.433
CRT: all wrong to all correct	0.798	0.286	7.803	1	0.005	2.221
Constant	-5.646	0.818	47.632	1	<0.001	0.004

The final model, shown in Table 6, shows the significant predictors for each of the four age-groups. For the 16–29 age-group, there are two significant predictors, ID-theft & credit-card misuse and Machiavellianism, which is a personality trait characterized by manipulateness and deceitfulness. These same two variables are predictors for the age-group 30–39, but only significant on a 0.1-level. For the age-group 40–49, education is a predictor, indicated by those have studied at a higher education institution and those who have not. A negative Beta reveals that it is those who have not studied at higher education are sexting more frequently. Finally, for the age-group 50–69 years, ID-theft & credit-card misuse together with data-disclosure (high willingness to share) and the CRT are significant predictors.

In binary logistic regression, there is one more result that should be reported. This is the classification table, showing correct and incorrect percentages for the groups ‘has shared nudes’ (coded as 1) and ‘has not shared nudes’ within the last 12 months (coded as 0). The higher the percentage, the better the variables function as predictors.

The youngest (70%) and the oldest (77%) age-groups have acceptable predictions. For the two other age-groups the percentages were lower, 57% and 64%.

In general, we assume that personal traits and decision style are of relevance for online behaviour; what people do, for instance, when no one’s watching. The results from the present study are in accordance with previous studies on impulse behaviour, sensation-seeking and potential vulnerabilities.

**Table 8.** Classification table - the binary logistics prediction of the two groups.

		Not sexting (numbers)	Sexting (numbers)	Percentage correct
16–29 years	0	119	35	77.3%
	1	25	22	46.8%
				<b>70.1%</b>
30–39 years	0	138	105	56.8%
	1	16	22	57.9%
				56.9%
40–49 years	0	111	60	64.9%
	1	13	16	55.2%
				63.5%
50–69 years	0	288	85	77.2%
	1	4	12	75.0%
				77.1%

The fact that the respondents are drawn from a panel representing a national population aged 16–69, and not only up to 30 years old, is important, and this is a study that can be replicated. Our results show that sexting is not only a phenomenon among young people, but it is more common amongst the younger generation. Of greater importance, the correlation between financial vulnerability and the sharing of graphic nude images indicates that sexting might be even riskier than anticipated.

## DISCUSSION

In the present study, with a large data set, self-reported online behavior together with personal traits and willingness to share personal data were assessed as predictors of the tendency towards sexting. Our study showed that ID-theft and credit-card misuse were more common among those that reported that they have been sexting compared to the non-sexting group. The ID-theft/credit-card variable was a significant predictor of sexting for the age-groups 16–29, 30–39 and 50–69 years old. For the youngest group, the manipulateness and deceitfulness trait is also a predictor, whereas for the oldest group, the intuitive decision style and a high willingness to share personal data are also significant predictors.

Our results indicate that personal traits do matter, but these findings were less clear-cut than those for the ID-theft and credit card misuse. In planned future studies, we will examine how and to what extent several personality variables, such as and including intuitiveness, are associated with and can therefore predict hazardous online behaviours such as sexting.

## ACKNOWLEDGMENT

This research was supported by Research Council Norway under the grant 310105 (NORCICS).

## REFERENCES

- Archer, K. J., Lemeshow, S., and Hosmer, D. W. (2007). Goodness of fit tests for logistic regression models when data are collected using a complex sampling design. *Computational Statistics & Data Analysis*, 51, 4450–4464.
- Barrense-Dias, Y., Berchtold, A., Suris, J. C., and Akre, C. (2017). Sexting and the definition issue. *Journal of Adolescent Health*, 61(5), 544–554.
- Christie, R. and Geis, F. L. (1970). *Studies in Machiavellianism*. New York NY, Academic.
- Cooper, K., Quayle, E., Jonsson, L., and Svedin, C. G. (2016). Adolescents and self-taken sexual images: A review of the literature. *Computers in Human Behavior*, 55, 706–716. doi: 10.1016/j.chb.2015.10.003.
- Dir, A. L., Cyders, M. A., and Coskunpinar, A. (2013). From the bar to the bed via mobile phone: A first test of the role of problematic alcohol use, sexting, and impulsivity related traits in sexual hookups. *Computers in Human Behavior*, 29, 1664–1670.
- Dodaj, A., and Sesar, K. (2020). Sexting categories. *Mediterranean Journal of Clinical Psychology*, 8(2).
- European Union, (2017). 5661. Special Eurobarometer 464a “European attitudes towards cyber security.” September 2017.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25–42.
- Hosmer, W. and Lemeshow, S. *Applied logistic regression*. (1989). New York: Wiley.
- Klettke, B., Hallford, D. J., Clancy, E., Mellor, D. J. and Toumbourou, J. W. (2019). Sexting and Psychological Distress: The Role of Unwanted and Coerced Sexes. *Cyberpsychology, Behavior, and Social Networking*. Apr 2019. 237–242.
- Krosnick, J. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys, *Applied Cognitive Psychology*, 5 3, 213–36.
- Luthans, F., Youssef, C. M., and Avolio, B. J. (2007). *Psychological capital: Developing the human competitive edge*. Oxford: Oxford University Press.
- Machimbarrena, J. M., Calvete, E., Fernández-González, L., Álvarez-Bardón, A., Álvarez-Fernández, L., and González-Cabrera, J. (2018). Internet risks: An overview of victimization in cyberbullying, cyber dating abuse, sexting, online grooming and problematic internet use. *International Journal of Environmental Research and Public Health*, 15(11), 2471.
- MacKenzie, S. B. and P. M. Podsakoff (2012). Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies. *Journal of Retailing*, 88: 542–555.
- March, E. and Wagstaff, D. L. (2017) Sending Nudes: Sex, Self-Rated Mate Value, and Trait Machiavellianism Predict Sending Unsolicited Explicit Images. *Front. Psychol.* 8: 2210.
- Rauthmann, J. F. (2013) Investigating the MACH-IV With Item Response Theory and Proposing the Trimmed MACH\* *Journal of Personality Assessment*, 95:4, 388–397.
- Reyns, B. W., Henson, B. and Fisher, B. S. (2014) Digital Deviance: Low Self-Control and Opportunity as Explanations of Sexting Among College Students. *Sociological Spectrum: Mid-South Sociological Association*. Vol. 34, (3): 273–292.
- Tjostheim I. and Waterworth J. A. (2020). Predicting Personal Susceptibility to Phishing. In: Rocha Á., Ferrás C., Montenegro Marin C., Medina García V. (eds.) *Information Technology and Systems. ICITS 2020. Advances in Intelligent Systems and Computing*, vol 1137. Springer, Cham. [https://doi.org/10.1007/978-3-030-40690-5\\_54](https://doi.org/10.1007/978-3-030-40690-5_54)
- Toplak, M.E. West, R.F., and Stanovich, K.E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20:2, 147–168, DOI: 10.1080/13546783.2013.844729



# GoodMaps: Assessing an Indoor Navigation App Built on Camera-Based Positioning

**Jennifer Palilonis**

Ball State University, Muncie, IN 47304, USA

## ABSTRACT

This paper presents GoodMaps, an AI-driven indoor navigation tool, and chronicles findings from formative focus groups and user acceptance testing with a diverse group of participants. GoodMaps is built on camera-based positioning integrated with a smartphone app that provides turn-by-turn navigation indoors. Originally designed to support blind and low-vision users in independent indoor navigation, GoodMaps was redesigned in 2023 to help all people navigate safely and efficiently with dynamic routing instructions, orientation aids, visual maps, augmented reality, and landmark recognition. This paper shares key challenges, tensions, and opportunities in designing assistive tools for differently-abled users and app-based navigation solutions.

**Keywords:** Accessibility design, Assistive technology, Indoor wayfinding

## INTRODUCTION

Being able to travel independently is critical to maintaining quality of life and participation in a community (Lindsay & Lamprey, 2019). However, doing so can be challenging for individuals who face physical or neurological disabilities. Of course, the outdoor world is effectively mapped, with apps like Google Maps, Apple Maps, MapQuest, and Waze providing accessible navigation support. A variety of techniques and devices also provide navigation support for indoor spaces. For example, Bluetooth trilateration is used to calculate someone's position and relies on a known distance between two or three reference points marked by installed beacons. Despite efforts to provide indoor navigation similar in ease of use and accuracy to Apple and Google Maps, these efforts are in their infancy.

Outdoor navigation is governed by GPS, which is relatively inaccurate and/or difficult to consistently implement indoors. GoodMaps, an AI-driven indoor navigation app built on a camera-based positioning system (CPS), implements geo-referenced images to determine a user's position. CPS can locate where someone is in a room within one meter of accuracy, a substantial improvement over GPS (10 meters outdoors), Bluetooth trilateration (five meters), and other approaches. To date, GoodMaps has mapped just over 23 million square feet of indoor space around the world, with live mapped square footage at just under 19 million ft<sup>2</sup>. U.S. availability includes train

stations, libraries, shopping malls, and sporting arenas, with that scope broadening. Three major airports in the U.S. are also mapped– Portland, OR, Louisville, KY, and Westchester, NY – which collectively serve about two million passengers per month. Outside the U.S., the app is available throughout the U.K. – including in large swaths of the country-wide rail network – in Canada, Continental Europe, and Central and South America, where users can access retail spaces, transit hubs, and corporate buildings. Originally designed to support blind and low-vision (BLV) users, GoodMaps was redesigned in 2023 to support indoor wayfinding for all by providing dynamic routing instructions, orientation aids, and landmark recognition.

In 2023, two focus groups and four rounds of User Acceptance Testing (UAT) were conducted with 100 participants to collect feedback about app design, functionality, and overall user experience. Focus groups were remote (conducted via Zoom) with 53 users. The first focus group included sighted participants who were given control of an interactive prototype and asked to provide feedback about interface design and feature functionality. About one-half of those participants identified as: anxious/nervous travellers, hearing impaired, neurodiverse, and/or wheelchair user. The second focus group included BLV users who provided insights about challenges they face with indoor navigation. A beta app was developed, and UATs were conducted with 47 individuals who are BLV (including guide dog users and white cane users), hearing impaired, neurodiverse, and mobility impaired, as well as users with no reported disability. These UATs were held at a museum in Louisville and were designed to collect feedback about the live routing experience, app functionality, and app design. All UATs resulted in iterative improvements to the user interface, user experience, and system functionality. Data from these sessions helped identify instances in which:

- users struggled to understand the nature and/or functionality of a feature;
- design was confusing, unattractive, and/or detrimental to user experience.

## **REVIEW OF BACKGROUND**

Indoor navigation is a technology-supported activity in which users can travel to specific indoor locations. Indoor navigation typically involves five key components: 1) generating navigation models from 3D building models or indoor maps; 2) identifying a user's location in a building through indoor positioning; 3) path-planning to align building models with location coordinates, spatial semantics, and room properties; 4) turn-by-turn navigation instructions based on routing visualizations; and 5) smartphone apps/other devices to support spatial cognition, wayfinding, and orientation (Liu, Li, Zlatanova & Oosterom, 2021).

### **State-of-the-Art in Indoor Navigation**

Accessible GPS devices introduced in the past 15 years, as well as development of smartphone and navigation applications, has enabled people

to easily navigate to desired destinations. However, GPS-enabled apps and devices are limited to outdoor navigation, rendering them ineffective for indoor wayfinding. Most assistive technologies developed to help people navigate indoor spaces require renovations to the space or installation of custom proprietary devices (Faria et al., 2010; Nakajima & Haruyama, 2012; Ahmetovic et al., 2016).

By 2024, more than 7 billion people worldwide reported owning a smartphone. Likewise, the use of assistive apps to help individuals with disabilities reach heightened levels of autonomy is on the rise (Abraham et al., 2022). Studies have shown that individuals with disabilities frequently use apps designed to help them accomplish daily activities (Griffin-Shirley et al., 2017). Several solutions exist for indoor mapping. However, there are a variety of approaches to the complex nature of dynamic wayfinding. Table 1 describes popular indoor wayfinding apps and includes how each differs from GoodMaps. A few additional solutions have existed, including Clew (users retrace a previously recorded route); Soundscape, a Microsoft solution sunsetted in 2022; and Nearby Explorer, sunsetted in 2019.

**Table 1.** State-of-the-art in most commonly used indoor navigation apps.

App Name & Description	Differentiation
<b>Hyper:</b> Indoor navigation powered by Augmented Reality sensors and WiFi	<ul style="list-style-type: none"> <li>• SDK approach requires Wi-Fi</li> <li>• No venue management system</li> </ul>
<b>Lazarillo:</b> App guides users through buildings with real-time voice messages	<ul style="list-style-type: none"> <li>• Requires installation of Bluetooth beacons for navigation</li> </ul>
<b>Lazarus:</b> App provides navigation assistance through geolocation, voice recognition, and device camera	<ul style="list-style-type: none"> <li>• No built-in navigation support</li> <li>• Android-only</li> </ul>
<b>Pointr:</b> AI-powered, beacon-based navigation and mapping platform; provides blue-dot location and visual map with routing information	<ul style="list-style-type: none"> <li>• No audio/voice instructions</li> <li>• Requires installation of Bluetooth beacons for navigation</li> <li>• No support for people with disabilities</li> </ul>
<b>RightHear:</b> App + portal for integration of audio descriptions of a location	<ul style="list-style-type: none"> <li>• Requires installation of Bluetooth beacons for navigation</li> </ul>
<b>Waymap:</b> Indoor navigation app provides turn-by-turn navigation support	<ul style="list-style-type: none"> <li>• Does not require any external signals</li> <li>• User must download venue maps</li> </ul>

### Independence, Self Determination, and Assistive Technology

Access to and use of assistive technology has had a positive psychosocial effect on feelings of competence, adaptability, and self-esteem among users with disabilities (McNicholl, Desmond & Gallagher, 2023). Studies have also demonstrated that when students with disabilities use assistive devices and apps in educational settings, it promotes heightened feelings of academic

self-efficacy (Depountis, Okungu & Molloy-Daugherty, 2019). Likewise, assistive technology has been found to “promote participation and inclusion in society, and support access to health, social services, education, work and other important life experiences for persons with disabilities, older people, and those with chronic conditions” (Smith, et al., 2024). Recently, organizations like WHO and UNICEF have called for improved access to assistive tech globally, asserting that it is critical to the achievement of differently abled people (UNICEF, 2022). When it comes to navigation and wayfinding, inaccessible spaces can have a profoundly negative affect on quality-of-life, health and safety, independence, and social participation among differently abled individuals (Kapsalis, Jaeger & Hale, 2024). Specifically, mobility, vision, and neurological impairments may lead to substantial disruptions in the ability to freely and easily move between places.

## METHODOLOGY

User Acceptance Testing (UAT) is a user-centered phase of development in which a digital product is tested in the real world by its intended audience. The GoodMaps app was developed through multi-phased design and development process that included two focus groups and four rounds of UAT conducted over the course of a year as part of an iterative process. Early research informed the vision, design, and development of the initial feature set. Subsequent UATs informed design and functionality and led to the identification of features for future development.

### GoodMaps

First launched in September 2020, GoodMaps is a smartphone app for iOS and Android devices that provides turn-by-turn navigation for indoor spaces. The app is designed to help people navigate safely and efficiently with dynamic routing instructions, orientation aids, and landmark recognition. To create an accessible building map, a GoodMaps technician scans a facility with a LiDAR camera that captures 360-degree images, measurements, and video. Second, an accurate, detailed map is created from the scan data. Points of interest—like restrooms, offices, dining, fire extinguishers, exits, etc.—are tagged in GoodMaps Studio, which processes and hosts all map data. Studio generates three map views: a 2D floorplan, a LiDAR point cloud, and a 3D model. Third, building supervisors can add location names in Studio, update points of interest, and set up access permissions for the map. Fourth, the building map is published to the GoodMaps navigation app. Once a map is published, app users can navigate indoor spaces using the app along with voice, tactile (haptic vibrations), or text prompts to find their desired destinations. Buildings mapped with this technology are explorable via the app, which includes step-by-step navigation to selected destinations. CPS determines a users’ position in a building, and the app offers information about indoor points of interest. The app relies on visually distinct surroundings, such as patterned carpets, ceiling panels, and art to provide accurate location information.

## Procedures & Participants

Early iterations of the GoodMaps app (formerly called *Explore*) were designed for use by individuals who are blind and/or visually impaired. In 2023, the app was redesigned to extend its reach to all people, with special attention to features for a wide range of ability. Table 2 outlines the nature of three phases of user-centered research and development for this redesign:

1. **Discovery Research** included two focus groups designed to understand sighted users' indoor navigation needs and to identify key features.
2. **Formative Research** included two sets of remote UAT interviews during which sighted participants provided feedback on an app prototype.
3. **Summative Research** included UATs with diverse users at a museum.

**Table 2.** Research participants and procedures.

Discovery Phase	
<b>Focus Group 1</b>	
3 remote sessions, 9–10 users each	N = 28 – No disability: 15; Neurodiverse: 8; Wheelchair: 5
<b>Focus Group 2</b>	
2 remote sessions, 12–13 users each	N = 25 – Blind: 15; Low vision: 10
<b>UAT 1</b>	
Remote interactive prototype testing	N = 13 – No disability: 7; Neurodiverse: 3; Wheelchair: 2
<b>UAT 2</b>	
Remote interactive prototype testing	N = 12 – No disability: 5; Neurodiverse: 3; Wheelchair: 3
Summative Phase: Sessions required each participant to provide feedback about specific features after completing tasks while using the app.	
<b>UAT 3: In-person task-based testing</b>	N = 10 – No disability: 6; Neurodiverse: 4
<b>UAT 4: In-person task-based testing</b>	N = 12 – Blind: 8; Low vision: 4

## Data Analysis

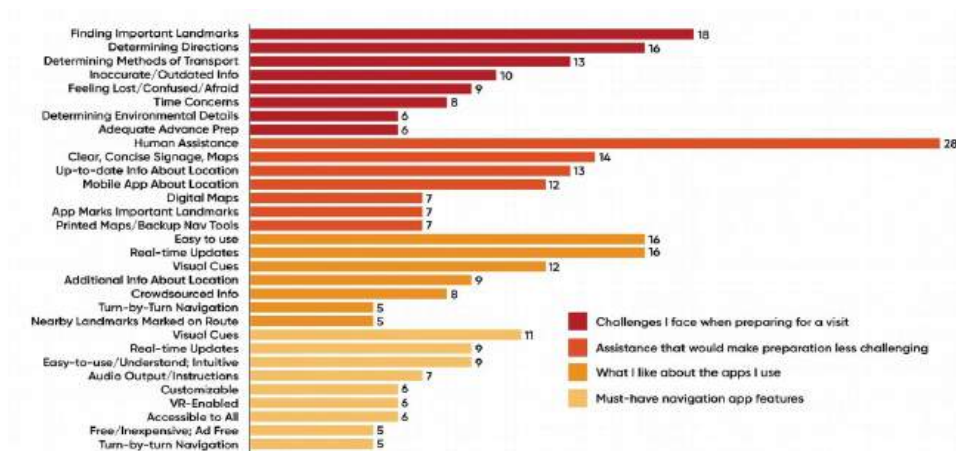
Each research session generated qualitative interview data, and task-based UATs generated observational and interview data. Although results from these sessions does not yield statistically significant results, interview and observational data analysis allows for the abstraction of raw data into more general insights. Because qualitative data is inherently unstructured, interview responses were assigned thematic tags to structure and help synthesize the data. Observational notes collected during UATs helped illuminate common user behaviors while using the app to travel routes in real-world settings. Thus, transcripts from semi-structured interviews and observational notes were coded to illuminate key themes related to participants' understanding of app functionality, user experience, and design preferences. Each unit of analysis—defined as a full sentence or sentence fragment that represented a complete thought—was assigned a code to represent a summative, salient, and/or essence-capturing attribute (Saldana, 2009).

## RESULTS

After coding all individual responses and observational notes, common user experience themes and recommendations for improvement emerged based on cross-participant commonalities.

### DISCOVERY PHASE: FOCUS GROUPS

The first round of focus groups explored needs and perspectives of sighted (no disability), wheelchair, and neurodiverse users. Figure 1 illustrates key themes that emerged from focus groups related to challenges users face when preparing to visit a new venue, assistance that would make preparation less challenging, features they appreciate in frequently used apps, and must-have navigation app features.



**Figure 1:** Key themes from sighted user focus groups.

A second focus group included BLV users and was designed to elicit feedback about challenges faced when visiting unfamiliar destinations, desired assistive tools, and features that motivate use of navigation apps. Participants reported using a wide range of apps for navigation. They also noted that most of these apps were built for the sighted world, with BLV users an afterthought (at best), making them cumbersome to use. Several additional themes surfaced:

- Traditional navigation apps don't provide enough information about a BLV user's surroundings. However, too much information delivered via screen readers or text-to-speech may result in information overload.
- Interface design is often not optimized for screen reader interaction, making apps difficult to navigate for BLV users.
- BLV users cannot confirm they are in the "right" place by sight, so concerns exist about the location accuracy a navigation app can provide.
- BLV users are concerned about safety, as well as how they are perceived by others who see them using apps for navigation in public spaces.

- Other recommended features included: tutorials on how to use the app; the ability to choose routes that avoid stairs; more built-in building/destination information; tones, haptics, and voice for routing confirmation; location sharing with others using the app; the ability to explore a venue ahead of a visit; and support for finding human assistance in a building.

### **Formative Phase: UATs 1 & 2**

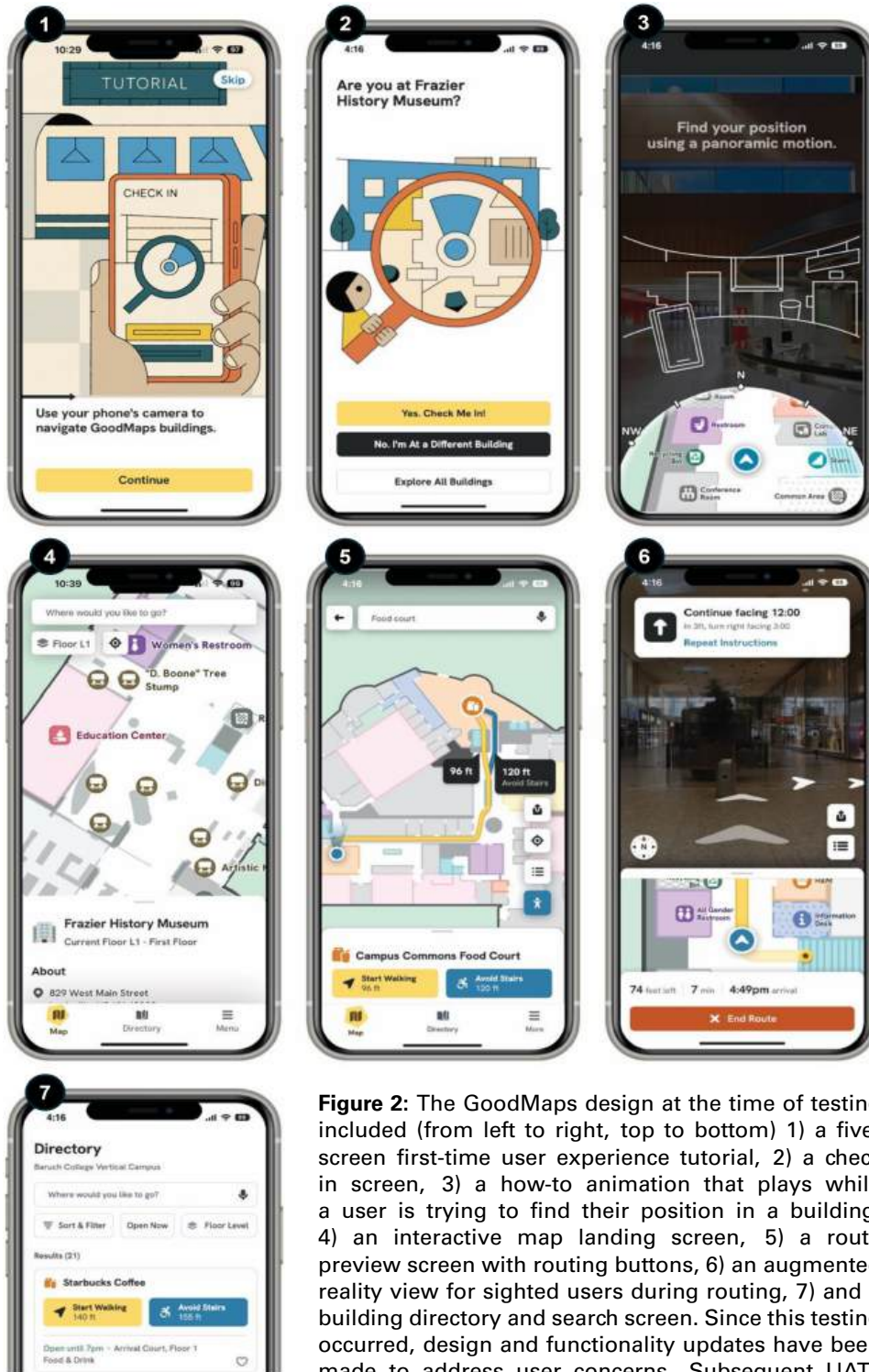
UATs 1 and 2 focused on features designed for sighted users. UAT 1 focused on the first-time user experience tutorial, building and destination information cards, and map/routing screens. UAT 2 focused on localize by checking in, notifications, directory and search, and augmented reality (AR) view. Participants were given remote control access to a prototype built in Axure RP, asked to examine each screen, and then asked to complete interaction tasks. Two key themes emerged:

**Participants sometimes struggled to understand the functional complexity associated with camera-based positioning (CPS).** In fact, participants often had to be reminded that the app relies on the phone's camera, not GPS for routing. This is significant because CPS functionality is directly tied to interactions required to effectively use the app. For example, a user must hold the phone vertically and scan slowly left and right so that the phone's camera can find their position (aka, "localize") in a building. If a user doesn't understand how CPS works, they may not understand how to hold the phone to achieve this goal.

**Some UI elements were also difficult to understand.** Specifically, some participants had trouble interpreting a few of the tutorial screens. Based on those screens, some thought they had to open the phone's camera separately to use the app. Also, some users thought tutorial illustrations were interactive instead of being static visual representations of how the app works. Several users also didn't understand the nature of a "re-center" button, which allows them to center the screen on a blue dot that represents their location on the map. The button that allows users to switch to a list view of routing instructions was also difficult for some to understand. Based on UAT 1 and 2 results, designs were updated, and a beta app was released in the Apple and Google Play stores for more testing.

### **Summative Phase: UATs 3 & 4**

The next two rounds of UAT were conducted at a museum in Louisville, Kentucky. UAT 3 included sighted participants, while UAT 4 included BLV participants. During individual, task-based navigation sessions, participants were asked to provide feedback about the first-time user experience, the check in process, find my position (localizing), destination search and directory, map view, augmented reality view, routing, app output (voice, haptic, tone), favorites, and the overall user experience. Figure 2 illustrates key screens users encountered during these task-based sessions. Five key themes emerged from UATs 3 and 4.



**Figure 2:** The GoodMaps design at the time of testing included (from left to right, top to bottom) 1) a five-screen first-time user experience tutorial, 2) a check in screen, 3) a how-to animation that plays while a user is trying to find their position in a building, 4) an interactive map landing screen, 5) a route preview screen with routing buttons, 6) an augmented reality view for sighted users during routing, 7) and a building directory and search screen. Since this testing occurred, design and functionality updates have been made to address user concerns. Subsequent UATs occur regularly to identify additional improvements and new features.



Both sighted and BVI users were uncertain about how to properly hold the phone while routing. Although the tutorial states that the phone must be held vertically, with the camera facing outward, sighted and low-vision users often let the phone drift to the left and right. Because the phone needs to be pointed in the direction a user is walking to provide accurate routing directions, this sometimes resulted in conflicting instructions. This issue was even more prevalent with completely blind users, who often struggle with the spatial awareness required to understand how to hold the phone. In extreme cases, users raised the phone sideways or held the camera backward. Sometimes blind participants blocked the camera with fingers. The most common BLV tendency was to let the wrist droop/tilt sideways while walking (Figure 3).



**Figure 3.** Common blind user phone positions.

Low-vision users expressed concern that some UI elements were too small and/or did not exhibit enough visual contrast. Although the app was designed to meet Web Content Accessibility Guidelines (WCAG) related to color contrast and UI sizing, low-vision users still had trouble discerning key elements on a smartphone screen, especially when screens contained a lot of visual information.

Some affordances – the floor switcher and routing buttons – were difficult to understand. Users noted that the standard labelling for the floor switcher, which included an icon and the floor number (Figure 4), didn't include enough information to be intuitive at first glance. Likewise, users reported that the “start walking” and “avoid stairs” buttons (see Figure 2) necessary to start a route were confusing because they don't follow routing paradigms already established by apps like Google and Apple maps (typically a GO button). They also shared that it wasn't immediately obvious that tapping one of these is necessary to start a route. Rather, some users thought the route should start as soon as they select a destination.



**Figure 4.** Floor switcher.

Several participants noted that panoramic scanning to localize took too long. If it took longer than 5 to 6 seconds for the app to find a user's position, they began to wonder if the app was broken and/or became impatient. If it took longer than 10 to 12 seconds, impatience became frustration, and users expressed that if it takes much longer than this, they would not want to use the app.

Many participants shared that the voice speed of routing instructions is too fast, making it difficult to mentally process while routing. BLV users – who are often able to process faster language speeds due to a heightened sense of hearing and language processing – were typically less concerned about voiceover/text-to-speech speeds. However, sighted users often noted that the default language speed was too fast for them to quickly process while in route.

The discussion that follows expands on how these findings can inform the design of other apps and assistive tools and includes information about how the GoodMaps app has evolved based on this research.

## DISCUSSION

It's worth noting that feedback about the GoodMaps app was generally positive, especially among individuals with disabilities. One focus group wheelchair user shared, "...you can't imagine how often I struggle to get where I'm going because I have to look for elevators or ramps. This app would make it so much easier for me to get where I want to go quickly, and without help." Likewise, according to one blind UAT participant, "...this app is magical. It's unreal. This could be a life changer. If I had access to this in the places I travel – like the grocery, or museums, or hospitals – I could go anywhere I wanted. I know that seems small to you because you can go anywhere you want any time you want. I have never had that kind of independence. This app would change my life." Users with no reported disability also shared positive feedback, indicating the GoodMaps app would prove useful when exploring an unfamiliar building in advance of a visit, navigating new spaces, and traveling independently. However, designing an assistive tool with the level of functional and design complexity inherent for indoor wayfinding requires a careful balance between intuitive design, clear communication of product affordances, managing user expectations, and understanding diverse user needs. In particular, three findings from this research allude to learnings that can be generalized for assistive tech designers and developers.

First, explaining how novel, complex technologies works in ways that are easy to understand and efficient for app design and experience can be tricky. This is evidenced in two key results from this study: 1) users generally don't understand how CPS works, and 2) they struggle to properly hold the phone. As previously noted, users *must* have some understanding of how CPS works to understand why certain actions are required by the GoodMaps app. This is intensified by the fact that most users are familiar with GPS, which functions very differently. Thus, for the user to understand how to use the app, they must also understand that the phone's camera – not a satellite signal or Bluetooth beacons – plays an integral role in the app's ability to find their position on the map and provide turn-by-turn directions to a destination. Relatedly, how a user holds the phone affects the app's ability to provide accurate routing information. All this leads to a wicked design challenge: What is the best way to communicate those critical, complex concepts on small screens? The answer to that question certainly depends on the product itself. But for GoodMaps, it has meant multiple iterations of tutorial design, visualization, and traditional instructions. In Fall 2024, GoodMaps will add an interactive tutorial that allows users to practice using the app. The interactive tutorial will provide visual, voice, haptic, and tonal feedback to alert users when they are "doing it wrong" and to confirm when they are "doing it right."

Second, balancing requirements and/or limitations of a particular technology with users' expectations for how it should work can also be challenging. For example, designing for smartphones is often impeded by the need to communicate many complex concepts (sometimes all at once) on a very small screen. On face value, this may seem like usability design 101. However, this really highlights a more advanced UI/UX design concept that relates to how designers decide what can be delivered at a higher level of abstraction and what needs to be more blatantly communicated to ensure user understanding. In other words, sometimes an abstract icon is perfect for communicating a complex concept on a small screen. However, when abstraction leads to confusion, more visual, textual, and/or verbal modes of communication may be necessary. Although this approach isn't likely to be space conservative, it may be the only way to ensure understanding. As such, future versions of the GoodMaps design include more detailed labelling on affordances like the floor switcher, routing buttons, and more.

Third, even when accessibility guidelines were followed for UI design, low-vision users in particular still struggled to discern information on maps and routing screens, with some calling for more color and sizing contrast than even WCAG requires. This illuminates the critical value that testing with differently abled users brings to the design and development process. The GoodMaps app represents a relatively complex visual interface, suggesting that perhaps, the more complex a small-screen interface is, the more a design should compensate with increased levels of visual contrast, spacing, sizing, etc.

It is worth noting that many research-informed improvements have been made to the GoodMaps interface since this study was conducted. These include: decluttering of the visual interface when possible, optimizing voice speed for all users, faster localization speeds (now 2 to 3 seconds), revised routing buttons to mirror familiar design paradigms (GO button), enhanced tutorials, redesigned UI elements for better recognition and recall, and more.

## CONCLUSION & FUTURE RESEARCH

Since the completion of this research, the GoodMaps app has undergone a continuous, iterative design and development to improve the experience. The design team engages in an Agile process, through which improvements are made and the app is updated in regular, two-week sprints. Future research will explore the extent to which indoor wayfinding support affects disabled users' feelings of independence and self-determination, as well as their quality of life and work.

## REFERENCES

- Abraham, C. H., Boadi-Kusi, B., Morny, E. K. A. and Agyekum, P., 2022. Smartphone usage among people living with severe visual impairment and blindness. *Assistive Technology*, 34(5), pp. 611–618.
- Ahmetovic, D., Gleason, C., Ruan, C., Kitani, K., Takagi, H. and Asakawa, C., 2016, September. NavCog: A navigational cognitive assistant for the blind.

- In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 90–99).
- Depountis, V., Okungu, P. and Molloy-Daugherty, D., 2019. The development of an application that supports body awareness for children with visual impairments and additional disabilities. *Vision Rehabilitation International*, 10(1), pp. 1–8.
- Faria, J., Lopes, S., Fernandes, H., Martins, P. and Barroso, J., 2010, September. Electronic white cane for blind people navigation assistance. In *2010 World Automation Congress* (pp. 1–7). IEEE.
- Griffin-Shirley, N., Banda, D. R., Ajuwon, P. M., Cheon, J., Lee, J., Park, H. R. and Lyngdoh, S. N., 2017. A survey on the use of mobile applications for people who are visually impaired. *Journal of Visual Impairment & Blindness*, 111(4), pp. 307–323.
- Kapsalis, E., Jaeger, N. and Hale, J., 2024. Disabled-by-design: effects of inaccessible urban public spaces on users of mobility assistive devices—a systematic review. *Disability and Rehabilitation: Assistive Technology*, 19(3), pp. 604–622.
- Lindsay, S., & Lamptey, D. L. (2019). Pedestrian navigation and public transit training interventions for youth with disabilities: A systematic review. *Disability and rehabilitation*, 41(22), 2607–2621.
- Liu, L., Li, B., Zlatanova, S. and van Oosterom, P., 2021. Indoor navigation supported by the Industry Foundation Classes (IFC): A survey. *Automation in Construction*, 121, p. 103436.
- McNicholl, A., Desmond, D. and Gallagher, P., 2023. Assistive technologies, educational engagement and psychosocial outcomes among students with disabilities in higher education. *Disability and Rehabilitation: Assistive Technology*, 18(1), pp. 50–58.
- Nakajima, M. and Haruyama, S., 2012, August. Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing. In *2012 1st IEEE International Conference on Communications in China (ICCC)* (pp. 524–529). IEEE.
- Saldana, J., 2011. *Fundamentals of qualitative research*. Oxford university press.
- Smith, E. M., Huff, S., Wescott, H., Daniel, R., Ebuenyi, I. D., O'Donnell, J., Maalim, M., Zhang, W., Khasnabis, C. and MacLachlan, M., 2024. Assistive technologies are central to the realization of the Convention on the Rights of Persons with Disabilities. *Disability and Rehabilitation: Assistive Technology*, 19(2), pp. 486–491.
- UNICEF (2022) *Global report on assistive technology*. Available at <https://www.unicef.org/reports/global-report-assistive-technology> (Accessed 11 March 2024).

# Deep Surface Liquid Crystal Displays for Extended Reality Applications

Ian Gonsher<sup>1</sup>, Maya Magavi<sup>1</sup>, Mina Kim<sup>2</sup>, Monica Bhyrappa<sup>2</sup>,  
Claire Yang<sup>1</sup>, Selia Jindal<sup>1</sup>, and John Finberg<sup>1</sup>

<sup>1</sup>Brown University, Providence, RI 02912, USA

<sup>2</sup>Rhode Island School of Design, RI 02903, USA

## ABSTRACT

Representing depth on a surface has been a design concern since at least the Renaissance. And yet, despite this long history, 3D displays have not become widely adopted. However, with the emergence of mixed reality applications, the development of new modalities for providing depth to displays has become even more relevant. We present a novel design that addresses these challenges by employing layered transparent LCD screens in configuration, which we call Deep Surface Liquid Crystal Displays (DS-LCD). DS-LCDs require no additional artifacts such as anaglyph glasses. Drawing on both stereoscopic and volumetric approaches, we present the DS-LCD prototype and discuss the visual elements that create the “deep surface” effect. We then describe the initial software applications that were developed to validate the prototype, as well as our initial user studies. We conclude with speculation about future work and applications.

**Keywords:** Extended reality, Augmented reality, Mixed reality, 3D display, Stereoscopic display, Volumetric display, Liquid crystal display

## INTRODUCTION

The ability to create the illusion of depth upon a two-dimensional surface has long been a concern of artists and technologists interested in the verisimilitude of visual representation. Painting was arguably the first attempt in the modern era to describe what we might now call virtual or augmented reality. Painters were and are world makers. By the early 15<sup>th</sup> century, artists were developing new strategies for representing deep, volumetric space on flat surfaces. These works of art are now esteemed because one finds, in the best instances, a masterful integration of linear perspective, atmospheric perspective, relative size, and chiaroscuro that create a space within which a story can take place.

In the 20<sup>th</sup> century, with the development of film and television, moving images could be orchestrated into time-based experiences that added greater realism than static two-dimensional representations such as painting. As these technologies matured, storytellers and technologists continued to explore ways of creating more immersive experiences.

Since then, the most common strategy for creating 3D effects has been the use of stereoscopic images, which produce a 3D effect by taking advantage of

binocular vision (Kooi and Toet, 2004). This is usually achieved by splitting the image into opposing colors, typically red and cyan, and filtering each color through anaglyph glasses. When the images are reassembled in the viewer's perception, it gives the illusion of the image popping off the surface. However, while anaglyph glasses have become widespread, public sentiment has not been overwhelmingly positive. A 2023 survey shows that demand for traditional 3D technologies has tapered off; however, it projects that with the integration of improved, anaglyph glasses free technologies, global demand, particularly driven by Asian markets, is projected to nearly quadruple for 3D cinema (YouGov, 2011). Many virtual reality headsets use variations of this technique, and therefore require the use of wearables, typically mounted onto the head and in front of the eyes.

Autostereoscopy, which does not require the use of anaglyph glasses or other similar wearables, employs a similar technique. Glasses are unnecessary because lenticular lenses or parallax barriers on the surface separate the image for each eye to process based on relative position. Binocular vision produces the effect relative to the position of the viewer. Because these kinds of screens are dependent on specific viewing zones and the relative position of the viewer, outside of which the effect is not seen, autostereoscopy also has significant limitations (Mphepo, 2020).

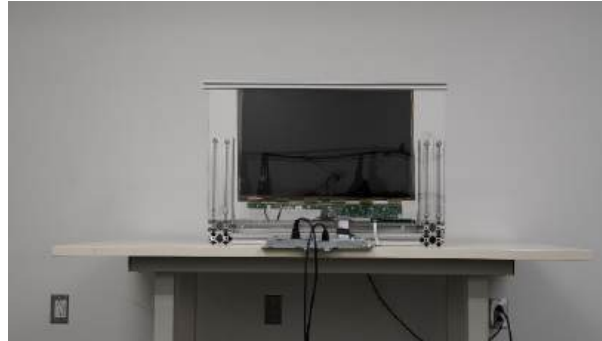
Volumetric displays and other holographic effects, such as Pepper's Ghost, are yet another approach to creating 3D user experiences (England, 2018; Gonsler et al., 2023). Typically, some kind of substrate is needed to project upon, which need not be flat, but does need to maintain a calibrated position between the viewer and the projected image. Like stereoscopic images, volumetric displays also have limitations in terms of viewing angles and resolution.

Deep Surface Liquid Crystal Displays (DS-LCD), as we call the prototypes presented here, offer an alternative design strategy with several advantages. DS-LCDs can be used as a kind of volumetric display in the z axis, employing the shallow space produced in the gap between the two displays. The two (or more) layers work as substrates upon which slices of the image can be produced. But DS-LCDs can also work stereoscopically, in the y and x axes, by taking advantage of binocular effects. It doesn't require anaglyph glasses, and so it can produce a deep surface experience that is integrated directly into the built environment. This strategy eschews the pitfalls of wearables in general, such as cumbersome AR and VR devices that attach to the head, and returns to the paradigm that painters developed centuries ago.

### **Prototype: Hardware**

Deep Surface Liquid Crystal Displays work much like a traditional display. However, instead of a single LCD display, our prototype features two LCD displays, one layered on top of the other with a small gap in between. The rear LCD screen is a standard unmodified LCD display. The front LCD display is modified by removing the opaque backing layer, allowing for the liquid crystal layer and polarizing layers to remain transparent (Figure 1 and Figure 2). This allows for the layering, or doubling, of additional visual

content onto the background screen. These LCD displays are mounted onto an adjustable aluminium frame, which allowed us to vary and test the size of the gap between screens. Initial testing determined that a gap of between 1.75 inches and 2.75 inches between the screens was most optimal for realizing the 3D effect (Figure 3).



**Figure 1:** Prototype of the DS-LCD.



**Figure 2:** DS-LCD prototype viewed from front.



**Figure 3:** DS-LCD prototype viewed from side and rear.

### Prototype: Software

In order to understand how media might be displayed onto the layered displays, and best enhance the depth of field effect, we developed software that could: 1) Play two videos files (e.g. mp4) on the two displays simultaneously. 2) Separate the background field from the foreground image using OpenCV, and play both videos simultaneously on both displays. 3) Play two video files in sequence in order to create a “leaping” effect of a smaller background image moving forward toward the viewer and becoming a larger foreground image on the transparent display.

The first Python program - *videoplayer* - simply played two video files at the same time on the two displays. This was accomplished by creating two windows for each video, defining the x and y coordinates for each window, and resizing and aligning each frame of the video to conform to the correct frame dimensions. This first iteration of the prototype allowed for the further development and enhancement of the deep surface effect.

The second Python program – *videoextractor* – took in a file path to read to and a file path to write to, then went through each frame of the input file and removed the background using the OpenCV module cv2. The extracted foreground image could then be played on the front transparent display, while the background image could then be played on the rear display. In videoconferencing applications, this separates the image of a person from the background, giving a deeper spatial effect (Figure 4).



**Figure 4:** Two displays, one transparent and one opaque, are layered in configuration to produce greater depth (depth not to scale).

The third Python program – *videoleaper* – split a video into two files that could be played on both screens sequentially. A smaller image on the rear screen jumps forward onto the front screen, increasing slightly in size to further enhance the effect of an object coming towards the viewer. We initially tested this effect by creating a simulation of a bouncing ball that “leaped” between screens at five second intervals, growing larger on the front display



and smaller on the rear display. From this initial test, we developed and tested other media to further enhance the effect (Figure 5 and Figure 6).



**Figure 5:** The plane leaps from the rear display to the front display.



**Figure 6:** The ball leaps from the rear display to the front display and back to the rear display in a circular motion.

Using these programs, a series of demos were developed, including:

1. Video conferencing demo: Using *vidoe extractor*, a local user engages with a remote user in a typical video conferencing scenario. The video image of the local user was separated into a foreground image and a background image, each displayed on its respective display.
2. Airplane animation demo: Using a single video file and *videoleaper*, a plane flying forward jumps from the rear display to the front display, slightly enlarging the image, and moving from display to display.
3. Ball animation demo: A bouncing ball jumps to the front screen as it enlarges and back to the back screen as it shrinks again using the same technique as the airplane animation demo.

## USABILITY TESTING

Ten graduate and undergraduate students from the Rhode Island School of Design (RISD) and Brown University participated in the user tests and surveys. They possessed a basic understanding of 2D digital screens and an interest in digital experiences. None of the participants had a disability that might affect the user tests. The user tests were conducted twice with different groups, and survey forms were distributed to all participants.

## Methodology

The user testing was conducted in a dark room to ensure optimal performance of the DS-LCD prototype, with participants seated in chairs positioned in front of the DS-LCD. Two of the user tests were conducted under the same conditions. Participants were given a general introduction to the DS-LCD

project, explaining that the display system is made of two layers of 2D digital screens, enabling depth beyond that of a flat screen and providing an enhanced dimensional experience. Two scenarios were presented to the user groups.

The first scenario presented subjects with a video conferencing experience. We asked participants to have a 2 to 3 minute video conference conversation, with one participant using a remote laptop in an adjacent room, and one participant using the prototype. Subjects were asked to compare the video conferencing experience of a traditional display with the experience of using a DS-LCD. Users followed up by filling out a survey. Users were asked if a more volumetric presentation of remote users enhanced their telepresence experience, and generally made for more natural interactions. Users were also asked if the ability to move content between the displays, from front to back and back to front, allowed for greater facility with presenting content on the screen.

The second scenario explored the experience of the 3D effects when animated between the displays. Using the *videoleaper* asset, participants were shown a series of animated videos where objects came towards the viewer and moved away from the user, shifting in size and between the rear display and the front display. For example, we showed users a ball animation that moved in a circular motion horizontally to the user, moving from the rear screen to the front screen. We also demonstrated an animation of a plane that flies toward the subject, moving from the rear screen to the front screen.

### **Videoconferencing Scenario**

Four participants (40%) reported being highly engaged (4 out of 5 on the scale) with this scenario, while two participants (20%) reported being even more highly engaged (5 out of 5). Additionally, all ten participants (100%) agreed that it was easier to empathize with the speaker on the DS-LCD teleconferences compared to traditional teleconferences. Their ratings varied, with 60% giving it a score of 3 out of 5, 30% rating it 4 out of 5, and 10% giving it the highest rating of 5 out of 5.

Furthermore, all ten participants (100%) expressed a preference for meeting via DS-LCD videoconferencing than with traditional approaches. Their ratings varied, with 50% giving it a score of 3 out of 5, 20% rating it 4 out of 5, and 30% giving it the highest rating of 5 out of 5.

Overall, when comparing the videoconferencing experience with DS-LCDs against traditional teleconferencing, participants indicated that they were more engaged, and enjoyed the meetings more than traditional modalities.

### **Animation Scenario**

Eight participants (80%) reported feeling engaged while watching visuals with the DS-LCD, rating their engagement as 4 out of 5 on the scale, and one participant (10%) rated it as 5 out of 5. This finding is particularly significant considering that only 30% of participants reported feeling engaged with traditional modalities (the sum of ratings 4 and 5 out of 5), whereas 90% of participants reported feeling engaged with the DS-LCD experience (the sum

of ratings 4 and 5 out of 5). This indicates a notable preference for a higher level of engagement with the DS-LCD visuals compared to the control.

## CONCLUSION

These initial prototypes validated the idea that layered transparent LCD displays can produce greater depth by combining both stereoscopic and volumetric approaches. The effect is something akin to looking into clear water from the shore. On the surface plane of the x and y axes, unadulterated images or stereoscopic images can be used to enhance binocular effects, while the shallow space in the z axis - in the physical gap between displays - can produce a very convincing volumetric effect. This iteration of the DS-LCD prototype challenges us to consider what other ways we might further integrate these two approaches in order to enhance the spatial effect?

In future iterations, one can imagine introducing more transparent LCD layers in the gap between the rear and front displays as a way to increase the resolution of the volumetric output. By further altering the images, including stereoscopic manipulation, one can imagine further enhancing the effect.

Mixed reality is also an area of interest for future interactions. Beyond the videoconferencing scenarios already explored, one can imagine a role for DS-LCDs in telepresence applications where the boundary between remote and local space is less distinct (Gonsher, 2024). At different scales and within different contexts, one can imagine a role for DS-LCDs that further blurs the boundaries between remote and local spaces, and is fully integrated into the built environment.

## ACKNOWLEDGMENT

The authors would like to acknowledge the Brown University Computer Science Department for supporting this project.

## REFERENCES

- England, J., 2018. Science of Pepper's Ghost Illusion. *Cosmos Magazine*. Available at: <https://cosmosmagazine.com/science/physics/the-science-behind-the-peppers-ghost-illusion/>.
- Gonsher, I., 2024. The telepresence of furniture in extended reality. *Interactions*, 31(2), pp. 44–48.
- Gonsher, I., Feng, P., McArn, T. and Christenson, A., 2023. Designing for wide adoption: An inexpensive and accessible extended reality (XR) device. In: *International Conference on Human-Computer Interaction*. Cham: Springer Nature Switzerland, pp. 185–191.
- Kooi, F. and Toet, A., 2004. Visual comfort of binocular and 3D displays. *Displays*, 25(2–3), pp. 99–108. Available at: <https://doi.org/10.1016/j.displa.2004.07.004>.
- Mphopo, W., 2020. Stereoscopia and autostereoscopia. In: *Mixed Reality and Three-Dimensional Computer Graphics*. Chapter 9. Available at: <https://doi.org/10.5772/intechopen.92633>.
- YouGov, 2011. YouGov survey results. Available at: [https://cdn.yougov.com/today\\_uk\\_import/yougov-archives-yougov-3dfilms-230911.pdf](https://cdn.yougov.com/today_uk_import/yougov-archives-yougov-3dfilms-230911.pdf).

# Digital Transformation for Sustainability: Industry 5.0 in UK SMEs

**Mohammad Rashed Khan**

International Business and Marketing, Salford Business School, University of Salford,  
Manchester, UK.

## ABSTRACT

In the era of Industry 5.0, digital transformation presents a critical opportunity for UK small and medium-sized enterprises (SMEs) to enhance sustainability while driving business growth. This paper will explore how SMEs can leverage advanced technologies, such as artificial intelligence (AI), the Internet of Things (IoT), and collaborative robots (cobots), to achieve sustainable operational practices and meet Environmental, Social, and Governance (ESG) goals. By examining case studies of successful digital transformations within UK SMEs, the research will highlight strategies for integrating green manufacturing techniques, optimising resource use, and reducing carbon footprints. The study will also address the challenges SMEs face in adopting these technologies, including financial constraints, skill gaps, and regulatory compliance. Furthermore, it will discuss the role of government initiatives and support programs in facilitating this transition. The expected outcomes include a strategic framework for SMEs, policy recommendations, and insights into the positive impact of Industry 5.0 on workforce inclusivity and ESG performance. This research aims to provide a practical roadmap for UK SMEs seeking to navigate the complexities of digital transformation and underscores the importance of adopting innovative, sustainable practices to remain competitive in a rapidly evolving market landscape.

**Keywords:** Artificial intelligence (AI), Talent management (TM), Human resources (HR), Human resource management (HRM), Future human resource management

## INTRODUCTION

In today's rapidly evolving market landscape, digital transformation has become a crucial driver for sustainability and business growth, particularly for small and medium-sized enterprises (SMEs) (Melo et al., 2023). The advent of Industry 5.0, which emphasises human-centric innovation and collaboration between humans and advanced technologies (Youssef and Mejri, 2023), presents a unique opportunity for UK SMEs to enhance their operational practices. By integrating technologies such as artificial intelligence (AI), the Internet of Things (IoT), and collaborative robots (cobots), SMEs can achieve significant improvements in efficiency, productivity, and sustainability.

The importance of sustainability in business operations has never been more pronounced. Companies are increasingly expected to meet Environmental, Social, and Governance (ESG) goals, not only to comply

with regulations but also to fulfil the growing demands of environmentally conscious consumers and stakeholders (Aldowaish et al., 2022). For UK SMEs, which represent a significant portion of the economy, embracing digital transformation is essential to maintaining competitiveness and ensuring long-term viability.

This paper explores how UK SMEs can leverage Industry 5.0 technologies to implement sustainable practices, optimise resource utilisation, and minimise environmental impact. Through detailed case studies of successful digital transformations, the research will uncover strategies for integrating green manufacturing techniques and reducing carbon footprints. It will also address the challenges SMEs face in adopting these technologies, including financial constraints, skill gaps, and regulatory compliance. Furthermore, the study will examine the role of government initiatives and support programs in facilitating this transition.

The expected outcomes of this research include the development of a strategic framework for SMEs, policy recommendations, and insights into the positive impact of Industry 5.0 on workforce inclusivity and ESG performance. By providing a practical roadmap for digital transformation, this study aims to empower UK SMEs to navigate the complexities of technological adoption and thrive in an increasingly sustainability-focused market.

## **METHODOLOGY**

This research adopts a mixed-methods approach to explore how UK SMEs can leverage Industry 5.0 technologies for sustainable transformation. The methodology comprises three key phases: literature review, case study analysis, and stakeholder interviews.

### **Phase 1: Literature Review**

A comprehensive literature review will be conducted to establish the theoretical framework for the study. This review will cover existing research on digital transformation, Industry 5.0 technologies (such as AI, IoT, and cobots), and sustainable practices within SMEs. It will also examine current trends, challenges, and best practices in the context of ESG goals. Sources will include academic journals, industry reports, government publications, and white papers.

### **Phase 2: Case Study Analysis**

The second phase involves an in-depth analysis of selected case studies of UK SMEs that have successfully implemented digital transformation initiatives. These case studies will be chosen based on criteria such as the extent of technology adoption, the impact on sustainability, and the achievement of ESG goals. Data will be collected through company reports, interviews with key personnel, and site visits where possible. The analysis will focus on identifying strategies, outcomes, and lessons learned from each case.

**Table 1.** Methodology flowchart.**Phase 1: Literature Review**

- Comprehensive literature review
- Focus on digital transformation, Industry 5.0, and sustainability
- Sources: Academic journals, industry reports, government publications

**Phase 2: Case Study Analysis**

- In-depth analysis of selected UK SMEs
- Criteria: Extent of technology adoption, sustainability impact, ESG goals
- Data collection: Company reports, interviews, site visits

**Phase 3: Stakeholder Interviews**

- Semi-structured interviews with key stakeholders
- Participants: SME owners, industry experts, technology providers, government representatives
- Analysis: Thematic analysis to identify common themes

**Phase 3: Stakeholder Interviews**

To complement the case study findings, semi-structured interviews will be conducted with key stakeholders, including SME owners, industry experts, technology providers, and representatives from government support programs. These interviews will provide qualitative insights into the challenges and opportunities faced by SMEs in adopting Industry 5.0 technologies. The interview data will be analysed using thematic analysis to identify common themes and patterns.

The combination of these methods will provide a robust and holistic understanding of how digital transformation can drive sustainability in UK SMEs, leading to practical recommendations for both practitioners and policymakers.

**DIGITAL TRANSFORMATION AND SUSTAINABILITY**

Digital transformation involves the adoption of digital technologies to fundamentally change how businesses operate and deliver value to customers (Gong and Ribiere, 2021). For SMEs, this transformation is crucial for maintaining competitiveness and enhancing operational efficiency. Studies by Costa et al. (2023) and Cichosz et al. (2020) highlight that SMEs often face unique challenges, such as limited financial resources, skill gaps, and a lack of strategic vision, which can hinder their digital transformation efforts. Despite these challenges, the adoption of digital technologies can lead to significant benefits, including improved productivity, cost savings, and enhanced customer experiences (Bharadwaj et al., 2013).

**INDUSTRY 5.0: A HUMAN-CENTRIC APPROACH**

Industry 5.0 builds on the foundation of Industry 4.0, which emphasised automation and data exchange in manufacturing technologies. However, Industry 5.0 places a greater emphasis on collaboration between humans and machines, aiming to create more personalised and sustainable production

processes. According to Fazal et al. (2022) and Nahavandi (2019), Industry 5.0 technologies such as artificial intelligence (AI), the Internet of Things (IoT), and collaborative robots (cobots) are pivotal in driving this transformation. These technologies enable real-time data collection and analysis, predictive maintenance, and efficient resource management, contributing to sustainable business practices.

## **SUSTAINABILITY AND ESG GOALS**

Sustainability has become a central focus for businesses worldwide, driven by increasing regulatory pressures and consumer demand for environmentally responsible practices. ESG goals provide a framework for companies to measure and report on their sustainability performance (Ozkan et al., 2023). Research by Eccles et al. (2014) and Friede et al. (2015) indicates that companies with strong ESG performance often experience better financial outcomes and greater resilience. For SMEs, integrating ESG principles into their operations can enhance their reputation, attract investment, and ensure long-term viability.

## **THE ROLE OF INDUSTRY 5.0 IN ACHIEVING SUSTAINABILITY**

Industry 5.0 technologies play a critical role in enabling SMEs to achieve sustainability and meet ESG goals. AI and IoT, for example, can optimise energy consumption and reduce waste through smart monitoring systems (Verdouw et al., 2016). Cobots can assist in implementing green manufacturing techniques by enhancing precision and efficiency, thus minimising material waste and energy use (Cherubini et al., 2016). Additionally, digital platforms can facilitate transparent reporting and tracking of ESG metrics, helping SMEs comply with regulatory requirements and meet stakeholder expectations.

## **CHALLENGES AND OPPORTUNITIES**

Despite the potential benefits, SMEs often face significant barriers to adopting Industry 5.0 technologies. Financial constraints, limited access to technical expertise, and regulatory compliance issues are among the most common challenges (Hansen et al., 2024; Muller et al., 2018). However, support from the government and other non-government organisations such as funding schemes and training programs, can play a crucial role in overcoming these barriers (Muller et al., 2024; Bascavusoglu-Moreau and Tether, 2011).

The integration of Industry 5.0 technologies offers a promising pathway for UK SMEs to enhance their sustainability and achieve ESG goals. While challenges remain, the strategic adoption of digital transformation practices can lead to substantial operational improvements and competitive advantages. This literature review underscores the importance of continued research and support for SMEs in navigating the complexities of digital transformation and sustainability.

## **EXPECTED FINDINGS AND ANALYSIS**

The research is anticipated to uncover several key findings regarding the adoption of Industry 5.0 technologies by UK SMEs for sustainable transformation. These findings will be analysed to provide a comprehensive understanding of the potential benefits, challenges, and strategic approaches necessary for effective digital transformation.

### **Enhanced Operational Efficiency and Sustainability**

It is expected that SMEs leveraging Industry 5.0 technologies, such as AI, IoT, and cobots, will demonstrate significant improvements in operational efficiency and sustainability. The integration of AI and IoT can lead to smarter energy management systems, optimising energy usage and reducing operational costs. Cobots, with their precision and efficiency, are likely to minimise material waste and improve production accuracy. This optimisation is anticipated to contribute to lower carbon footprints and a reduction in overall environmental impact.

### **Achievement of ESG Goals**

The study is expected to show that SMEs implementing these advanced technologies will be better positioned to achieve their Environmental, Social, and Governance (ESG) goals. Enhanced data collection and analysis capabilities provided by IoT and AI will facilitate more accurate tracking and reporting of ESG metrics. This transparency will not only help in regulatory compliance but also improve stakeholder trust and attract environmentally conscious customers and investors.

### **Case Studies of Successful Digital Transformations**

Through the analysis of case studies, the research will likely highlight successful instances of digital transformation within UK SMEs. These case studies are expected to reveal best practices and strategies that have been effective in overcoming common challenges such as financial constraints and skill gaps. For example, partnerships with technology providers and government support programs might emerge as critical enablers of successful digital transformations. The case studies will provide practical examples of how SMEs have integrated green manufacturing techniques and optimised resource use to achieve sustainability.

### **Challenges in Technology Adoption**

The research is expected to identify several challenges that UK SMEs face in adopting Industry 5.0 technologies. Financial constraints are likely to be a significant barrier, as advanced technologies often require substantial upfront investment. Additionally, a lack of technical expertise and skills among the workforce can impede the adoption process. Regulatory compliance might also present challenges, particularly for smaller enterprises with limited resources to navigate complex regulatory environments.



## **Role of Government Initiatives and Support Programmes**

The study is anticipated to underscore the importance of government initiatives and support programmes in facilitating the digital transformation of SMEs. Funding schemes, tax incentives, and training programs provided by the government can play a crucial role in mitigating financial and skill-related barriers. The research is expected to recommend the continuation and expansion of such programmes to support the widespread adoption of Industry 5.0 technologies among SMEs.

## **Positive Impact on Workforce Inclusivity**

An important finding expected from the study is the positive impact of Industry 5.0 on workforce inclusivity. By automating routine tasks and enabling more flexible and collaborative working environments, these technologies can create new job opportunities and enhance job satisfaction. The adoption of cobots, in particular, is anticipated to complement human workers rather than replace them, fostering a more inclusive and supportive workplace culture.

## **ANALYSIS**

The analysis will involve a detailed examination of the collected data to validate these expected findings. The operational improvements and sustainability gains will be quantified through metrics such as energy consumption, production efficiency, and carbon footprint reduction. The achievement of ESG goals will be evaluated through compliance records, stakeholder reports, and market performance indicators.

Case studies will be analysed qualitatively to extract actionable insights and strategies. Common themes and patterns across different case studies will be identified to develop a strategic framework for SMEs. Challenges and barriers will be analysed to understand their root causes and potential solutions, with a particular focus on the role of government support.

The analysis will also explore the broader implications of digital transformation for workforce inclusivity and business competitiveness. By synthesising these findings, the research will provide a holistic view of how Industry 5.0 technologies can drive sustainable growth in UK SMEs, offering practical recommendations for both practitioners and policymakers.

In summary, the expected findings and analysis aim to provide a robust understanding of the transformative potential of Industry 5.0 technologies for UK SMEs. By addressing the challenges and highlighting successful strategies, the research will contribute valuable insights into the sustainable digital transformation of SMEs.

## **CONCLUSION**

This research explores the transformative potential of Industry 5.0 technologies for UK SMEs, highlighting the critical role of digital transformation in achieving sustainability and enhancing operational efficiency. By leveraging advanced technologies such as AI, IoT, and cobots,

SMEs can optimise resource use, reduce environmental impact, and meet stringent ESG goals. The expected findings underscore the significant benefits of these technologies, including improved energy management, reduced waste, and enhanced ESG performance.

The research also identifies key challenges faced by SMEs in adopting Industry 5.0 technologies, such as financial constraints, skill gaps, and regulatory compliance issues. However, it emphasises the pivotal role of government initiatives and support programmes in mitigating these barriers, providing essential funding, training, and resources.

Through detailed case studies and stakeholder interviews, the study aims to uncover successful strategies and best practices, offering a strategic framework and policy recommendations for SMEs. Additionally, the research highlights the positive impact of Industry 5.0 on workforce inclusivity, fostering a collaborative and supportive work environment.

Overall, this research provides a practical roadmap for UK SMEs to navigate the complexities of digital transformation, emphasizing the importance of adopting innovative, sustainable practices to remain competitive in a rapidly evolving market landscape.

## REFERENCES

- Aldowaish, A., Kokuryo, J., Almazyad, O., & Goi, H. C. (2022). Environmental, social, and governance integration into the business model: Literature review and research agenda. *Sustainability*, *14*(5), 2959.
- Bascavusoglu-Moreau, E., & Tether, B. S. (2011). Resilient to what? The resilience of small firms to major external shocks. *Industry and Innovation*, *18*(5), 437–458.
- Ben Youssef, A., & Mejri, I. (2023). Linking digital technologies to sustainability through industry 5.0: A bibliometric analysis. *Sustainability*, *15*(9), 7465.
- Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. *MIS Quarterly*, *37*(2), 471–482.
- Cherubini, A., Passama, R., Crosnier, A., Lasnier, A., & Fraise, P. (2016). Collaborative manufacturing with physical human-robot interaction. *Robotics and Computer-Integrated Manufacturing*, *40*, 1–13.
- Cichosz, M., Wallenburg, C. M., & Knemeyer, A. M. (2020). Digital transformation at logistics service providers: Barriers and motivators. *International Journal of Physical Distribution & Logistics Management*, *50*(5), 472–494.
- Costa, A., Crupi, A., De Marco, C. E., & Di Minin, A. (2023). SMEs and open innovation: Challenges and costs of engagement. *Technological Forecasting and Social Change*, *194*, 122731.
- Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857.
- Fazal, N., Haleem, A., Bahl, S., Javaid, M., & Nandan, D. (2022). Digital management systems in manufacturing using industry 5.0 technologies. In *Advancement in Materials, Manufacturing and Energy Engineering, Vol. II: Select Proceedings of ICAMME 2021* (pp. 221–234). Singapore: Springer Nature Singapore.

- Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.
- Gong, C., & Ribiere, V. (2021). Developing a unified definition of digital transformation. *Technovation*, 102, 102217.
- Hansen, A. K., Christiansen, L., & Lassen, A. H. (2024). Technology isn't enough for Industry 4.0: On SMEs and hindrances to digital transformation. *International Journal of Production Research*, 1–21.
- Melo, I. C., Queiroz, G. A., Junior, P. N. A., de Sousa, T. B., Yushimito, W. F., & Pereira, J. (2023). Sustainable digital transformation in small and medium enterprises (SMEs): A review on performance. *Heliyon*, 9(3).
- Muller, J. M., Buliga, O., & Voigt, K. I. (2018). Fortune favors the prepared: How SMEs approach business model innovations in Industry 4.0. *Technological Forecasting and Social Change*, 132, 2–17.
- Müller, J. M., Islam, N., Kazantsev, N., Romanello, R., Olivera, G., Das, D., & Hamzeh, R. (2024). Barriers and Enablers for Industry 4.0 in SMEs: A combined integration framework. *IEEE Transactions on Engineering Management*.
- Nahavandi, S. (2019). Industry 5.0—A human-centric solution. *Sustainability*, 11(16), 4371.
- Ozkan, S., Romagnoli, S., & Rossi, P. (2023). A novel approach to rating SMEs' environmental performance: Bridging the ESG gap. *Ecological Indicators*, 157, 111151.
- Verdouw, C. N., Wolfert, J., Beulens, A. J. M., & Rialland, A. (2016). Virtualization of food supply chains with the internet of things. *Journal of Food Engineering*, 176, 128–136.

# Stress, Anxiety, and Depression in Young Adults: Findings From a User Diversity-Based Analysis

Victoria Knieling<sup>1</sup> and Nirmalya Thakur<sup>2</sup>

<sup>1</sup>Program in Linguistics, Emory University, Atlanta, GA 30322, USA

<sup>2</sup>Department of Computer Science, Emory University, Atlanta, GA 30322, USA

## ABSTRACT

The work of this paper presents multiple novel findings from a comprehensive analysis of a dataset that includes the stress, anxiety, and depression levels experienced by 95 young adults, computed using the Depression Anxiety Stress Scale (DASS). First, for age groups, 18-20, 21-25, and 26-30, average stress and anxiety levels were higher in females as compared to males. Second, for all these age groups, the percentages of females who experienced a higher level of depression as compared to anxiety or stress were 15%, 16%, and 33.33%, respectively - indicating an increasing trend. However, such an increasing trend was not observed for males across different age groups. Third, for all these age groups, the percentages of females who experienced a higher level of stress as compared to anxiety or depression were 80%, 64%, and 66.67%, respectively. The pattern was observed to be different for males as for all these age groups, the percentages of males who experienced a higher level of stress as compared to anxiety or depression were 41.66%, 59.09%, and 28.57%, respectively. Finally, Pearson's correlation was used to analyze the nature of correlations between stress, anxiety, and depression for each of these diversity groups of young adults which revealed multiple novel insights. For example, for the age group of 18-20, the correlation between the DASS Stress Score and the DASS Depression Score was observed to be statistically significant for males but not for females. For the age group of 26-30, the correlation between the DASS Anxiety Score and the DASS Depression Score was observed to be statistically significant for females but not for males. In addition to this, for this age group, the correlation between the DASS Stress Score and the DASS Depression Score was also observed to be statistically significant for males but not for females.

**Keywords:** Data analysis, User diversity, Stress, Depression, Anxiety, Correlation analysis

## INTRODUCTION

Young adulthood is a stage of development marked by several changes in social and interpersonal relationships (Konstam et al., 2015). It is also marked by relatively frequent occurrences of stress, anxiety, and depression (Qualter et al., 2015). Furthermore, during the phase of young adulthood, there is a heightened vulnerability to the development of several mental health-related problems (Kessler et al., 2005). Early-onset depression is considered a more severe manifestation of the disorder as late-onset

depression is linked to fewer psychosocial scars and a lower prevalence of simultaneous mental illnesses (Rohde et al., 1994). Furthermore, the presence of significant depression or anxiety throughout early adulthood may increase the likelihood of developing drug misuse or dependency later in life (Chilcoat et al., 1998). Therefore, analysis of stress, anxiety, and depression in young adults is crucial to investigate.

Anxiety is an incessant and persistent worry that affects the daily lives of people who suffer from it. It can manifest in different ways and have different causes (Dymond et al., 2009). People with anxiety are likely to display disorder-influenced behaviors, like avoidance, of topics that people without anxiety would most likely not dwell on (Dymond et al., 2009 and Lehto et al., 2009). Unsurprisingly, a common trigger topic for anxiety is death (Lehto et al., 2009). Another trigger topic for people with anxiety is the climate crisis, also defined as “eco-anxiety” (Pihkala et al., 2020).

Stress is a human state wherein a person displays physical or emotional reactions to situations deemed high-pressure or beyond control. The stress an individual experiences can also be exacerbated by different events (Sangalang et al., 2019). Stress is often found in one’s work life, especially those who work in healthcare (Godin et al., 2005). In an analysis of job stress, researchers found stress associated with poor mental health in both men and women (Tajvar et al., 2015 and Galvin et al., 2015). Depression is a common mental health disorder characterized by persistent negative moods and feelings. The severity of depression varies from person to person and severe cases of depression are also associated with a lack of self-care and thoughts of suicide. An individual who suffers from depression may also suffer from a decrease in cognitive functions such as memory and processing speed (McDermott et al., 2009).

In view of the above, this area of research has attracted the attention of researchers from several disciplines (as discussed in Section 2). However, none of the prior works in this area have focused on the analysis of stress, anxiety, and depression in young adults by considering user diversity with a specific focus on age group and gender. The work presented in this paper aims to address this research gap by presenting the findings of a comprehensive analysis of a dataset (Senaratne et al., 2021) that includes the stress, anxiety, and depression levels experienced by 95 young adults, computed using the Depression Anxiety Stress Scale (DASS). The rest of this paper is organized as follows. Section 2 presents a review of recent works in this field. Section 3 discusses the methodology and presents the results of this research work. It is followed by Section 4 where a summary of the scientific contributions of this paper is discussed and the scope for future work in this area of research is outlined.

## LITERATURE REVIEW

This section presents a review of the findings of recent works in this field. Anxiety in combination with depression can increase the risk of cardiovascular disease even further (Chalmers et al., 2014). Identifying and addressing anxiety early on may be a preventative measure against

cardiovascular events such as stroke. Furthermore, individuals with cardiovascular disease may be more likely to experience anxiety (Emdin et al., 2016). Studies have shown that among people with autism spectrum disorders, 39.6% also had an anxiety disorder, and among medical students, the prevalence rate of anxiety was 33.8% (van Steensel et al., 2011 and Quek et al., 2019). Specific jobs may be associated with varying levels of stress or anxiety. For example, nurses usually display a high level of stress due to the high-demand nature of their job and among ICU nurses, approximately 83.9% experience high levels of stress (Tajvar et al., 2015).

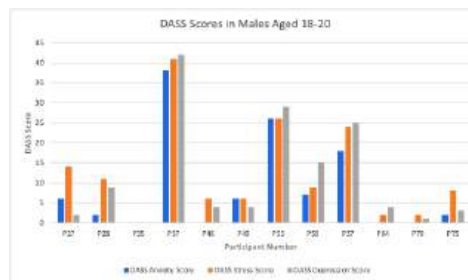
One or more factors in an individual's personal life may be associated with depression. For instance, in a study of adults in South Korea, it was found that unmarried men experienced higher rates of depression as compared to married men (Jang et al., 2009). Depression has also been recorded in people with diabetes (Gonzalez et al., 2008). Studies have shown that physical exercise is an effective way to offset the symptoms of depression, and regular exercise alongside medication has also proven to be effective against depression (Gerber et al., 2014 and Kvam et al., 2016). With the advancement of neural networks, Deep Convolutional Neural Networks (DCNN) can now help clinicians diagnose depression by analyzing a person's speech (He et al., 2018). Most depressive disorders can predict anxiety disorders and vice versa, and depressive disorders can also predict social or specific phobias (Jacobson et al., 2017 and Kotov et al., 2010). Depression and anxiety are often comorbid, and they usually have co-occurring symptoms. Generally, common symptoms include a sad or negative mood and worry, and these symptoms can influence and exacerbate each other. Comorbid disorders create a "network" of co-existing symptoms, making it difficult for a person to overcome the struggles of anxiety and depression. Anxiety and depression together can also preclude substance abuse (Beard et al., 2016). To understand the full extent of depression and anxiety and how the same impacts a person, it is important to understand the severity of the same and the underlying causes (Knowles et al., 2020). As depression is a strong predictor of anxiety, analysis of the severity of depression is critical to assess a person's likelihood of experiencing anxiety (Knowles et al., 2020).

In summary, there has been a considerable amount of work performed in this area of research. However, none of the prior works in this area have focused on the analysis of stress, anxiety, and depression in young adults by considering user diversity with a specific focus on age group and gender. Addressing this research gap serves as the main motivation for performing this research work.

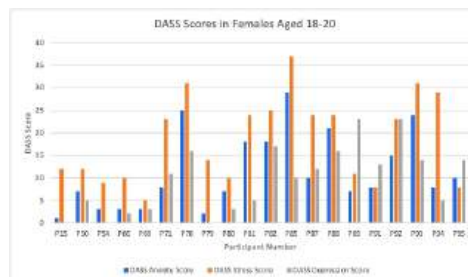
## **METHODOLOGY AND RESULTS**

The dataset (Senaratne et al., 2021) used for this work presents the stress, anxiety, and depression levels experienced by 95 young adults, computed using the Depression Anxiety Stress Scale (DASS). To develop this dataset, the researchers used social media groups, such as those for mental health support, psychological clinics, and university advertising to recruit 95 individuals between the ages of 18 and 35. The racial backgrounds of

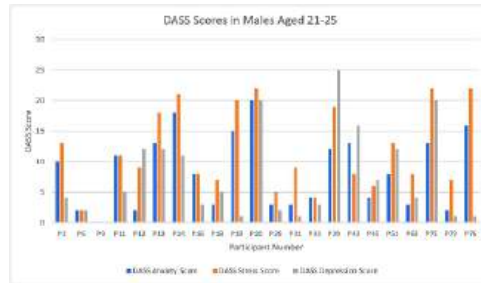
the participants were Caucasian (23.16%), Black or African American (2.1%), East Asian (27.37%), and South Asian (47.37%). For anxiety-related problems, 41 participants had recently sought out professional or specialized help. Others did not suffer from anxiety disorders and had not previously taken anxiety medication or received outside assistance. In order to minimize the impact of unknown factors on physiological and behavioral measurements, the researchers excluded individuals according to several factors which included the use of beta-blocker drugs, drug abuse, impairments in mobility, neurological and cardiovascular disorders, lack of ability to speak English, uncorrected vision or hearing impairments, and intellectual disabilities. As an additional precaution against potential risks, the researchers also disqualified individuals having a diagnosis of panic disorder. For performing the research work presented in this paper, the file with the filename – “participants\_details.csv” from Version 2 of this dataset (the most recent version at the time of conducting this work) was used. First, the DASS scores per gender per age group were analyzed. The results of this analysis are presented in Figures 1-7. It is worth mentioning that in the dataset file, there were no individuals in the age group of 31–35 who identified as female, so a visualization for the same is not presented. In all these figures, the X-axis represents the participant number (as presented in this dataset file) and the Y-axis represents the DASS score.



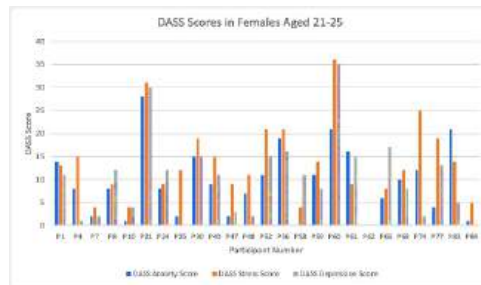
**Figure 1:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in males (age group 18 to 20).



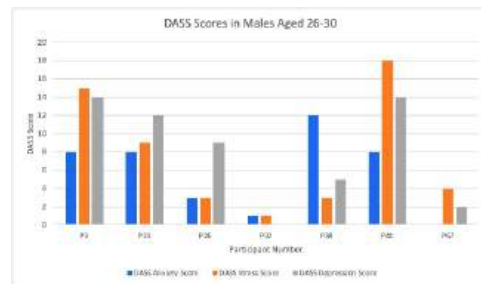
**Figure 2:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in females (age group 18 to 20).



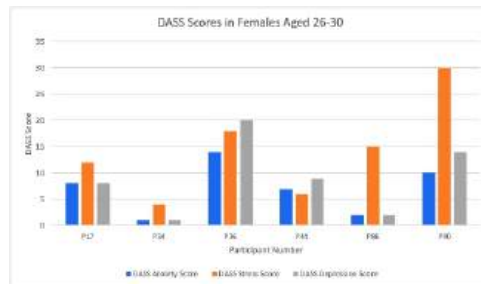
**Figure 3:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in males (age group 21 to 25).



**Figure 4:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in females (age group 21 to 25).

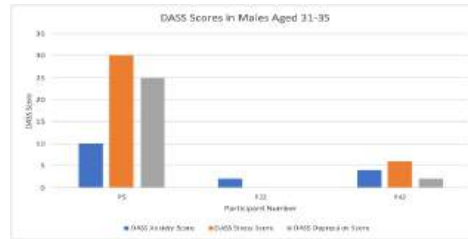


**Figure 5:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in males (age group 26 to 30).



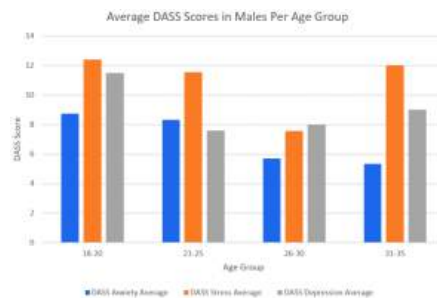
**Figure 6:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in females (age group 26 to 30).



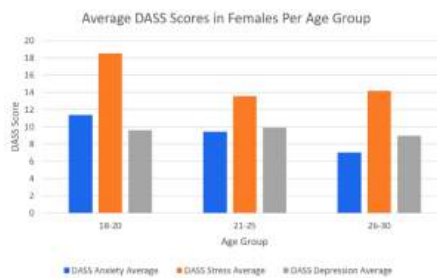


**Figure 7:** Analysis of DASS anxiety score, DASS stress score, and DASS depression score in males (age group 31 to 35).

Thereafter, the average DASS score in males and females was computed. The results of the same are shown in Figures 8 and 9. The analysis shown in Figures 1 to 9 revealed multiple novel insights. First, for age groups, 18–20, 21–25, and 26–30, average stress and anxiety levels were higher in females as compared to males. Second, for all these age groups, the percentages of females who experienced a higher level of depression as compared to anxiety or stress were 15%, 16%, and 33.33%, respectively - indicating an increasing trend. However, such an increasing trend was not observed for males across different age groups. Third, for all these age groups, the percentages of females who experienced a higher level of stress as compared to anxiety or depression were 80%, 64%, and 66.67%, respectively. The pattern was observed to be different for males as for all these age groups, the percentages of males who experienced a higher level of stress as compared to anxiety or depression were 41.66%, 59.09%, and 28.57%, respectively.



**Figure 8:** Analysis of the average DASS anxiety score, DASS stress score, and DASS depression score in males across different age groups.



**Figure 9:** Analysis of the average DASS anxiety score, DASS stress score, and DASS depression score in females across different age groups.

---

**Algorithm 1:** Correlation Analysis of DASS Scores per age group per gender
 

---

**Input:** specific demographic data

**Output:** heatmap, correlation matrix, nature of correlation

Import numpy, pandas, matplotlib.pyplot, csv, scipy.stats, seaborn

df:= read data

**def** check\_correlation(col1, col2):

    stat := Pearson's correlation between col\_1 and col\_2

    p\_value:= p-value from the stat variable

**If** (p\_value < 0.05) **then:**

        correlation between <col\_1> and <col\_2> is statistically significant

**end of function**

list:= list for demographic information

**for** each element in the list:

    check\_correlation ('DASS\_Anxiety\_Score', 'DASS\_Stress\_Scores')

    check\_correlation ('DASS\_Anxiety\_Score', 'DASS\_Depression\_Scores')

    check\_correlation ('DASS\_Stress\_Scores', 'DASS\_Depression\_Scores')

    corr:= Pearson's correlation matrix of the data

    plot:= heatmap with correlation matrix

    display plot and output the nature of correlations

**end of for loop**

---

Thereafter, correlation analysis was utilized to examine the characteristics of the DASS scores within each age group and gender. A correlation coefficient is a numerical metric that quantifies the extent of correlation, which represents a statistical relationship between two variables. The variables can refer to either two features of a collection of observations, commonly known as a sample, or two parts of a multivariate random variable with a known distribution. Based on the methodology of similar works in this field (Thakur et al., 2024), the specific approach that was followed involved calculating the Pearson correlation coefficient (PCC) to determine the correlation among the specific DASS scores. The Pearson correlation coefficient (PCC) is a statistical measure that precisely assesses the magnitude and nature of the linear relationship between two sets of data. The Pearson correlation coefficient (PCC) is computed by dividing the covariance of two variables by the product of their standard deviations. The normalized value it offers ranges from  $-1$  to  $1$ , showing both the magnitude and nature of the connection (Sedgwick et al., 2012). The pseudocode of the program that was written in Python 3.10 to perform this correlation analysis is shown in Algorithm 1. This program computed the results of correlation analysis per age group per gender and presented the results in the form of a correlation matrix. The correlation matrix generated by this program for all males in the age group of 18–20 is shown in Figure 10. For paucity of space, the correlation matrices of all the other diversity groups are not presented as separate images, and the results of the same are summarized in Table 1.



**Figure 10:** Correlation matrix for all males in the age group of 18 to 20.

**Table 1.** Summary of the results of correlation analysis from Algorithm 1.

Attributes for Correlation Analysis	Nature of Correlation
Males in the age group of 18 to 20	
DASS_Anxiety_Score and DASS_Stress_Scores	statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	statistically significant
Females in the age group of 18 to 20	
DASS_Anxiety_Score and DASS_Stress_Scores	statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	not statistically significant
Males in the age group of 21 to 25	
DASS_Anxiety_Score and DASS_Stress_Scores	statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	statistically significant
Females in the age group of 21 to 25	
DASS_Anxiety_Score and DASS_Stress_Scores	statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	statistically significant
Males in the age group of 26 to 30	
DASS_Anxiety_Score and DASS_Stress_Scores	not statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	not statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	statistically significant
Females in the age group of 26 to 30	
DASS_Anxiety_Score and DASS_Stress_Scores	not statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	not statistically significant
Males in the age group of 31 to 35	
DASS_Anxiety_Score and DASS_Stress_Scores	statistically significant
DASS_Anxiety_Score and DASS_Depression_Scores	not statistically significant
DASS_Stress_Scores and DASS_Depression_Scores	not statistically significant

The results of the correlation analysis revealed multiple novel insights. For instance, for the age group of 18–20, the correlation between the DASS Stress Score and the DASS Depression Score was observed to be statistically

significant for males but not for females. For the age group of 26–30, the correlation between the DASS Anxiety Score and the DASS Depression Score was observed to be statistically significant for females but not for males. In addition to this, for this age group, the correlation between the DASS Stress Score and the DASS Depression Score was also observed to be statistically significant for males but not for females. The work presented in this research paper has a limitation. The results presented in this paper are based on the data available in the dataset (Senaratne et al., 2021) which was used for this research work. There are multiple factors that affect anxiety, stress, and depression, which may vary from individual to individual. For instance, during the COVID-19 outbreak, a rapid transition to online learning was shown to be associated with a degree of stress in educators as well as in students (Hall et al., 2022 and Thakur et al., 2023). So, it is possible that if this study is conducted again by recruiting participants during an ongoing virus outbreak or any other event of global concern, the results may vary as a virus outbreak or an event of global concern may cause anxiety, stress, and depression in individuals in multiple ways.

## **CONCLUSION**

Many changes in social and interpersonal connections characterize the developmental stage of young adulthood. Furthermore, young adulthood is characterized by increased susceptibility to mental health issues and instances of stress, anxiety, and depression. Research works in this field have shown that depression or anxiety throughout the initial stages of adulthood may raise the probability of developing drug misuse or addiction in the future. Therefore, the analysis of stress, anxiety, and depression in young adults has garnered the interest of researchers from several disciplines. However, none of the prior works in this field have examined the analysis of stress, anxiety, and depression in young adults while taking into account user diversity, particularly in terms of age group and gender. The work presented in this research paper aims to address this research gap by presenting the findings of a comprehensive analysis of a dataset that contains the DASS Anxiety Score, DASS Stress Score, and DASS Depression Score of 95 young adults in the age range of 18 to 35. The gender identity of these young adults is also available in the dataset. The methodology that was followed in this research work involved multiple forms of data analysis and correlation analysis which revealed several novel insights in terms of stress, anxiety, and depression in an age-group-specific and gender-specific manner. Future work in this area would involve performing a similar data collection of middle-aged and elderly individuals and repeating this study to analyze the variation of stress, anxiety, and depression in an age-group-specific and gender-specific manner in middle-aged and elderly individuals.

## **DATA AVAILABILITY STATEMENT**

A publicly available dataset was analyzed in this study. The data can be found at [https://bridges.monash.edu/articles/dataset/Anxiety\\_Phases\\_Dataset/15176082](https://bridges.monash.edu/articles/dataset/Anxiety_Phases_Dataset/15176082), accessed on 7 December 2023.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## REFERENCES

- Beard, C. et al. (2016) "Network analysis of depression and anxiety symptom relationships in a psychiatric sample," *Psychological medicine*, 46(16), pp. 3359–3369. doi: 10.1017/s0033291716002300.
- Chalmers, J. A. et al. (2014) "Anxiety disorders are associated with reduced heart rate variability: A meta-analysis," *Frontiers in psychiatry*, 5. doi: 10.3389/fpsyt.2014.00080.
- Chilcoat, H. D. and Breslau, N. (1998) "Posttraumatic stress disorder and drug disorders: Testing causal pathways," *Archives of general psychiatry*, 55(10), p. 913. doi: 10.1001/archpsyc.55.10.913.
- Dymond, S. and Roche, B. (2009) "A contemporary behavior analysis of anxiety and avoidance," *The Behavior analyst*, 32(1), pp. 7–27. doi: 10.1007/bf03392173.
- Emdin, C. A. et al. (2016) "Meta-analysis of anxiety as a risk factor for cardiovascular disease," *The American journal of cardiology*, 118(4), pp. 511–519. doi: 10.1016/j.amjcard.2016.05.041.
- Galvin, J. et al. (2015) "Mental health nursing students' experiences of stress during training: a thematic analysis of qualitative interviews," *Journal of psychiatric and mental health nursing*, 22(10), pp. 773–783. doi: 10.1111/jpm.12273.
- Gerber, M. et al. (2014) "Physical activity in employees with differing occupational stress and mental health profiles: A latent profile analysis," *Psychology of sport and exercise*, 15(6), pp. 649–658. doi: 10.1016/j.psychsport.2014.07.012.
- Godin, I. et al. (2005) "A prospective study of cumulative job stress in relation to mental health," *BMC public health*, 5(1). doi: 10.1186/1471-2458-5-67.
- Gonzalez, J. S. et al. (2008) "Depression and diabetes treatment nonadherence: A meta-analysis," *Diabetes care*, 31(12), pp. 2398–2403. doi: 10.2337/dc08-1341.
- Hall, I., Thakur, N. and Han, C. Y. (2022) "Trends in Remote Learning-based Google Shopping in the United States due to COVID-19," arXiv [cs. CY]. Available at: <https://arxiv.org/abs/2204.12654> (Accessed: April 22, 2024).
- He, L. and Cao, C. (2018) "Automated depression analysis using convolutional neural networks from speech," *Journal of biomedical informatics*, 83, pp. 103–111. doi: 10.1016/j.jbi.2018.05.007.
- Jacobson, N. C. and Newman, M. G. (2017) "Anxiety and depression as bidirectional risk factors for one another: A meta-analysis of longitudinal studies," *Psychological bulletin*, 143(11), pp. 1155–1200. doi: 10.1037/bul0000111.
- Jang, S.-N. et al. (2009) "Marital status, gender, and depression: Analysis of the baseline survey of the Korean Longitudinal Study of Ageing (KLoSA)," *Social science & medicine* (1982), 69(11), pp. 1608–1615. doi: 10.1016/j.socscimed.2009.09.007.
- Kessler, R. C. et al. (2005) "Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication," *Archives of general psychiatry*, 62(6), p. 593. doi: 10.1001/archpsyc.62.6.593.
- Knowles, K. A. and Olatunji, B. O. (2020) "Specificity of trait anxiety in anxiety and depression: Meta-analysis of the State-Trait Anxiety Inventory," *Clinical psychology review*, 82(101928), p. 101928. doi: 10.1016/j.cpr.2020.101928.
- Konstam, V. (2015) *Emerging and young adulthood: Multiple perspectives, diverse narratives*. Cham: Springer International Publishing.

- Kotov, R. et al. (2010) "Linking 'big' personality traits to anxiety, depressive, and substance use disorders: A meta-analysis," *Psychological bulletin*, 136(5), pp. 768–821. doi: 10.1037/a0020327.
- Kvam, S. et al. (2016) "Exercise as a treatment for depression: A meta-analysis," *Journal of affective disorders*, 202, pp. 67–86. doi: 10.1016/j.jad.2016.03.063.
- Lehto, R. H. and Stein, K. F. (2009) "Death anxiety: An analysis of an evolving concept," *Research and theory for nursing practice*, 23(1), pp. 23–41. doi: 10.1891/1541-6577.23.1.23.
- McDermott, L. M. and Ebmeier, K. P. (2009) "A meta-analysis of depression severity and cognitive function," *Journal of affective disorders*, 119(1–3), pp. 1–8. doi: 10.1016/j.jad.2009.04.022.
- Pihkala, P. (2020) "Anxiety and the ecological crisis: An analysis of Eco-anxiety and climate anxiety," *Sustainability*, 12(19), p. 7836. doi: 10.3390/su12197836.
- Qualter, P. et al. (2015) "Loneliness across the life span," *Perspectives on psychological science: A journal of the Association for Psychological Science*, 10(2), pp. 250–264. doi: 10.1177/1745691615568999.
- Quek et al. (2019) "The global prevalence of anxiety among medical students: A meta-analysis," *International journal of environmental research and public health*, 16(15), p. 2735. doi: 10.3390/ijerph16152735.
- Rohde, P., Lewinsohn, P. M. and Seeley, J. R. (1994) "Are adolescents changed by an episode of major depression?," *Journal of the American Academy of Child and Adolescent Psychiatry*, 33(9), pp. 1289–1298. doi: 10.1097/00004583-199411000-00010.
- Sangalang, C. C. et al. (2019) "Trauma, post-migration stress, and mental health: A comparative analysis of refugees and immigrants in the United States," *Journal of immigrant and minority health*, 21(5), pp. 909–919. doi: 10.1007/s10903-018-0826-2.
- Sedgwick, P. (2012) "Pearson's correlation coefficient," *BMJ (Clinical research ed.)*, 345(jul04 1), pp. e4483–e4483. doi: 10.1136/bmj.e4483.
- Senaratne, H. et al. (2021) *Anxiety Phases Dataset*. Monash University. doi: 10.26180/15176082. V2.
- van Steensel, F. J. A., Bögels, S. M. and Perrin, S. (2011) "Anxiety disorders in children and adolescents with autistic spectrum disorders: A meta-analysis," *Clinical child and family psychology review*, 14(3), pp. 302–317. doi: 10.1007/s10567-011-0097-0.
- Tajvar, A. et al. (2015) "Occupational stress and mental health among nurses in a medical intensive care unit of a general hospital in Bandar Abbas in 2013," *Electronic Physician*, 7(3), p. 1108. doi: 10.14661/2015.1108–1113.
- Thakur, N. et al. (2023) "Investigation of the gender-specific discourse about online learning during COVID-19 on Twitter using sentiment analysis, subjectivity analysis, and toxicity analysis," *Computers*, 12(11), p. 221. doi: 10.3390/computers12110221.
- Thakur, N. et al. (2024) "Investigation of the misinformation about COVID-19 on YouTube using topic modeling, sentiment analysis, and language analysis," *Computation (Basel, Switzerland)*, 12(2), p. 28. doi: 10.3390/computation12020028.

# Establishing Sustainable Health Services for the Medication of Elderly Chronic Diseases: An Analysis Based on SAPAD

Yuxin Sheng, Tingmin Yan, Yanhao Geng, Zijing Dai, and Yinan Li

School of Art, Southeast University, Nanjing, Jiangsu 211189, China

## ABSTRACT

Over 85% of elderly individuals suffer from one or more chronic diseases, with the primary treatment approach involving long-term use of targeted medications. The medication process for afflicted elderly individuals comprises three essential steps: medical evaluation, obtaining prescribed medication, and adhering to the prescribed regimen. Due to various factors such as environment, systems, and cognitive aspects, the average medication adherence rate among the elderly is below 45%, leading to significant loss in terms of both health and life and wastage of medical resources each year. As the aging population continues to grow, the need for sustainable and effective medication services becomes increasingly urgent. This study employs SAPAD analysis to establish touch points for sustainable health medication services. Starting from the behaviors of five typical elderly individuals with chronic diseases and employing quantitative research methods, objective issues within the service system are identified. In response to the identified issues, a design thinking approach is employed to construct sustainable health services for the medication of chronic diseases in the elderly. The aim is to mitigate the medical burden brought about by chronic diseases and actively leverage the roles of communities, pharmacies, online medical platforms, and the Internet of Things to enhance the healthcare experience for elderly individuals with chronic diseases. Furthermore, as an outcome of this research, the optimized strategies are applied to the practical design of the healthcare service system for elderly medication, with the goal of promoting sustainable well-being for the elderly and societal development.

**Keywords:** Chronic disease medication, Health services, SAPAD model, User behavior, Design strategies, Sustainable development

## INTRODUCTION

With advancing age, elderly individuals often suffer from chronic underlying conditions, necessitating the intake of multiple types of medications on a regular basis. However, due to the decline in sensory perception and cognitive function associated with aging, as well as a reduction in memory capacity, the elderly frequently experience situations where they forget to take their medication or take it incorrectly. However, medication non-adherence remains an enduring and intricate issue (Liu and Varshney, 2020). Additionally, due to age-related physiological changes and susceptibility to various chronic underlying conditions among the elderly (Chiang-Hanisko

et al., 2014), they face a higher risk of medication non-adherence compared to younger counterparts. This can lead to exacerbated health conditions and a proportionate surge in healthcare expenses (Lam and Fresco, 2015). Medication adherence stands as a significant health concern, posing considerable challenges to the sustainable development of an aging society (Faisal et al., 2021). Consequently, enhancing medication adherence is pivotal for improving health outcomes and effectively managing chronic diseases among the elderly.

Research predicts that the worldwide population comprising both the elderly and the young will live in smart homes, smart communities, and smart cities in the coming years (Thakur and Han, 2022). The advent of intelligent healthcare is driving the evolution of medical products, wherein intelligent medication adherence solutions utilize various connectivity methods to capture and transmit real-time medication intake data. This, in turn, enables the provision of smart medication services for the elderly (Faisal et al., 2021). Regrettably, despite the numerous design objectives proposed to address objective medication adherence challenges within the medical field (Jimmy and Jose, 2011), certain existing products fail to effectively translate these design goals into practical product requirements, often lacking a consideration of user experience. Existing medical research suggests that the development of optimal solutions involves utilizing new products, services, and experiences. Abookire (2020) proposed the application of design thinking to lead medical transformation. When compared to conventional problem-solving methods in healthcare and other public health-related fields, design thinking has the potential to yield more successful and sustainable intervention measures (Annweiler et al., 2023). Consequently, this article views design thinking as an effective problem-solving approach, aiming to transform objective clinical challenges into service designs driven by the principles of design and sustainability.

Therefore, the design of a medication service system requires a multi-layered and multidimensional approach. In our research, we introduce the SAPAD method, which stands for Semiotics Approach of Product Architecture Design. SAPAD is a model framework proposed by Hu (2013, 2015) that centers around solving practical problems with a user-centric approach. This model, combining product semiotics theory, delves into the deep meanings, values, and potential contradictions in the complex interaction among various components of systems, users, and stakeholders. Through multidimensional analysis and mapping of “behavior-meaning-object,” SAPAD reconstructs the elements and relationships within the service system, including software, hardware, and environmental factors (Hu et al., 2020 and Zhang et al., 2017). The application of SAPAD involves three main stages: behavioral observation and analysis, meaning analysis and construction, and service/product development and design. It spans various domains such as tangible products, information interaction, and service and system design. Yang and He (2023) proposed an SAPAD-based interactive product design method and applied it to the context of online education for international students; Liu and wan (2024) apply the SAPAD-AHP method to propose a systematic digital design for triage services for elderly patients.



In this study, by introducing the SAPAD model framework, we can map the existing medication process to medication behavior. Additionally, the model consists of multiple layers including the semantic layer, the experiential layer, and the syntactic layer. The syntactic layer involves constructing various systems from an objective existence perspective, the experiential layer pertains to the user's software and hardware touchpoints in the process from a temporal dimension, and the semantic layer arises from the user's actual medication needs from their viewpoint. This integration across interdisciplinary perspectives forms the foundation to effectively reconstruct and innovate the complex elderly intelligent medication service system. This is crucial for achieving efficiency and sustainability within the service system.

In this study, through the application of the SAPAD model and under the design-thinking driven integration of new technologies, we have explored an approach to optimize intelligent medication services for the elderly. These contributions collectively represent our effort to enhance the quality and effectiveness of medication services for the elderly, incorporating innovative methodologies and technologies, while also promoting sustainable health practices.

## **PARTICIPANTS**

### **Selection of Typical Samples**

In the initial phase of optimizing and improving medication services for the elderly with chronic diseases, 20 elderly individuals within the community showed a positive attitude towards the research topic and were willing to provide details of their illnesses and medication journeys, thereby seeking welfare for their group.

We conducted a preliminary investigation of the medication-related contexts of chronic disease patients, summarizing and organizing the obtained information into primary data. Since relatives and attending physicians play an important role in the care process of patients, we conducted in-depth interviews with elderly chronic disease patients, their relatives, and attending physicians based on a preliminary understanding of the relevant care contexts.

To comprehensively understand the medication process of elderly chronic disease patients, we selected typical observational subjects based on the following criteria: (1) Suffering from common chronic diseases, such as diabetes, hypertension, stroke, etc., and requiring regular and timely medication; (2) The initial data can be corroborated by relatives and attending physicians; (3) The data and medication samples provided by the patients have a high degree of authenticity, and they are willing to discuss their true thoughts. Therefore, we selected five typical patients, A, B, C, D, and E: A, 67 years old, male, with hypertension, requiring daily timed antihypertensive medication; B, 78 years old, female, with diabetes, requiring daily medication and insulin injections; C, 75 years old, male, with stroke; D, female, with dual comorbidities, suffering from both hypertension and diabetes, requiring daily medication and insulin injections; E, with triple comorbidities, suffering from heart disease, kidney disease, and diabetes, requiring daily medication.

### **Confirmation of User Paths**

Approved by the Ethics Committee, the observational samples voluntarily joined the research topic, and with the consent of their family members, we conducted informal interviews with the five typical patients through field visits combined with non-participant observation and recording of medical behavior, pathways, and emotions. By tracking the medication process, we identified medication-related activities: seeking medical treatment, prescribing, obtaining medication/buying medication, payment, and taking medication.

### **Summary of Experimental Materials**

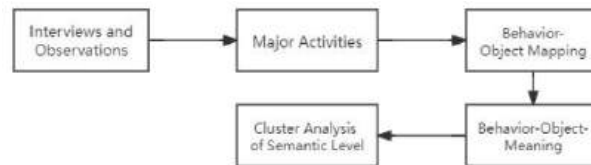
Based on the behavioral observation experiment, we discovered the following issues: during the medical check-up phase, elderly individuals faced challenges in physically queuing up at hospitals due to the inability to make online appointments. Complex department structures and building layouts further hindered efficient navigation. For patients undergoing check-ups, the registration, payment, and consultation processes were also cumbersome, with issues like repeated tests and difficulty in effective communication with doctors. Regarding medication acquisition, obtaining medication from hospitals was accompanied by lengthy commutes and intricate procedures. Community pharmacies, on the other hand, had problems related to prescription verification, limited medication options, and medication errors. For medication adherence, forgetfulness emerged as the primary reason for non-adherence, followed by high medication costs. Additionally, elderly individuals with conditions requiring short-acting medication expressed difficulty in obtaining their physiological data, such as hypertensive patients who couldn't time their medication accurately. Problems in medication counting and replenishment were also observed.

From an industry and technology perspective, microelectronics and information technology have matured. Applying these technologies through design thinking can enhance medical efficiency and service quality in the realm of sustainable healthcare. Therefore, for sustainable health services in chronic disease medication, we identified opportunity points based on the above problems. These included navigation systems, electronic health records, medication management systems, emergency assistance systems, welfare systems, caring systems, and medication adherence systems. These points were analyzed for backend and hardware-software composition, service touchpoints, and key design considerations across the entire service process.

### **METHODS**

Based on interviews and observations with participants, this study summarizes and generalizes the major activities of medication use among the elderly with chronic diseases. Subsequently, these major activities are further subdivided and mapped against the physical objects associated with the activities. On this basis, the study delves into the significance of user behaviors, mapping them at the levels of syntax, experience, and semantics.

Following this, a cluster analysis is conducted to identify the core clusters of meaning of Semantic Level.



**Figure 1:** Flowchart of the study method.

### Analysis of Behavior-Object Mapping

The analysis of behavior-object mapping primarily pertains to intervention activities under the community and well-being context. Firstly, there are medical activities encompassing various online and offline healthcare services. Secondly, medication supply activities cover medication services provided by community hospitals, pharmacies, online deliveries, and more. Thirdly, there are social care and welfare actions, including medical insurance, disease subsidies, and assistance for vulnerable groups. Lastly, product services consist of a range of elderly care and rehabilitation products aimed at promoting the well-being of the elderly.

**Table 1.** Mapping of behavior to objects.

Activity	Environment	Task	Subtask
Medical Activities	Registration Counter	Register	Book appointment, Queue for registration, Wait for consultation
	Outpatient Department	Consultation	Describe symptoms, Medical tests, Doctor's diagnosis, Prescription issuance
	Payment Counter	Payment	Registration payment, Consultation payment, Medication payment, Test payment
Medication Supply	Pharmacy/ Hospital Pharmacy	Collect Medication	Provide prescription, Pharmacist dispensing, Payment, Collect medication
	Online	Medication Delivery	Provide prescription, Online payment, Medication preparation, Home delivery
Social Care and Welfare	Community	Volunteering	Day care, Personal care, Housekeeping, Emotional support
	Government and Social Assistance Organizations	Financial Assistance	Medication subsidy, Disease subsidies, Medical insurance
Product Services	Home	Smart Services	Physiological monitoring, Medication reminders, Family supervision, Medication record- keeping, Medication management
	Outdoors	Medication Assurance	Pill counting, Pharmacy map, Emergency call, Portable carrying
	Hospital	Information Services	Electronic card simulation, Synchronized medical records, Automated payment, Data provision
	Server Room	Establish Medication Cloud	Big data monitoring, Artificial intelligence analysis, Medication logistics management

To deconstruct the common activities of medication among elderly individuals with chronic diseases, and to prepare for the mapping of meanings. As shown in Table I, the four major activities were further divided into 11 tasks and 40 subtasks.

### **Mapping of Behavior-Object-Meaning**

Conducting meaning mining and mapping analysis based on the deconstructed user behaviors is a pivotal step within the SAPAD theoretical framework. The meanings are divided into three levels: Syntax, Experience, and Semantics. Thus, we begin by mapping user behaviors to these three levels.

Syntax Level primarily deals with the interaction among various physical attributes and serves as a foundational component in constructing the service system. It encompasses seven meaning points: (1) Navigation System, (2) Electronic Health Records Cloud, (3) Medication Cloud, (4) Emergency Assistance System, (5) Welfare System, (6) Care System, and (7) Medication Adherence System.

Experience Level focuses on the interaction between users and products, as well as between users and stakeholders (on the production side) within the service system. The Experience Level comprises twenty-two meaning points: (1) Queue Algorithm, (2) Augmented Reality (AR) Navigation, (3) Path Planning, (4) Synchronized Medical Records, (5) Remote Analysis, (6) Electronic Prescription Issuance, (7) Logistic Deployment, (8) Medication Management, (9) Medication Records, (10) Prescription Verification, (11) Physiological Monitoring, (12) One-Touch Emergency Call, (13) Subsidy Disbursement, (14) Medical Insurance Coordination, (15) Fee Reduction, (16) Precise Assistance, (17) Family Supervision, (18) Disease Progress Tracking, (19) Caregiver Appointment, (20) Medication Reminder, (21) Adherence Assessment, and (22) Side Effect Management.

The Semantic level focuses on individual emotional factors and user experiences that span the entire elderly medication process. The Semantic level comprises 32 meaning points: (1) Queue Captain, (2) Long Waiting Time, (3) Heavy Physical Burden, (4) Poor Medical Experience, (5) Frequent Payment, (6) Cannot book Appointments, (7) Bad Hospital Signage, (8) Communication Barriers, (9) Instruction Confusion, (10) Missed Med-time, (11) Missed Med-type, (12) Excessive Carrying, (13) Forget Med-effects, (14) High Retesting Costs, (15) Inadequate Med Subsidies, (16) Inadequate Disease Subsidies, (17) High Med Costs, (18) Incomplete Healthcare, (19) Lack of Monitoring, (20) No Daytime Care, (21) Inadequate Reminders, (22) Lack of family Guardianship, (23) Lack of Med Records, (24) Insufficient Service info, (25) Low Emergency Support, (26) Unable to carry Med, (27) Improper Med storage, (28) Inadequate Med Inventory, (29) Limited Med-types, (30) Community Med Counting, (31) Uneven Med Distribution, (32) Assessing Med Source.



The result is a symmetric cluster matrix, leading to the identification of meaning clusters. In the cluster analysis of the Syntax level, five meaning clusters were obtained: (1) Medicine collection process, (2) Cognitive Load, (3) Economic Burden, (4) Drug Resource, and (5) Medication contact.

### Construction of the SAPAD Mapping Model for Elderly Medication Services System

By employing the SAPAD model, we accomplished the mapping of the multi-level, multi-dimensional elderly medication service system (see Figure 3). This mapping process enabled us to extract the core meanings of the medication system, software and hardware touch points. The central circle represents the spatial distribution of the medication system. Progressing outward, we obtained the outer circle, which pertains to software and hardware touch points distributed over time, and the inner circle, which represents user medication requirements. Additionally, we conducted clustering on the medication requirements within the inner circle, yielding five meaning clusters.

According to above behavior processes and five meaning clusters, operation of the service system can be expressed in the form of service blueprint, as shown in Figure 4.

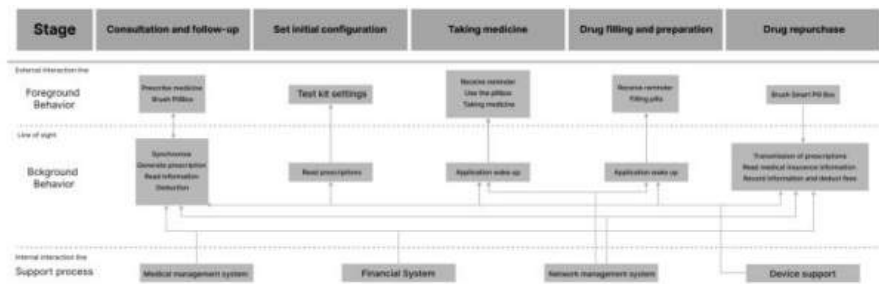


Figure 4: Service blueprint of the elderly medication service system.



Figure 5: Design of smart medicine box service system.

### Design Example of Intelligent Medicine Box

Based on the SAPAD mapping model, a smart medicine box wristwatch was designed, which was developed into a wearable device by integrating its intelligent module into a wristwatch.

Medication reminders, synchronous doctor order prescriptions, and NFC-simulated medical insurance cards are the three characteristic functions of the smart pillbox watch in terms of intelligence and information (see Figure 5).

The wearable smart pillbox watch adopts a magazine-type loading structure. When taking medicine, the medicine-taking buttons on both sides are clicked, where after the dial opens and the pills in the magazine are automatically sent out by the spring. The user can press the “taking medicine button” to take medicine according to the number of medicines displayed on the dial.



**Figure 6:** Design of smart kit dial.



**Figure 7:** Wearable smart pillbox watch.

## CONCLUSION

The primary focus of this paper is to explore the optimization of intelligent drug treatment services to address medication adherence issues among elderly patients with chronic diseases. The study proposes a design thinking-driven method based on the SAPAD model, integrating services and experiences into drug services to establish a sustainable medical service system. Based on the aforementioned research, a design example of a wearable smart pillbox watch is proposed to optimize the behavior of the elderly during drug treatment.

In summary, the paper presents the following innovative points: In terms of research methodology, this paper employs the SAPAD model to extract precise and detailed optimization strategies from chaotic systems,

driving design thinking to integrate new technologies across disciplines into medication services and sustainable health industries; In terms of design practice, the paper demonstrates the potential of wearable technology to improve medication management among the elderly, and has preliminarily constructed a sustainable health service system for the chronic medication of the elderly, providing reference cases for future research and development in this field.

However, due to inherent limitations, the study presented in this paper has areas that necessitate further enhancement:

- a. A more comprehensive and extensive sample is required for observation and validation purposes.
- b. Design practices should concentrate on elderly individuals suffering from various chronic conditions and undertake targeted investigations.
- c. The proposed design methodology requires broader application and refinement across a range of practical cases to substantiate and elevate its applicability and generalizability.

Subsequent research will be conducted with the aim of refining and advancing upon this groundwork.

## REFERENCES

- Abdi, M. R., & Labib, A. W. (2003). A design strategy for reconfigurable manufacturing systems (RMSs) using analytical hierarchical process (AHP): A case study. *International Journal of production research*, 41(10), 2273–2299.
- Abookire, S., Plover, C., Frasso, R., & Ku, B. (2020). Health design thinking: An innovative approach in public health to defining problems and finding solutions. *Frontiers in public health*, 8, 459.
- Annweiler, C., Noublanche, F., Boureau, A. & Berrut, G. (2023). Design thinking and clinical living labs: Two complementary approaches for addressing complex health challenges in geriatric medicine. *Maturitas*, 173, 121–122.
- Chiang-Hanisko, L., Tan, J. Y., & Chiang, L. C. (2014). Polypharmacy issues in older adults. *Hu li za zhi*, 61(3), 97.
- Faisal, S., Ivo, J., & Patel, T. (2021). A review of features and characteristics of smart medication adherence products. *Canadian Pharmacists Journal/Revue des Pharmaciens du Canada*, 154(5), 312–323.
- Hu, F., Li, J., Wang, W., & Sato, K. (2019, July). Meaningful experience in service design for the elderly: SAPAD framework and its case study. In *Proceedings of the Design Society: International Conference on Engineering Design (Vol. 1, No. 1, pp. 3081–3090)*. Cambridge University Press.
- Hu, F., Sato, K., Zhang, X., & Zhu, T. (2013). Semiotic basis for designing product architecture. In *DS 75-4: Proceedings of the 19th International Conference on Engineering Design (ICED13), Design for Harmonies, Vol. 4: Product, Service and Systems Design, Seoul, Korea, 19-22.08. 2013*.
- Hu, F., Zhang, X., Sato, K., Teeravarunyou, S., & Lin, H. (2015, October). Human-centred Product Architecture from UPPA to SAPAD. In *2015 International Conference on Sustainable Energy and Environmental Engineering (pp. 112–115)*. Atlantis Press.
- Jimmy, B., & Jose, J. (2011). Patient medication adherence: Measures in daily practice. *Oman medical journal*, 26(3), 155.



- Karayalcin, I. I. (1982). *The analytic hierarchy process: Planning, priority setting, resource allocation*: Thomas L. SAATY McGraw-Hill, New York, 1980, xiii+287 pages, £ 15.65.
- Lam, W. Y., & Fresco, P. (2015). *Medication adherence measures: An overview*. BioMed research international, 2015.
- Lin, M. C., Wang, C. C., Chen, M. S., & Chang, C. A. (2008). Using AHP and TOPSIS approaches in customer-driven product design process. *Computers in industry*, 59(1), 17–31.
- Liu Yan & Wan Xuan. (2024) Design and Research of a Guided Medical Consultation Service System to Alleviate Healthcare Anxiety among the Elderly. *Packaging engineering* 1–13.
- Liu, X., & Varshney, U. (2020). Mobile health: A carrot and stick intervention to improve medication adherence. *Decision Support Systems*, 128, 113165.
- Roebuck, J. (1979). When does “old age begin?": The evolution of the English definition. *Journal of social history*, 12(3), 416–428.
- Thakur, N.; Han, C. Y. (2022) A Simplistic and Cost-Effective Design for Real-World Development of an Ambient Assisted Living System for Fall Detection and Indoor Localization: Proof-of-Concept. *Information*, 13, 363.
- Vaidya, O. S., & Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of operational research*, 169(1), 1–29.
- Yang, B., & He, R. (2023). Research on online learning interactive product design based on SAPAD and scenario-based thinking. *Packaging Works* (02), 167–179. doi: 10.19554/j.cnki.1001–3563.2023.02.019.
- Zhang, X., HU, F., Zhou, K., & Sato, K. (2017, July). Reflecting Meaning of User Experience Semiotics Approach to Product Architecture Design. In *Trans disciplinary Engineering: A Paradigm Shift: Proceedings of the 24th ISPE Inc. International Conference on Trans disciplinary Engineering*, July 10–14, 2017 (Vol. 5, p. 329). IOS Press.

# Conceptual Approach of an Online Correction System for the Stent Production

Yuna Haas and Eric Sax

Institut fuer Technik der Informationsverarbeitung (ITIV), Karlsruhe Institute of Technology (KIT), Karlsruhe, 76131 BaWü, Germany

## ABSTRACT

Stents are the most common form of treatment for coronary heart disease (CHD). Therefore, in Germany, in 2020, 298,557 stents were implanted. Nonetheless, they are relatively expensive. According to the German fee-per-case system, the cost of a single stent can range from 42.17 € up to 1,391.27 €. One possible reason for these costs is the lack of an automated inspection and correction system for maypole braided stents. In this paper, a concept for an online correction system of the stents' geometry during production is proposed. In contrast to existing proposals, the concept does include the un- and re-braiding of the stent if necessary. This leads to existing errors in the stent being corrected rather than only focussing on the future braiding process. This can on the downside lead to a recursive un-braiding of the complete stent. Therefore, a recursion-prevention is included. Further, multiple options to compute the adapted take-up speed of the Mandel, including a mathematical as well as an AI-based approach, are discussed. Moreover, the concept can handle a complete description of the geometry to be produced, as well as a description based on the mandrels' take-up speed, which is more common for stent producers. All in all, the concept contains three steps. In the first steps, it is detected, if a correction is necessary and the recursion-prevention is applied. In the next step, the number of braid cells, that have to be un-braided, as well as the adapted take-up speed are computed. In the last step, the communication of the changed braiding parameters to the maypole braider, as well as the propagation of the take-up speed regarding the remaining production process, are handled.

**Keywords:** Stents, Braiding, Braid control, AI, Error correction

## INTRODUCTION

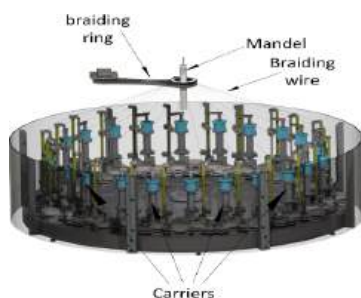
In Germany, 985,572 citizens died reportedly in the year 2020 (Radtke, 2023). The leading cause of death are diseases of the circulatory system (36.4 %) (Radtke, 2023). The coronary heart disease (CHD) belongs to this group of diseases and represents, with 121,725 deaths, the most common (single) cause of death (Herzstiftung, 2022). The main cause of CHD is a narrowing of the coronary vessels, which will lead to an insufficient supply of oxygen and nutrients to the heart muscle (Schlüssel et al., 2022). This undersupply can cause heart failure, cardiac arrhythmia or heart attack (Schlüssel et al., 2022). A possible treatment for this narrowing, and therefore the CHD, is the implantation of a stent (Schlüssel et al., 2022). The stent will widen the

narrowing and therefore restore the supply of oxygen and nutrients (Schlüssel et al., 2022). Due to this reason, there were 298,557 stent implantations in Germany (Herzstiftung, 2022). Incorporating that, according to the German fee-per-case system, a single stent costs 42.17 € up to 1,319.27 € (section 17b of the German Hospital Financing Act, in the version for the year 2023), the financial burden on the public sector is evident. A possible reason for these high costs must be the lack of an automated inspection and correction system during stent production with a maypole braider (Bermudez et al., 2017). Currently, the stent is manually inspected for errors after its production (Bermudez et al., 2017). If there is an error in e.g. its geometry, the stent is disposed of (Haas et al., 2022). Therefore, an inspection system, which can detect and correct geometry errors, during production, is desirable.

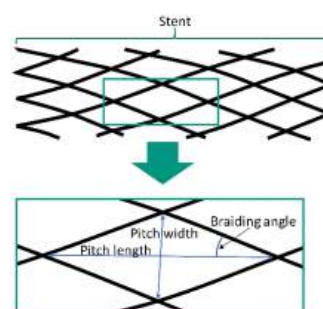
## FUNDAMENTALS

### Stent Production

A maypole braider (s. Figure 1), which can be used to produce stents, contains three main components. The first one is the carriers, which move in a pairwise opposite direction and are propelled by horn gears. In addition, they store the braiding wire. The braiding wires' other end is fixed at the mandrel, which corresponds to the second main component. During braiding, the mandrel moves vertically away from the carriers' movement plane. Its movement speed is called take-up speed. The last component is the braiding ring. It is being used to control the stents' geometry. Apart from that, the mentioned take-up speed, the tension, as well as the diameter of the braiding wire, the to-be-braided pattern, and the mandrels' shape, influence the stents' final geometry. This geometry can be described as a mesh (s. Figure 2). A single mesh, or braiding cell, can be characterized using its pitch length and width or braiding angle. The pitch length and width are defined as the distance between the top- and bottommost respectively the left- and rightmost interlacing point of the braiding wire of a braiding cell. The braiding angle describes the angle between the braiding wire and the stents' axis (Kyosev, 2014) Based on this mesh structure, a row and column index can be assigned to each braid cell (Haas, Stang et al., 2021).



**Figure 1.** Schematic visualization of a maypole braider and its main components (Haas, Braeuner et al., 2021).



**Figure 2.** Schematic visualization of a stent as a mesh as well as its braiding angle, pitch length and width.

## RELATED WORK

Currently, there aren't any systems or conceptual work regarding an inspection and correction system for maypole braided stents. However, there are systems, which measure the braiding angle using e.g. a Fourier transformation (e.g. (Jiyong et al., 2023; Ershov et al., 2022; Yang et al., 2022) or (Vollbrecht et al. 2021)), a Hough Transformation (Monnot et al., 2017) or Convolutions Neural Networks (CNNs) (Şerban & Barsanescu, 2020) of composite braids. These methods have in common, that the measurement is performed over multiple braid cells. Therefore, an error localization on braid cell level is not possible. In Addition, some of these systems adapt the future braiding process based on braiding angle measurements (e.g. (Jiyong et al., 2023; Yang et al., 2022) or (Vollbrecht et al., 2021)). This adaption does not include a correction of the observed error and lacks a precise error localization.

## REQUIREMENTS

If an error correction system is implemented, then requirements regarding the maypole braider as well as the measurement system arise. The maypole braider needs to be able to start, stop, and adapt the braiding process at will and in real time. In doing so, the correction process can be started as soon as an error is detected. Moreover, it must be able to braid and un-braid to a specific braid cell at a specified index with or without adapted parameters. This is necessary because the stent has to be un-braided up to a specific braid cell and afterward resume the braiding process using adapted parameters. This requires the braider to know the index of the currently braided braid cell. Otherwise, it cannot un-braid to a specific one. If the maypole braider does not know the current braid cell index, then the error correction system has to compute it. This conversion would require insight into the maypole braiders' setup and parameterization. Following this, additional computing resources would be necessary. Therefore, to operate more resource-efficient and preserve a decoupled system, the maypole braider needs to be able to operate braid cell indexed. Moreover, the braider needs to offer some kind of programmable interface (API). Otherwise, the correction system cannot communicate with the machine. In this paper, only changes in the mandrels' take-up speed are addressed. Therefore, the API needs to provide read and write access to the mandrels' braid cell indexed take-up speed.

Further, the measurement system must be able to measure and locate single-braid cells in real-time, e.g. as proposed in Haas & Sax (2023). The location needs to contain a row and column index (s, t). Otherwise, a comparison with the desired structure is not possible. In Addition, it should be executed independently from the correction system. This is necessary because both systems have to be executed in parallel. Otherwise, it would be impossible to measure during a correction.

Lastly, the expected geometry specifying the braid angle and/or the pitch length and width as well as additional braid-specific parameters, like the mandrels' diameter, are needed. One possible implementation method would be to extend the braiders API.

## CONCEPT

### Approach Classification

Before presenting the concept, two characteristics of the proposal must be specified:

1. **Expected geometry:** Is the expected geometry a description based on single braid cell definitions or is it a time-dependant take-up speed profile?
2. **Range:** Should the system operate area-based or single braid cell based?

The first point addresses the required description of the expected braids' geometry. This paper covers two different description methods. The first one is based on a description of each single braid cell. This can be interpreted as some sort of table including the index of each cell and its corresponding size. The second one is the description of the mandrels' take-up speed as well as the horn gears' speed. Using these two factors, the expected geometry can be computed (idealized).

The second point discerns if the system corrects every single braid cell (single braid cell based) or if it accepts deviations (area-based). The area-based approach is restricted to accepting deviations only

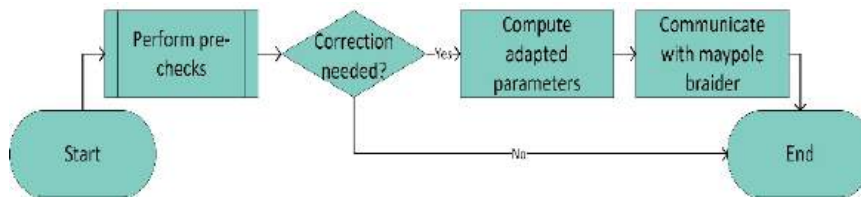
1. If the braid cell is located in an area of a stent where a geometry change (geometry changing area) occurs, and
2. If the deviation converges to the expected value over a specified number of braid cells.

This can lead to a reduction of correction cycles, accepting deviations from the specification.

In practice, the two characteristics correlate with each other. Meaning, that if the expected geometry is based on single braid cell definitions, then one would assume a single braid cell based correction to happen. Otherwise, the geometry would not match the specification. If the expected geometry is based on the mandrel's take-up speed, then the implicit braid cell description is idealized and therefore faulty. In this case, the system can, using the area-based approach, accept deviations from this description to avoid unnecessary correction cycles.

### Top-Level Concept

The proposed concept contains three major steps (s. Figure 3). In the first step, pre-checks are performed. Then, the parameters to adjust the braiding process are computed. Lastly, the results have to be communicated. Due to the measurement system measuring a single braid cell per measurement, the correction system is based on the correction of a single braid cell as well. The case of measuring multiple braid cells at once isn't part of this paper.



**Figure 3:** Visualization of the three major steps in the correction process.

### Pre-Checks

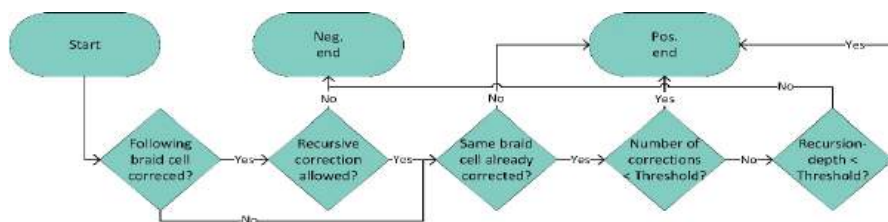
In the scope of the pre-checks, four different aspects will be handled:

1. Does the deviation of the measured from the expected geometry exceed a defined threshold?
2. Is the stent being un-braided?
3. Is a recursive correction happening and allowed?
4. Are the approaches prerequisites fulfilled?

The first step tests, if a deviation occurred, which should be corrected. In specific, the measured braid cell is compared to the expected one. If the deviation is smaller than a certain value (e.g. as the error described in (Haas, Braeuer et al., 2021)), then it is ignored and no correction will be applied.

The second step is to verify if an un-braiding is happening. Because of the inversed movement speed of the mandrel during the un-braiding, all inspected braid cells have been inspected before. Therefore, there shouldn't be any corrections necessary.

The third step handles recursions (s. Figure 4). At first, it is checked if a braid cell with a higher row index exists, which was previously corrected and if recursions are allowed (by the user). If both questions are answered with yes, then it is checked, whether the same braid cell was previously corrected. If so, the number of previous corrections is compared to a threshold value. The same procedure will be repeated for the depth of recursion. This will prevent an infinite correction loop of a single braid cell as well as the whole stent.

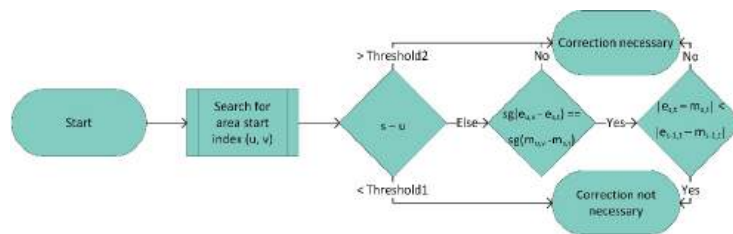


**Figure 4:** Recursion prevention procedure.

In the last step, it is analyzed, whether the approaches' prerequisites are fulfilled. In the single braid cell case, this step can be skipped, because the deviation of a single braid cell suffices to trigger a correction. The area-based one contains the following steps (s. Figure 5):

1. Search for the row and column index of the braid cell  $(u, v)$  at the start of the geometry changing area. This braid cell will be used as a reference.
2. Compute the difference between the row index of the currently braided cell  $s$  and  $u$ . Then, compare the value:
  - a. If  $s - u < \text{Threshold1}$ , then a correction is not necessary.
  - b. If  $s - u > \text{Threshold2}$ , then a correction is necessary.
  - c. Otherwise continue with the procedure.  
Using  $\text{Threshold1}$  and  $\text{Threshold2}$ , a minimal and maximal size of the geometry changing area can be defined.
3. Check if the sign of the difference of the expected geometry at  $(u, v)$  and  $(s, t)$  is the same as the measured one at  $(u, v)$  and  $(s, t)$ . If it isn't, then a correction is necessary. This ensures that the change in geometry has the same sign. This means that if e.g. the expected geometry increases, the measured one should as well.
4. Compute the absolute difference between the expected and measured geometry at  $(s, t)$  and  $(s-1, t)$ . If the value at  $(s, t)$  is smaller than the one at  $(s-1, t)$ , then a correction is not necessary. Otherwise, the error increases, and a correction is necessary. If  $\text{Threshold1}$  is smaller than one or not being used, then this step should be skipped if  $s$  is equal to  $u$ .

If a correction is necessary, the correction system uses the maypole braiders' API to stop the braiding process.



**Figure 5:** Prerequisites check of the area-based approach with  $e_{a,b}$  and  $m_{c,d}$  being the expected and measured braid cell at index  $(a,b)$  and  $(c,d)$ .

### Compute Adapted Parameters

In the next step, the number of un-braided braid cells as well as the take-up speed adjustment(s) are computed. The procedure changes depending on the chosen approach (single braid cell vs. area-based). The cause for this is, that in the single braid cell case, all braid cells are assumed to be correct up to the incorrect one. Therefore, the un-braiding will always be performed up to the incorrect one. In the area-based case, the un-braiding can be performed up to any braid cell in the geometry changing area.

### Single Braid Cell Approach

To compute the take-up speed adaption using the single braid cell approach, multiple procedures are possible. In this paper, three possibilities are

explained briefly. The first one is mathematical<sup>1</sup>. It uses the deviation between the expected and measured geometry. In the case of the pitch length being the measurement criteria, the new take-up speed  $v_{new}$  can be computed as shown in Equation 1. In this equation  $v_{old}$ ,  $a_e$ ,  $a_m$  and  $rpm$  describe the previously used take-up speed, the expected and measured pitch length as well as the horn gears RPM.

$$v_{new} = v_{old} + (a_e - a_m) * rpm \quad (1)$$

The second one is a regression based on deep learning (DL). In specific, a deep neural network will be used to compute the deviation directly. The input to the networks needs to cover important information like the deviation of the measured to the expected geometry or the mandrels' diameter. The last one is (deep) reinforcement learning (RL). The (RL) states correspond with the expected and measured geometry, machine parameters like the current take-up speed, and other relevant parameters like the mandrels' diameter. The (RL) action is defined as an adaption of the take-up speed. Lastly, the reward corresponds to the deviation between the expected and the measured geometry.

### Area-Based

The area-based approach consists of two steps: First compute the number of braid cells being un-braided and second, compute the take-up speed adaption, except for (deep) RL, which will compute both parameters at once.

Two possibilities to compute the number of un-braided braid cells are:

1. Rule-based: Always un-braid to the start of the geometry change area. This is valid because deviations are allowed, as long as the deviation is lower than a specified threshold.
2. Deep Learning: Using a DL-based regressor, the value can be computed using braid-relevant information like the expected and measured braid cell and the mandrels' diameter.

Afterward, one of the methods described in the previous chapter (Single braid cell approach) can be used. In the second case (DL-based regressor), the regressors' input will be extended such that the number of un-braided braid cells will be used as well.

Alternatively, RL can be used to compute both parameters at once. In this case, the (RL) state corresponds to the expected and measured geometry and relevant machine parameters like the mandrels' diameter. An (RL) action will be to compute the number of un-braided braid cells as well as the adapted mandrels' take-up speed profile for these braid cells. The (RL) reward will be defined as the negative absolute difference between the expected and measured geometry at the correction triggering braid cell.

<sup>1</sup>Due to the scope of this paper, the mathematical approach is held rather simple. More (mathematical) complex approaches can be found in e.g. (Jiyong et al., 2023).



### Communicate With Maypole Braider

In the last step, the communication with the maypole braider is handled. This communication can be divided into four steps: 1. Offset-computation, 2. Communicate adjustments, 3. Reset internal states and 4. Resume braiding. In the first step, an offset has to be added to the number of un-braided braid cells. The reason for this is, that the measurement system detects finished braided cells. Therefore, there is a discrepancy between the currently braided and detected braid cells. This discrepancy can be computed by comparing both indices. The first one can be retrieved using the braiders API, while the second one is a result of the measurement system. The difference between those two values corresponds to the mentioned offset. In the second step, the adjusted take-up speed profile has to be communicated to the maypole braider using its API. In the third step, internal states have to be reset. E.g. if a braid cell is un-braided, then all previous measurements of its geometry are invalid. In the last step, the braiding has to be resumed.

### DISCUSSION

Revisiting the requirements, it is assumed, that the maypole braider can operate indexed based. This can, in a practical implementation, be problematic. Due to the restrictions of the embedded system controlling the braider, it is possible, that there aren't enough computing resources to realize this functionality. Therefore, the decision, whether or not the braider can be controlled using an index, is dependent on its as well as the controlling systems computing and network resources. If the correction system could compute the index, then the localization system would not be necessary. This would require, that the measurement system is not noised. Otherwise, the mapping of the measurement to its index would be biased, even if both computations are correct in themselves.

Another requirement is that the measuring system must be able to measure the row and column index. This is necessary for correct localization. Conventional systems often do not provide control over single carriers. Therefore, all braid cells of one row are identical. In this case, a column-wise localization is not necessary and can therefore be discarded.

Further, the two possibilities single braid cell and area-based were introduced. The first one assumes to be a correct description on braid cell level. The second one is interpreted as a faulty description, especially if changes in the braid pattern occur. Following this, one could mistake them for two opposing procedures. In reality, they represent two different use cases. In the first one, a final description of the stent to be produced is available. The second one is based on the braiders' take-up speed profile and therefore cannot represent the stents' geometry correctly.

Regarding the approach prerequisites, there are two thresholds to be set. Both values can be set to trivial values like 1 (Threshold1) and infinity (Threshold2). However, this would make them incapable of controlling the geometry changing areas' size. To choose useful values, insight into the stent being produced is necessary. Due to the reason, that this check is only applicable to the area-based approach, which, per definition, lacks this

knowledge, it is likely, that trivial values will be chosen making the checks obsolete.

Furthermore, the correction system is not generally applicable. The area-based approach requires the stent to reach the expected value before another change is applied. Otherwise, the approach prerequisites computations are faulty. In Addition, in this case, it is not ensured, that the expected value of the initial geometry change is reached.

## SUMMARY AND OUTLOOK

In summary, a concept is proposed, which corrects a stent's geometry in case of an error during its production using a maypole braider. This concept contains three major steps: At first, pre-checks are performed. This includes e.g. a check if a correction is necessary as well as a (infinite) recursion prevention. If a correction is necessary, then the adapted parameters are computed in the second step. This includes the number of braid cells being unbraided as well as the take-up speed adaption after the un-braiding. For both parameters, different methods (mathematical, DL, and RL) are presented. In the last step, the parameters are communicated to the braider.

Regarding future work, the described methods, which compute the specified parameters, could be implemented and evaluated. In Addition, the proposed system could be extended to support the detection of multiple braid cells at once. This can be realized in two different ways. The first one would be to run this proposal in parallel. In this case, a harmonization of all correction values is needed if an error occurs. Alternatively, the procedure could be changed such that multiple braid cells are processed at once. This could be advantageous because the single braid cells can be set in context.

## ACKNOWLEDGMENT

Parts of this work have been developed in the project Stents4Tomorrow. Stents4Tomorrow (reference number: 02P18C022) is partly funded by the German Federal Ministry of Education and Research (BMBF) within the research program ProMed.

## REFERENCES

- Bermudez, C., Laguarda, F., Cadevall, C., Matilla, A., Ibañez, S., & Artigas, R. (2017). Automated stent defect detection and classification with a high numerical aperture optical system. In *Automated Visual Inspection and Machine Vision II* (p. 103340C).
- Deutsche Herzstiftung. (2022). Mortalitätsrate aufgrund von Herzkrankheiten in Deutschland nach Diagnosegruppen in den Jahren von 1990 bis 2020 (je 100.000 Einwohner).
- Ershov, S. V., Reimer, V., Zastrow, T., Kalinin, E. N., & Gries, T. (2022). Method for measuring the braid angle and its deviation from the specified value in braided preforms using image analysis. *Fibre Chemistry*, 53(5), 346–354.
- Haas, B., Braeuner, M., Lehmann, K., & Sax, E. (2021, October). Evaluation of different methods to measure a stent's pitch length in an industrial environment.

- In 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) (pp. 1–6). IEEE.
- Haas, B., Erlinghagen, L., & Sax, E. (2022, November). Pitch length measurement of stents using dynamically cropped images. In 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) (pp. 1–6). IEEE.
- Haas, B., & Sax, E. (2023). Conceptual measurement of individual pitches during the stent production. *Proceedings of 17th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing*, 115–122.
- Haas, B., Stang, M., Khan-Blouki, V., & Sax, E. (2021). Introduction of an Algorithm Based on Convolutional Neural Networks for an Automated Online Correction of Braided Cardiovascular Implants. In *Human Interaction, Emerging Technologies and Future Applications IV: Proceedings of the 4th International Conference on Human Interaction and Emerging Technologies: Future Applications (IHET-AI 2021)*, April 28-30, 2021, Strasbourg, France 4 (pp. 28–36). Springer International Publishing.
- Hoffstetter, M., Pfeifer, S., Schratzenstaller, T., & Wintermantel, E. (2009). Stenting und technische Stentumgebung. *Medizintechnik: Life Science Engineering*, 1263–1296.
- Jiyong, F., Yujing, Z., Zhongwei, W., & Zhiguo, Y. (2023). Traction control of space tubular shaped mandrel and detection of preform braiding angle. *Textile Research Journal*, 93(1-2), 392–408.
- Kyosev, Y. (2014). *Braiding technology for textiles: Principles, design and processes*. Elsevier.
- Monnot, P., Lévesque, J., & Lebel, L. L. (2017). Automated braiding of a complex aircraft fuselage frame using a non-circular braiding model. *Composites Part A: Applied Science and Manufacturing*, 102, 48–63.
- Rainer Radtke. (2023). Anzahl der Todesfälle nach den häufigsten Todesursachen in Deutschland in den Jahren 2019 bis 2021.
- Schüssel, K., Weirauch, H., Schlotmann, A., Brückner, G., & Schröder, H. (2022). *Gesundheitsatlas Deutschland. Koronare Herzkrankheit: Verbreitung in der Bevölkerung Deutschlands: Ursachen, Folgen und Präventionsmöglichkeiten [White paper]*. Wissenschaftliches Institut der AOK (WIdO).
- Şerban, A., & Barsanescu, P. D. (2020, December). Automatic detection of fibers orientation on composite laminates using convolutional neural networks. In *IOP Conference Series: Materials Science and Engineering* (Vol. 997, No. 1, p. 012107). IOP Publishing.
- Vollbrecht, B., Kohler, C., Kolloch, M., Jung, F., Grigat, N., & Gries, T. (2021). Developing a camera-based measuring system to feedback control the fibre orientation for the braiding process of CFRP. *Advances in Industrial and Manufacturing Engineering*, 3, 100059.
- Yang, M., Yu, L., Wong, C., Mineo, C., Yang, E., Bomphray, I., & Huang, R. (2022). A cooperative mobile robot and manipulator system (Co-MRMS) for transport and lay-up of fibre plies in modern composite material manufacture. *The International Journal of Advanced Manufacturing Technology*, 1–17.

# Sustainable Use of Resources in Hospitals: A Machine Learning-Based Approach to Predict Prolonged Length of Stay at the Time of Admission

Paolo Perlitì Scorzoni, Anita Giovanetti, Federico Bolelli,  
and Costantino Grana

Università degli Studi di Modena e Reggio Emilia - Dipartimento di Ingegneria “Enzo Ferrari”, Modena, Italy

## ABSTRACT

**Introduction:** Length of Stay (LOS) and Prolonged Length of Stay (pLOS) are critical indicators of hospital efficiency. Reducing pLOS is crucial for patient safety, autonomy, and bed allocation. This study investigates different machine learning (ML) models to predict LOS and pLOS.

**Methods:** We analyzed a dataset of patients discharged from a northern Italian hospital between 2022 and 2023 as a retrospective study. We compared sixteen regression algorithms and twelve classification methods for predicting LOS as either a continuous or multi-class variable (1–3 days, 4–10 days, >10 days). We also evaluated pLOS prediction using the same models, having pLOS defined as any hospitalization with LOS longer than 8 days. We further analyzed all models using two versions of the same dataset: one containing only structured data (e.g. demographics and clinical information), whereas the second one also containing features extracted from free-text diagnosis.

**Results:** Our results indicate that ensemble models achieved the highest prediction accuracy for both LOS and pLOS, outperforming traditional single-algorithm models, particularly when using both structured and unstructured data extracted from diagnoses.

**Discussion:** The integration of ML, particularly ensemble models, can significantly improve LOS prediction and identify patients at increased risk of pLOS. This information can guide healthcare professionals and bed managers in making informed decisions to enhance patient care and optimize resource allocation.

**Keywords:** LOS, pLOS, Machine learning, Hospital admissions, Public healthcare, Sustainability

## INTRODUCTION

Over the past twenty years, the number of beds in Italian hospitals has decreased by 30 percent (OECD, 2021) while bed occupancy has notably increased, causing congestion and lengthening patient stays. With 11.6 beds per 100,000 inhabitants in 2021, Italy still maintains a provision below the Organisation for Economic Cooperation and Development average of 16.9 (OECD, 2023). Moreover, a significant escalation in national health

expenditure, from approximately 80 billion euros in 2002 to 129 billion euros in 2022 (Italian Ministry of Economy and Finance, 2023), with 20% attributed to inpatient expenses, has placed hospitals under considerable pressure to implement effective cost containment strategies. Several trends, including the rise of patient volumes and changing working practices (as the COVID-19 pandemic forced to do), contribute to the complexity of managing patient flow. Furthermore, an aging population, higher co-morbidity rates, and delayed discharges (especially into nursing and residential care) add to the challenges.

Current bed modeling techniques, often based on midnight occupancy census, lack the information needed for effective space management. At the same time, the advent of AI and Machine Learning (ML) may assist bed managers in their daily routine, offering valuable tools.

## BACKGROUND

The Length of Stay (LOS), represented as the interval time between admission and discharge (i.e., total bed-days occupied by a patient), plays a fundamental role in evaluating the quality of care services.

In the literature, it has been demonstrated that LOS does have correlations to disease severity, readmission rates, and mortality (Marfil-Garza et al., 2018). Reducing LOS in public healthcare systems (i.e. enabling early discharge and fast turnover) benefits both patients – preventing complications, limiting the risk of adverse events (Ackroyd-Stolarz et al., 2011) such as falls, thrombosis, drug reactions and hospital-acquired infections, improving patient autonomy (Hauck and Zhao, 2011) – and hospitals – optimizing treatment plans and resource utilization (such as bed allocation) (Molloy et al., 2017), up to maintaining control over the growth of waiting lists. A prolonged Length of Stay (pLOS), in fact, is associated with cognitive impairment, functional limitations, and elevated burdens of comorbidity (Bo et al., 2016), and it generally leads to cancellations of elective operations, additional use of resources and thus increased medical costs (especially in ICU units); pLOS could also adversely affect the admission process for critically ill patients and hinder timely access to treatment. Moreover, a small percentage of patients with pLOS could consume a large proportion (up to 50%) of the limited resources available (Evans et al., 2018).

With an aging population, the anticipated increase of LOS underscores the urgency of responsive healthcare planning to ensure safety, satisfaction, and accessibility. This is particularly evident in Italy, where demographic shifts like rising life expectancy (reaching 80.6 years for men and 84.8 for women in 2022) (ISTAT, 2022) contribute to the prevalence of chronic and degenerative diseases. Identifying inpatients with extended stays – often referred to as ‘bed-blockers’ – necessitate early recognition for appropriate treatment planning, including potential admission to long-stay wards. Hence, prolonged LOS serves as a key metric, being associated with escalated costs and reduced capacity.

The present study aimed to extensively evaluate ML-based models to predict both LOS and pLOS for general patients. A comprehensive comparison among these methods was performed to assess their respective capabilities. Additionally, an investigation was conducted to identify the most relevant features for predicting LOS and pLOS.

## RELATED WORKS

Over the last two decades, various prediction models employing statistical techniques have been developed to investigate LOS and the influence of covariates such as age, gender, diagnosis, illness severity, type of admission, and hospital characteristics. Nevertheless, the use of Machine Learning and Deep Learning (DL) has recently gained attention in health service research as a viable alternative to those established methodologies (Gholipour et al., 2015; Barnes et al., 2016). There is a quite heterogeneous literature exploring LOS patterns (Stone et al., 2022), often focusing on broader patient cohorts (Mekhaldi et al., 2020), with particular emphasis on limited age intervals (Thompson et al., 2018; Hesselink et al., 2019), explicit discipline areas (Chen et al., 2023) and medical specialties (e.g., cardiology, Daghistani et al., 2019), surgical procedures (Chuang et al., 2018) and cancer surgeries (Jo et al., 2021), and specific medical units.

Unfortunately, only a minority of these reviews take Italian public health into consideration. In a study by Zeleke et al. (2023) six classification algorithms were developed to study the prediction of pLOS for 12,858 inpatients admitted through the emergency department (ED) of an Italian hospital (“Sant’Orsola-Malpighi University Hospital”, Bologna). The pLOS threshold was defined as any stay longer than the average LOS (6 days). The authors also developed eight regression models for LOS prediction. The Gradient Boosting classifier best predicted pLOS (accuracy 75%, AUC 75.4%). Ridge and XGBoost regressors best predicted LOS, with an overall prediction error between 6 and 7 days.

Trunfio et al. (2022) analyzed LOS for 2,515 patients undergoing hip-replacement surgery at the “San Giovanni di Dio e Ruggi d’Aragona” University Hospital of Salerno. Several regression and classification algorithms were implemented in order to predict the total length of stay. The results from the regression analysis showed that the best model was Multiple Linear Regression ( $R^2$  0.616). As for classification analysis, LOS was divided into 3 classes (LOS  $\leq$  6 days, 6 days  $<$  LOS  $\leq$  12 days, LOS  $>$  12 days). In terms of overall accuracy, Random Forest achieved a value of 71.76%.

Olivato et al. (2022) developed a machine learning-driven system to forecast pLOS for COVID-19 patients. Their model, trained on demographic information and laboratory test values from over 6,000 patients admitted to the Spedali Civili di Brescia, in northern Italy, achieved a ROC-AUC of 0.76.

In another study conducted at A.O.R.N. “Antonio Cardarelli”, Naples (Italy) (D’Onofrio et al. 2023), the Electronic Medical Records (EMRs) of 989 patients who underwent mastectomy surgery were used to predict LOS as a binary outcome. Random Forest showed the best accuracy (77.79%).

Age and the presence of comorbidities such as hypertension, diabetes, obesity, and tumor stage were considered factors affecting the LOS.

Di Matteo et al. (2023) implemented a custom neural network to forecast the length of stay (as a binary outcome) for patients undergoing hip or knee arthroplasty in “Humanitas Research” Hospital, Milan (Italy). The model, trained on a dataset of 1,517 patients, leveraged a combination of clinical and textual data. They achieved an AUC of 0.789 in predicting “Short LOS” ( $\leq 6$  days) over “Long LOS” ( $> 7$  days).

Altogether, these findings suggest that the ML approach could assist hospital systems in forecasting and addressing bed capacity requirements. However, most of these studies focus only on a specific department or rely on data not readily available at admission (such as lab results).

Our research employs a variety of supervised ML algorithms to predict the length of stay for patients in general inpatient settings, examining LOS as a continuous, multi-class, and dichotomous variable using data extracted from a medico-administrative platform. This investigation intentionally encompasses all medical-surgical departments with the aim of developing robust and adaptable models for effective generalization. This decision is supported by the observation that when a department reaches full capacity, patients are relocated to alternate wards with available beds (often regardless of their primary service). Consequently, evaluating all medical units collectively provides greater consistency. Additionally, our analysis solely considers data available at the time of admission. While this approach may potentially overlook a substantial portion of clinical data, it ensures good performance from the very beginning of hospitalization and can be implemented across diverse hospital settings.

## **MATERIALS AND METHODS**

The present study examines the departments of a general hospital located in Emilia-Romagna, Italy. The facility is organized by intensity care and structured around 19 clinical units. We analyzed a dataset of 12,471 hospitalizations from 10,145 unique patients discharged between February 2022 and November 2023, with a length of stay of at least 24 hours. Data were extracted from an EBMS (Electronic Bed Management System) which included information on patient demographics, admission type, clinical features, and hospitalization features.

Patients undergoing Day Surgery or Day Hospital procedures were excluded from the analysis due to their predetermined LOS. This exclusion aimed to mitigate potential biases in the model’s performance arising from an overrepresentation of cases with a fixed LOS of one day. To ensure the quality and integrity of the data, we also excluded patients deceased during hospitalization, inpatients with LOS values exceeding the 99.95th percentile of the LOS distribution (outliers), and maternity/infancy wards due to their distinct clinical characteristics and potential data collection biases.

### **Models Development**

The study employed two versions of the dataset: the first one (A) containing only structured data (demographics, clinical information, admission details)

and the second one (B) incorporating features extracted from unstructured free-text diagnoses documented by practitioners.

Initially, fourteen regression algorithms were employed for predicting LOS as a continuous value. Performance was evaluated using mean absolute error (MAE), root mean squared error (RMSE), R-squared ( $R^2$ ), and adjusted R-squared scores. Additionally, ten classification methods were utilized to predict LOS as a multi-class target (1–3 days, 4–10 days, >10 days). Evaluation metrics included accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUROC). The same ten classification models were also used to predict pLOS, defined as any hospitalization longer than 8 days, corresponding to the 75th percentile.

To extract features from diagnoses, a text cleaning process was implemented, removing stop words and irrelevant or non-domain-specific terms. A pre-trained BERT model was then employed to tokenize the text and generate embeddings. Moreover, principal component analysis (PCA) was applied to embeddings from dataset B to reduce the dimensionality of the data to 100 components, thereby decreasing computational complexity.

Each dataset was randomly partitioned into a training set comprising 9,976 admissions (80%) and a holdout/validation set encompassing 2,495 admissions (20%). A five-fold cross-validation approach was employed for each task to compare algorithms and identify the top performers. Hyperparameter tuning was then performed for each model. Finally, performance was assessed on the test set.

Additionally, a Voting ensemble and a Stacking ensemble were included in the final evaluation. These methods aggregate predictions from multiple base models (using voting and stacking aggregation respectively) to improve overall accuracy and reduce bias, yielding a superior ensemble model.

## RESULTS

### Regression Models

In predicting LOS as a continuous variable (Table 1), the ensemble Stacking-Regressor demonstrated the highest accuracy for dataset A (MAE 2.81,  $R^2$  score 0.635), followed by VotingRegressor (MAE 2.82,  $R^2$  score 0.634) and XGB-Regressor (MAE 2.84,  $R^2$  score 0.632). When considering datasets containing unstructured data (B), CatBoost outperformed other models in predicting LOS (MAE 2.73,  $R^2$  score 0.649), followed closely by VotingRegressor (MAE 2.72,  $R^2$  score 0.647) and XGBRegressor (MAE 2.78,  $R^2$  score 0.639).

Notably, incorporating embedded representations derived from free-text diagnoses led to a modest yet noticeable performance improvement across various models. This enhancement can be attributed, in part, to the ability of embeddings to capture the semantic meaning of diagnoses, which can be challenging to represent using traditional structured features. Additionally, embeddings offer the advantage of modeling the relationships between different diagnoses, which becomes particularly important in the presence of comorbidities. It is noteworthy that including the admitting diagnosis does not introduce bias or confound the study endpoint (e.g., data leakage), given its inherent availability at the time of hospitalization.



**Table 1.** Results of tuned models on datasets A and B (regression task).

Model	Dataset A				Dataset B			
	MAE	RMSE	R <sup>2</sup>	Ad. R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>	Ad. R <sup>2</sup>
Stacking Regressor	<b>2.806</b>	<b>4.614</b>	<b>0.635</b>	<b>0.633</b>	2.705	4.622	0.633	0.632
Voting Regressor	2.824	4.617	0.634	0.633	2.722	4.537	0.647	0.645
XGB Regressor	2.844	4.634	0.632	0.630	2.776	4.585	0.639	0.638
CatBoost Regressor	2.831	4.639	0.631	0.629	<b>2.726</b>	<b>4.520</b>	<b>0.649</b>	<b>0.648</b>
Linear Regression	2.976	4.692	0.622	0.621	2.911	4.622	0.633	0.632
Ridge	2.963	4.706	0.620	0.618	2.862	4.620	0.634	0.632
GB Regressor	2.946	4.723	0.617	0.616	2.923	4.739	0.615	0.613
LGBM Regressor	2.903	4.770	0.609	0.608	2.756	4.616	0.634	0.633
Elastic-net	2.995	4.776	0.609	0.607	2.887	4.679	0.624	0.623
SVR	2.883	4.784	0.607	0.606	2.792	4.694	0.622	0.620
Lasso	3.051	4.867	0.594	0.592	3.004	4.811	0.603	0.601
RF Regressor	2.968	4.903	0.588	0.586	2.843	4.787	0.607	0.605
KNN Regressor	2.989	5.012	0.569	0.567	2.868	4.957	0.578	0.577
AdaBoost Regress.	3.601	5.275	0.522	0.521	3.576	5.285	0.521	0.519
MLP Regressor	3.298	5.587	0.464	0.462	3.090	5.350	0.509	0.507
DT Regressor	3.882	6.617	0.249	0.246	3.757	6.604	0.252	0.249

### Classification Models

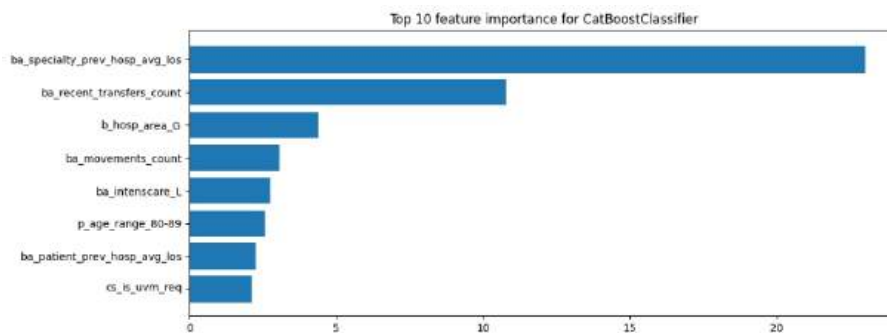
Among the ensemble-based models (Table 2), VotingClassifierSoft exhibited superior performance in predicting LOS for both datasets A (accuracy 73.55%, AUROC 87.94%) and B (accuracy 76.27%, AUROC 89.60%). As observed in the regression analysis, the employment of embeddings extracted from diagnoses evidences a modest augmentation in performance.

**Table 2.** Results of tuned models on datasets A and B (multi-class classification task).

Model	Dataset A				Dataset B			
	Acc.	F1	ROC	PRC	Acc.	F1	ROC	PRC
Voting Soft	<b>0.735</b>	<b>0.732</b>	<b>0.879</b>	<b>0.777</b>	<b>0.763</b>	<b>0.760</b>	<b>0.896</b>	<b>0.794</b>
Stacking	0.732	0.730	0.879	0.774	0.763	0.761	0.893	0.789
CatBoost	0.733	0.729	0.878	0.776	0.758	0.754	0.894	0.788
XGB	0.728	0.724	0.875	0.775	0.762	0.760	0.892	0.789
RF	0.730	0.728	0.873	0.767	0.750	0.745	0.886	0.778
GB	0.729	0.726	0.872	0.764	0.760	0.757	0.890	0.783
LGBM	0.724	0.722	0.870	0.760	0.748	0.745	0.887	0.778
Log. Regress.	0.730	0.727	0.864	0.750	0.740	0.737	0.883	0.773
KNN	0.702	0.697	0.847	0.726	0.718	0.710	0.862	0.733
MLP	0.678	0.679	0.827	0.707	0.719	0.717	0.852	0.729
AdaBoost	0.710	0.707	0.787	0.651	0.738	0.737	0.817	0.677
DT	0.654	0.652	0.724	0.519	0.655	0.655	0.720	0.513

CatBoost was also employed to determine the relative significance of features (Figure 1). Analyzing dataset A, the most salient features in descending order of importance were: the overall average length

of stay for same-service hospitalizations within the previous 30 days (*ba\_specialty\_prev\_hosp\_avg\_los*); the number of recent transfers across wards (*ba\_recent\_transfers\_count*), possibly indicative of increasing medical complexity (i.e. patient cases that require multiple departments or specialized care); the surgery hospitalization area (*b\_hosp\_area\_G*); the number of bed movements during hospitalization (*ba\_movements\_count*), including those within the same ward; low-intensity care level (*ba\_intenscare\_L*), supporting the idea that stays associated with critical conditions require only the time necessary to stabilize the patient's condition; age in range 80–89 years (*p\_age\_range\_80-89*), suggesting that inpatients in this age group may have an increased risk of prolonged stays, potentially due to age-related vulnerabilities or comorbidities; the average LOS for same-patient hospitalizations in the prior year (*ba\_patient\_prev\_hosp\_avg\_los*); and the need for a multidimensional geriatric assessment (*cs\_is\_uvm\_req*), typically associated with elderly or frail individuals.



**Figure 1:** Feature importance for CatBoostClassifier on dataset A (multi-class classification task).

### Binary Classification Models

Similar to the multi-class classification task (Table 3), Area Under the Receiver Operating Characteristic Curve (AUROC) and Area Under the Precision-Recall Curve (AUPRC) were employed to assess the performance of the models. Ranging from 0 to 1, AUROC effectively captures the trade-off between true and false positives across all possible thresholds. Conversely, AUPRC prioritizes the identification of positive samples, making it particularly advantageous for imbalanced datasets.

LogisticRegression provided the most accurate predictions for prolonged length of stay in the A dataset (accuracy 86.61%, AUROC 90.54%), followed by VotingClassifierSoft and CatBoostClassifier. On the other hand, for dataset B, VotingClassifierSoft (accuracy 86.53%, AUROC 91.67%), CatBoostClassifier (accuracy 86.69%, AUROC 91.24%), and StackingClassifier demonstrated superior predictive capabilities for pLOS.

**Table 3.** Results of tuned models on datasets A and B (binary classification task).

Model	Dataset A				Dataset B			
	Acc.	F1	ROC	PRC	Acc.	F1	ROC	PRC
Log. Regress.	<b>0.866</b>	<b>0.646</b>	<b>0.905</b>	<b>0.794</b>	0.866	0.654	0.911	0.798
Voting Soft	0.863	0.636	0.905	0.789	<b>0.865</b>	<b>0.645</b>	<b>0.917</b>	<b>0.808</b>
CatBoost	0.867	0.654	0.903	0.785	0.867	0.654	0.912	0.799
Stacking	0.867	0.654	0.903	0.785	0.867	0.654	0.912	0.799
GB	0.861	0.636	0.901	0.782	0.863	0.650	0.908	0.793
LGBM	0.861	0.642	0.898	0.777	0.862	0.642	0.907	0.791
AdaBoost	0.862	0.642	0.896	0.778	0.853	0.634	0.888	0.769
RF	0.863	0.635	0.893	0.777	0.857	0.584	0.904	0.788
XGB	0.842	0.606	0.884	0.751	0.862	0.655	0.901	0.785
KNN	0.848	0.545	0.884	0.743	0.862	0.516	0.874	0.726
MLP	0.827	0.618	0.867	0.733	0.862	0.625	0.874	0.745
DT	0.799	0.569	0.719	0.423	0.862	0.558	0.713	0.412

## DISCUSSION

A key strength of this study lies in its comparative analysis of different data sources. The integration of text extracted from diagnoses proved instrumental in capturing subtle nuances not represented in structured data alone. This is particularly relevant to leveraging the expertise of medical staff and the outcomes derived from physical examinations. Our findings align with prior investigations (Zhang et al., 2020; Jiang et al., 2023), which have demonstrated that text-derived features may enhance predictive performance.

This study also provides further evidence of the effectiveness of machine learning ensembles in improving LOS prediction accuracy. By aggregating the outputs of multiple base learners, these ensemble models effectively reduce inherent biases and significantly improve overall predictive power. Moreover, their decision tree structure facilitates the conversion of complex models into transparent decision rules, fostering understanding among practitioners. This enhances model acceptance and facilitates integration into healthcare workflows, which is pivotal for the adoption of ML-driven systems in medical practice.

Compared to previous works focused on specific patient groups, such as hip- or knee-replacement patients or heart failure patients, our model is not limited to narrow clinical subsets and encompasses a heterogeneous patient spectrum, expanding its applicability beyond specific diagnoses or conditions. Furthermore, our approach avoids the limitations of singleward-based studies (e.g. ICU or general medicine) that may be constrained by specific internal dynamics influencing patient outcomes. While this methodology poses a greater challenge, it also yields broader generalizability, increasing the potential for real-world implementation. Most importantly, our models leverage readily available data from institutional Electronic Health Records (EHR), collected within the first 24 hours of hospitalization, allowing bed managers to utilize them as decision-making aids promptly upon patient admission.

Despite its strengths, our study also exhibits certain limitations that should be acknowledged. Firstly, it is a retrospective study employing historical data. While it utilizes minimal historical information, such as the number of previous hospitalizations and the average LOS for same-specialty past stays, this may potentially introduce biases. Secondly, the exclusion of vital signs and laboratory test results limits the comprehensiveness of the predictor variables. While it is desirable to have input variables as general as possible, avoiding specific test or laboratory results, these vital data points could potentially enhance the prediction models' accuracy and reduce the risk of false positives (overestimation of LOS) and false negatives (underestimation of LOS).

Eventually, the monocentric nature of the study, relying on data from a single hospital, hinders external validation of the prediction models. While this may restrict their applicability to other healthcare settings, it should be recognized that LOS is also influenced by factors that may vary across facilities: striving for absolute generalizability may not be feasible or optimal in this context.

## CONCLUSION

Our findings underscore the potential of ensemble-based prediction models to assist medical personnel in identifying patients who are at risk of prolonged LOS across different services and wards. By recognizing these inpatients, healthcare providers can proactively implement targeted interventions, such as close monitoring and early discharge planning, to reduce the likelihood of lengthy stays. This preemptive approach can contribute to smoother patient flow, increased bed availability, a lower rate of rescheduled interventions, improved patient satisfaction, and, ultimately, reduced overall healthcare expenditures. Furthermore, the use of readily available data from EHR in conjunction with algorithms that do not necessitate resource-intensive procedures or specialized hardware encourages the potential integration of this methodology in other settings and workflows (e.g. embedding the inference model within EBMSs as a second-opinion tool, to support both medical staff and healthcare management in their daily tasks).

These results emphasize the priority for hospitals to adapt and innovate in order to meet the demands of an aging population with chronic disorders, while containing costs and optimizing resources to improve the sustainability of public healthcare systems.

## REFERENCES

- Ackroyd-Stolarz, S., Guernsey, J. R., Mackinnon, N. J., & Kovacs, G. (2011). The association between a prolonged stay in the emergency department and adverse events in older patients admitted to hospital. *BMJ Quality & Safety*.
- Barnes, S., Hamrock, E., Toerper, M., Siddiqui, S., & Levin, S. (2016). Real-time prediction of inpatient length of stay for discharge prioritization. *Journal of the American Medical Informatics Association*, 23(e1), e2–e10.
- Bo, M., Fonte, G., Pivarolo, F., Bonetto, M., Comi, C., Giorgis, V., Marchese, L., Isaia, G., Maggiani, G., Furno, E., Falcone, Y., and Isaia, G. C. (2016). Prevalence of and factors associated with prolonged length of stay in older hospitalized medical patients. *Geriatrics & gerontology international*, 16(3), 314–321.

- Chen, R., Zhang, S., Li, J. et al. (2023). A study on predicting the length of hospital stay for Chinese patients with ischemic stroke based on the XGBoost algorithm. *BMC Medical Informatics and Decision Making*, 23(1), 1–10.
- Chuang, M. T., Hu, Y. H., & Lo, C. L. (2018). Predicting the prolonged length of stay of general surgery patients: a supervised learning approach. *International Transactions in Operational Research*, 25(1), 75–90.
- Daghistani, T. A., Elshawi, R., Sakr, S., Ahmed, A. M., Al-Thwayee, A., & Al-Mallah, M. H. (2019). Predictors of in-hospital length of stay among cardiac patients: a machine learning approach. *International journal of cardiology*, 288, 140–147.
- Di Matteo, V., Tommasini, T., Morandini, P., Savevski, V., Grappiolo, G., Loppini, M. Machine Learning Prediction Model to Predict Length of Stay in Patients undergoing Hip or Knee Arthroplasties: Results from a High Volume Single Center Multivariate Analysis. Preprints 2023, 2023110915.
- D’Onofrio G., D’Amore A., Onofaro F., Caputi E., Napoli A., Triassi M., Marino M. R. Prediction of Hospital Length Stay for Patients Undergoing Mastectomy. *Stud Health Technol Inform*. 2023 Jun 29;305: 261–264.
- Evans, J., Kobewka, D., Thavorn, K., D’Egidio, G., Rosenberg, E., & Kyeremanteng, K. (2018). The impact of reducing intensive care unit length of stay on hospital costs: evidence from a tertiary care hospital in Canada. *Canadian Journal of Anesthesia*, 65(6), 627–635.
- Gholipour, C., Rahim, F., Fakhree, A., & Ziapour, B. (2015). Using an artificial neural networks (ANNs) model for prediction of intensive care unit (ICU) outcome and length of stay at hospital in traumatic patients. *Journal of clinical and diagnostic research: JCDR*, 9(4), OC19.
- Hauck, K., & Zhao, X. (2011). How dangerous is a day in hospital? A model of adverse events and length of stay for medical inpatients. *Medical care*, 1068–1075.
- Health at a Glance 2023 OECD Indicators <https://www.oecd-ilibrary.org/citeas/10.1787/7a7afb35-en>, last accessed 2024/01/08
- Hesselink, G., Van Den Bogaert, M., Akkermans, R. P., & Schoon, Y. (2019). Risk factors for prolonged length of stay of older patients in an academic emergency department: A retrospective cohort study. *Emergency medicine international*, 2019.
- ISTAT (Istituto Nazionale di Statistica). Annual Italian Statistics 2022 (2022). <http://dati.istat.it/>, last accessed 2024/01/08.
- Italian Ministry of Economy and Finance (2023). Rapporto n. 10: Il Monitoraggio della Spesa Sanitaria. [https://www.rgs.mef.gov.it/\\_Documenti/VERSIONE-I/Attivit-V-i/Spesa-soci/Attivit-monitoraggio-RGS/2023/IMDSS-RS2023.pdf](https://www.rgs.mef.gov.it/_Documenti/VERSIONE-I/Attivit-V-i/Spesa-soci/Attivit-monitoraggio-RGS/2023/IMDSS-RS2023.pdf)
- Jiang, L. Y., Liu, X. C., Nejatian, N. P. et al. (2023). Health system-scale language models are all-purpose prediction engines. *Nature*, 1–6.
- Jo, Y. Y., Han, J., Park, H. W., Jung, H., Lee, J. D., Jung, J., Cha, H. S., Sohn, D. K., Hwangbo, Y. (2021). Prediction of prolonged length of hospital stay after cancer surgery using machine learning on electronic health records: Retrospective cross-sectional study. *JMIR medical informatics*, 9(2), e23147.
- Marfil-Garza, B. A., Belaunzarán-Zamudio, P. F., Gulias-Herrero, A., Zuñiga, A. C., Caro-Vega, Y., Kershenobich-Stalnikowitz, D., & Sifuentes-Osornio, J. (2018). Risk factors associated with prolonged hospital length-of-stay: 18-year retrospective study of hospitalizations in a tertiary healthcare center in Mexico. *PLoS one*, 13(11).

- Mekhaldi, R. N., Caulier, P., Chaabane, S., Chraibi, A., & Piechowiak, S. (2020, April). Using machine learning models to predict the length of stay in a hospital setting. In *World conference on information systems and technologies* (pp. 202–211). Cham: Springer International Publishing.
- Molloy, I. B., Martin, B. I., Moschetti, W. E., & Jevsevar, D. S. (2017). Effects of the length of stay on the cost of total knee and total hip arthroplasty from 2002 to 2013. *The Journal of bone and joint surgery. American volume*, 99(5), 402.
- OECD Indicators on Healthcare Resources: Hospital beds by sector between 2001 and 2021 <https://stats.oecd.org/>
- Olivato, M., Rossetti, N., Gerevini, A. E., Chiari, M., Putelli, L., & Serina, I. (2022). Machine Learning Models for Predicting Short-Long Length of Stay of COVID-19 Patients. *Procedia Computer Science*, 207, 1232–1241.
- Stone, K., Zwigelaar, R., Jones, P., & Mac Parthaláin, N. (2022). A systematic review of the prediction of hospital length of stay: Towards a unified framework. *PLOS Digital Health*, 1(4), e0000017.
- Thompson, B., Elish, K. O., & Steele, R. (2018, December). Machine learning-based prediction of prolonged length of stay in newborns. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1454–1459). IEEE.
- Trunfio TA, Borrelli A, Improta G. Is It Possible to Predict the Length of Stay of Patients Undergoing Hip-Replacement Surgery? *International Journal of Environmental Research and Public Health*. 2022; 19(10):6219.
- Zelege AJ, Palumbo P, Tubertini P, Miglio R and Chiari L (2023) Machine learning-based prediction of hospital prolonged length of stay admission at emergency department: A Gradient Boosting algorithm analysis. *Front. Artif. Intell.* 6:1179226.
- Zhang, D., Yin, C., Zeng, J., Yuan, X., & Zhang, P. (2020). Combining structured and unstructured data for predictive models: a deep learning approach. *BMC medical informatics and decision making*, 20(1), 1–11.

# Systematic Gathering of Requirements for Macroergonomic Analysis and Design for Organizations in Brazil

Elton Moura Nickel, Flávio Anthero Nunes Vianna dos Santos, Patrícia Teixeira Parrela, and Isadora Pedreira de Assis

State University of Santa Catarina - UDESC, Florianópolis, SC 88035901, Brazil

## ABSTRACT

Around the world, several organizations benefit from studies on Macroergonomics. This area of research proposes methods and tools for optimizing socio-technical systems, through the analysis and design of work systems, so that they are suitable for human beings. As a concept, sociotechnical systems refer to interactions between humans and technology, which can be as simple as a single individual using a hand tool or as complex as a multinational organization. In this way, Macroergonomics becomes an important resource for the success of organizational management because it is centred on the human being, that is, it systematically considers professional and psychosocial characteristics in the design or redesign of work systems, thus being a humanized approach in the allocation of functions and tasks. At the same time, it is observed that the main macroergonomic methods currently available in the literature were designed at a specific time and context, in this case, predominantly in the 90s and to mainly meet the demand of North American companies. Therefore, the objective of this paper is to present the process of obtaining requirements that support better performance of macroergonomic analysis and design activities, so that they adequately cover organizations in Brazil. To achieve this objective, the main method used in this research, both for data collection and analysis, was the Systematic Literature Review (SLR), which included the investigation of studies already published in indexed databases about the use of macroergonomic methods in Brazilian organizations. The Systematic Literature Review (SLR) method used was divided into six stages: 1) Definition of the research question and conceptual framework; 2) Search strategy; 3) Search, eligibility and coding; 4) Quality assessment; 5) Summary of results and; 6) Presentation of the study. The result obtained and presented in this paper is a compilation of peer-reviewed scientific studies, which were analysed with the purpose of identifying the main characteristics, as well as the benefits and/or limitations of macroergonomic methods applied in different organizations in the last two decades. The analysis carried out generated conclusions that made it possible to draw up a list of requirements so that macroergonomic analysis and design activities can be conducted appropriately in this specific context. It is expected that these results will be useful to support the construction of a new methodological approach in the context of organizational management that meets the needs of organizations in Brazil more fully.

**Keywords:** Macroergonomics, Organizational management, Methods, Systematic review

## INTRODUCTION

Macroergonomics, or Organizational Ergonomics, makes up, together with Physical Ergonomics and Cognitive Ergonomics, the three major fields of activity in Ergonomics today, as established by the International Ergonomics Association - IEA. This area of knowledge deals with the optimization of socio-technical systems, including their organizational structures, policies and processes (IEA, 2023). As a concept, sociotechnical systems refer to interactions between humans and technology, which can be as simple as a single individual using a hand tool or as complex as a multinational organization. Thus, Macroergonomics values the criteria for an effective approach to the design of work systems, which involve a more integrated project, a humanized approach to the task and consideration of the socio-technical characteristics of organizations. This area of knowledge emerges with a top-down approach to sociotechnical systems for the design of work systems and defines the specifications for the design of the general work system, for the design of human-work, human-machine and human-software. Above all, Macroergonomics is human-centred, because it systematically considers professional and psychosocial characteristics in the design of the work system, being a humanized approach in the allocation of functions and tasks (Hendrick and Kleiner, 2002).

At the same time, it appears that researchers in this area in Brazil currently have a relatively old bibliography, which is predominantly from the 90s, as a reference to methods and tools for use in macroergonomic analysis and design. Furthermore, there is a greater dissemination of foreign case studies that were successful through the implementation of the methodology in the field of Macroergonomics and only incipient studies in Brazil. Added to this, certain organizational tools were designed in commercial software format that aim to assist with specific problems in North American organizations and are not available to the general public or, when they are, they are systems with outdated technology (Hendrick and Kleiner, 2002; Stanton et al., 2005). Therefore, this investigation is based on the assumption that both the teaching field of Macroergonomics in Brazil and national organizations can benefit from the implementation of improvements if an updated methodological foundation is established and adapted to organizational needs and strategies, aiming at analysis and the design of your work systems. Another research confirms this hypothesis. For example, Shahnava, Helali and Emami (2000) talk about the need for more Macroergonomics studies in the context of developing countries and Lawson et al. (2021) explains that Macroergonomics emerged in the West to address the American context and then presents the challenges and solutions for applying it in developing countries.

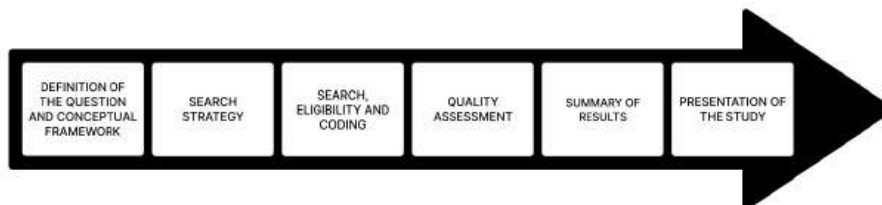
Therefore, given this gap, this paper proposes to present a process for systematically gathering requirements that aims to provide the basis for proposing a new methodological approach in the area of Macroergonomics that meets the needs of organizations in Brazil.



## METHOD

This study adopts bibliographical research as the main technique for data collection and analysis. According to Marconi and Lakatos (2010, p. 166), “bibliographical research is not a mere repetition of what has already been said or written on a certain subject, but rather allows the examination of a topic under a new focus or approach, reaching innovative conclusions”. In the case of this research, the innovative conclusions result in the identification of requirements for the development of a methodological approach that encompasses macroergonomic analysis and design activities in a way that is more appropriate to the reality of organizations in Brazil. This result will be obtained from the systematic study on the use of macroergonomic methods in national organizations that have already been published.

Given the above, it was decided to carry out a Systematic Literature Review (SLR), which is the main method presented in this paper. In fact, Saunders, Lewis and Thornhill (2012) argue that any research project should consider carrying out a systematic review of the available literature as one of its steps. Thus, although there are several methods for carrying out systematic reviews, such as those available in Khan et al. (2003), Cooper, Hedges and Valentine (2009), Smith et al. (2011) and Gough, Oliver and Thomas (2012), this study adopted an adaptation of the integrated method proposed by Dresch, Lacerda and Júnior (2015), whose steps can be seen in Figure 1.



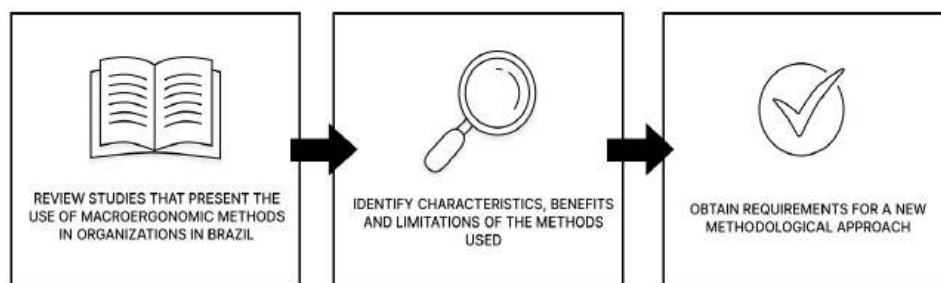
**Figure 1:** Method for systematic literature review (adapted from Dresch, Lacerda and Júnior, 2015, p. 146).

The authors of the method explain that SLR should be used to “map, find, critically evaluate, consolidate and aggregate the results of relevant primary studies on a specific research question or topic” (Dresch, Lacerda and Júnior, 2015, p. 142). The same authors also clarify that the SLR should make it possible to identify gaps to be filled, resulting in a coherent report or a synthesis. Thus, this synthesis must be much more than a mere compilation of the different elements researched, but it is expected that it will result in new knowledge. Therefore, in order to systematize the search for this new knowledge, the following topic presents the results obtained from conducting the steps of the proposed method.

## RESULTS

### Definition of the Question and Conceptual Framework

The first stage in carrying out this systematic review was the definition of the central research theme, which involves explaining the review question and defining the scope of the review through the development of a conceptual framework. Thus, based on the objectives of the present study, the review question can be presented as follows: What are the main characteristics that a macroergonomic analysis and design method should have to address the specificities of current organizations in Brazil? Based on this fundamental question, the conceptual framework was developed, which reveals how it will be answered through research, as shown in Figure 2.



**Figure 2:** Conceptual research framework (developed by the authors).

The three procedures presented in the previously defined conceptual framework make up the structure of a systematic review classified as an aggregative review, as it is a closed question that seeks to test a theory based on the collection of empirical observations using a hypothetical-deductive method, in which the results of primary studies are aggregated to obtain results (Dresch, Lacerda and Júnior, 2015). Thus, once the research question and conceptual framework have been defined, the search strategy will be detailed.

### Search Strategy

The first step in a search strategy is defining the search sources and terms to be used. For this research, according to the conceptual framework, studies that presented the use of macroergonomic methods in organizations in Brazil should be reviewed. Therefore, it was decided to limit the search sources to the main Brazilian databases that provide peer-reviewed scientific research: the Catalog of Theses and Dissertations and the Periodicals Portal of CAPES (Coordination for the Improvement of Higher Education Personnel of the Education Ministry). In these databases, the following search terms were used, in Portuguese: “Macroergonomics” OR “Organizational Ergonomics” AND “Method” OR “Methodology”.

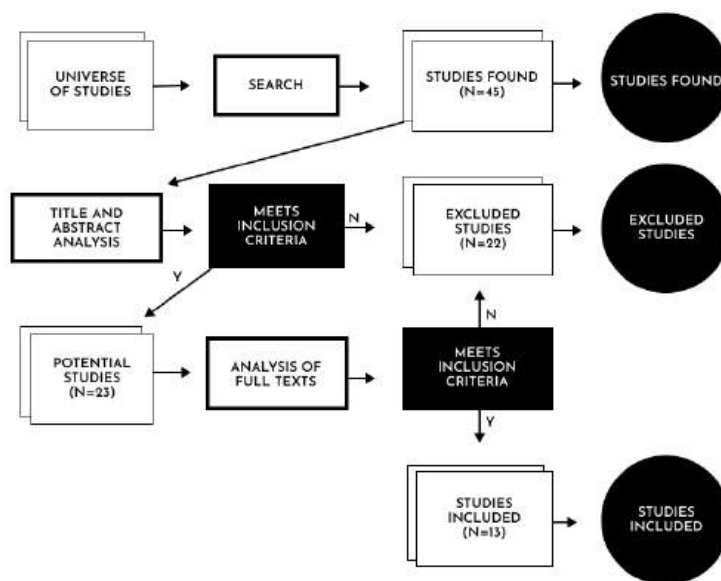
Based on the results found, inclusion criteria were applied, with the aim of verifying whether the analysed work actually presented the use or application of at least one macroergonomic method in an organization in Brazil. This

information was obtained by reading the title and summary of the paper or monograph.

### Search, Eligibility and Coding

After defining the search strategy, we moved on to its operationalization, that is, to search for primary studies, their selection and coding for subsequent evaluation, synthesis and presentation of results. This process is presented in Figure 3.

Based on the research included in this systematic review, a more in-depth analysis of the content was carried out in order to identify the characteristics, benefits, and/or limitations of the methods used in each study.



**Figure 3:** Search, eligibility and coding process (adapted from Dresch, Lacerda and Júnior, 2015, p. 154).

### Quality Assessment

The validity of the results produced from this systematic review also involves assessing the quality and relevance of the research, which must consider the selected primary studies and the review process in a more holistic way. Thus, with the purpose of minimizing study bias and considering all possible contributions from the databases researched, information already validated by the scientific community in this specific field was added. For example, when reading the included studies in full, two important references were used to help define the set of macroergonomic methods to be considered for this research. Initially, the work of Stanton et al. (2005), which presents 15 methods that are part of the section called Macroergonomic Methods, being one of the most complete compilations of methods available and validated

by the literature in the area. Additionally, the paper by Ferreira, Merino and Figueiredo (2017) was used as a reference for research in Brazil, who carried out a review of traditional literature to identify the main methods used in Organizational Ergonomics.

## Summary of Results

Table 1 presents a summary of the results obtained in this Systematic Literature Review (SLR).

**Table 1.** Summary of SLR results (developed by the authors).

Reference and method(s)	Considerations about the method(s)
Simoni and Zerbetto (2010) Macroergonomic Work Analysis (MWA) and Participatory Ergonomics (PE)	Applied in a construction company, the MWA method proved to be efficient in implementing an ergonomics program in the company. However, it was found that, in conjunction with the PE method, ergonomic demands are more easily identifiable compared to simple observation, with better interpretation and organization of the company's diagnosis.
Sampaio and Souza (2012) Interviews and Questionnaires	Applied in a restaurant, the results of the methods allowed us to conclude that the concepts, techniques and tools of macroergonomics can only be applied and implemented to their fullest when they are recognized by the manager as indispensable for the well-being of workers and for improving their performance. Furthermore, there was a need to develop a more detailed questionnaire to better monitor workers.
Campos and Oliveira (2013) Macroergonomic Design (MD)	Applied in a plastic recycling company, the method was defined as efficient for the study objectives as it was participatory in nature. However, despite having seven steps, only the first two steps of the method were carried out. The application of the method and the generalization of the results were limited to just one company. It was suggested to use it in conjunction with other methodologies.
Andrade (2016) Focus Groups	Applied in a public educational institution, the use of the method proved to be adequate for the purposes of the study. Records of satisfaction and dissatisfaction in relation to the organizational environment could be perceived. The limitation of the research was that it did not address motivation in reference to the individual's behaviour or personal satisfaction. The need to carry out studies with other methods for a complete macroergonomic approach was perceived.
Bischoff (2018) Macroergonomic Analysis of Structure (MAS)	Applied to clothing manufacturing companies, the results of the method did not present concrete actions to be taken by the company, but indicated the type of path that can guide decisions in order to obtain positive transformations in its organizational structure. Although MAS was considered flexible and versatile, it was realized that no change is easy in the business context, requiring time to adapt and carry out training and simulations.
Silva (2018) Macroergonomic Analysis of Structure (MAS) and HiTOP Analysis	Applied to Startups, the methods did not prove to be completely suitable for this type of company, making it difficult for the interviewed managers to understand the objectives of the methods. The MAS was considered more suitable for an initial analysis of organizations. The HiTOP Analysis was considered complex and time-consuming to apply, but it helped in the allocation of employees to tasks and training. Both methods were considered complete, but would need to be simplified for greater efficiency.

**Table 2.** Summary of SLR results (developed by the authors).

Reference and method(s)	Considerations about the method(s)
Wojcikiewicz (2018) Participatory Ergonomics (PE) and Macroergonomic Organizational Questionnaire Survey (MOQS)	Applied in a public educational institution, the methods proved to be complete and versatile, allowing not only the collection of data on organizational discrepancies, but also data on satisfaction and performance. The MOQS was considered easy to apply, but generated a large amount of data. With the PE method, it was found that participants' satisfaction increases due to the feeling of responsibility and power to change.
Mejias-Herrera (2018) Participatory Ergonomics (PE)	Applied at a biotechnology institute, it was found that the method must be used by people trained, not only in ergonomics, but in the procedure itself. The benefits of the application were not quickly realized and the participation of all hierarchical levels of the organization was recommended for greater effectiveness.
Goya (2019) Interviews and Questionnaires	Applied to coworking spaces, the methods are not limited to numerical representation, maintaining the focus on deepening the understanding of a group or organization. It was noticed that in-person visits were more efficient, as physical spaces could be better analysed and people felt more comfortable expressing themselves. The study was limited to the results obtained in the interviews and consisted of an initial approach to solving the problems.
Pires (2020) Interviews and Questionnaires	Applied to urban space, macroergonomic methods needed to be combined with other methodological approaches to obtain satisfactory results in this specific context. Although the studied city environment was considered as an organization, the study proved to be complex to be approached with traditional macroergonomics tools.
Nascimento (2021) Macroergonomic Analysis of Structure (MAS)	Applied in marketing companies, the method was considered flexible and subjective. However, it was found that, without the support of other studies to develop the research materials and methods, it would not be possible to obtain the necessary data. There were also doubts about the number of workers who should be interviewed, as it was noticed that the greater the number of workers interviewed, the greater the detail of the information obtained.
Oliveira (2021) Interviews and Questionnaires	Applied to nursing teams, the methods used were limited to specific instruments in the health area and did not support a complete macroergonomic analysis. On the other hand, the research results were statistically significant and can be generalized, although working conditions are peculiar in each occupational environment.
Girardi (2022) Macroergonomic Analysis of Structure (MAS)	Applied in a health rehabilitation centre, the method helped to identify the interaction between the specific structure of the organization and the ideal way of working. However, it was found that the assessment of sociotechnical variables and the degree of importance attributed to them is a subjective process that requires training and experience in conducting organizational assessments. It was also noted that the time required to apply the MAS varies depending on the scope of the organization.

As shown in Table 1, from the thorough analysis of each study, it was possible to identify the characteristics, benefits, and/or limitations of the macroergonomic methods used – this information is included in the column “considerations about the method(s)”. Thus, from the compilation and analysis of this information, it is possible to generate requirements for a methodological approach that is more suitable for organizations in Brazil.

## Presentation of the Study

All the methods presented have basically the same objective: to help an organization carry out, even partially, an analysis and a possible macroergonomic design or redesign. With this objective, the main result of this paper is now presented, that is, a list of requirements for the elaboration of a new methodological approach to analysis and macroergonomic design for organizations in Brazil, based on the studies carried out:

- Use Participatory Ergonomics (PE) as the main approach to macroergonomic analysis and design;
- People's participation must occur at all hierarchical levels;
- The same methodological approach must analyse both the processes and the organizational structure;
- Provide for the participation of senior management as a requirement for implementing a macroergonomic program;
- In addition to presenting the stages, the tasks required in each stage must be detailed and support tools suggested to carry them out;
- Indicate which stakeholders can participate in each application stage;
- The best practices of the main macroergonomic methods available should be adopted, but in a more intuitive and less complex way;
- Indicate the application step by step, without depending on the analyst's level of experience;
- Although the results monitoring process is constant and based on continuous improvement, the application of the methodological approach must be simple and quick;
- A new macroergonomic method must be flexible enough to adapt to different types of organizations.

## CONCLUSION

This paper presented the description of a process for obtaining requirements to support better performance of macroergonomic analysis and design activities that are adapted to organizations in Brazil. To achieve this objective, a Systematic Literature Review (SLR) was carried out, which included the investigation of studies already published in indexed databases on the use of macroergonomic methods in Brazilian organizations, as well as the presentation of considerations about these methods.

For future work, we intend to create a new methodological approach based on the requirements obtained, thus meeting the existing demand in the research context. Therefore, faced with such a gap, what this research proposed to develop were the methodological foundations that support a complete macroergonomic approach, in order to contemplate in a cohesive and easy-to-apply manner the activities of macroergonomic analysis and design, both of structures and of organizational processes. To achieve this, the proposal was limited to the requirements and management strategies of organizations in the contemporary Brazilian scenario.

## ACKNOWLEDGMENT

The authors would like to acknowledge FAPESC (2023TR632) for the financial support.

## REFERENCES

- Andrade, M. F. P. (2016). A condição do trabalho na relação entre satisfação e frustração em ambiente público de ensino superior sob olhar da ergonomia organizacional. 2016. 125 f. Mestrado Profissional em Ergonomia. Universidade Federal de Pernambuco.
- Bischoff, B. M. (2018). Macroergonomia na cadeia da moda: diagnóstico de comunicação em empresas com fabricação descentralizada. 2018. 355 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.
- Campos, G. C. and Oliveira, P. G. (2013). Macroergonomia aplicada em postos de trabalho de uma empresa de reciclagem mecânica de plástico. GEPROS - Gestão da Produção, Operações e Sistemas, Bauru, Ano 8, abr-jun/2013, no. 2, pp. 83–94.
- Cooper, H. M.; Hedges, L. V. and Valentine, J. C. (2009). The handbook of research synthesis and meta-analysis. London: Sage.
- Dresch, A.; Lacerda, D. P. and Antunes Júnior, J. A. V. (2015). Design science research: método de pesquisa para avanço da ciência e tecnologia. Porto Alegre: Bookman.
- Ferreira, A. S.; Merino, E. A. D. and Figueiredo, L, F, G. (2017). Methods used in organizational ergonomics: literature review. Human Factors in Design, vol. 6, no. 12, pp. 58–78.
- Girardi, E. O. L. (2022). Ergonomia organizacional para prevenção do abandono de tecnologia assistiva em centros especializados em reabilitação. 2022. 100 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.
- Goya, J. Y. L. (2019). A importância dos espaços colaborativos universitários para a formação do designer. 2019. 117 f. Mestrado em Design. Universidade Estadual Paulista Júlio de Mesquita Filho (Bauru).
- Gough, D.; Oliver, S. and Thomas, J. (2012). An introduction to systematic reviews. London: Sage.
- Hendrick, H. W. and Kleiner, B. M. (2002). Macroergonomics: Theory, methods, and applications. New Jersey: Lawrence Erlbaum Associates.
- IEA. International Ergonomics Association. (2023). Available at: <<https://iea.cc/>>. Accessed on: 10 Dec. 2023.
- Khan, K. S. et al. (2003). Five steps to conducting a systematic review. Journal of the Royal Society of Medicine, vol. 96, no. 3, pp. 118–121.
- Lawson, G. et al. (2021). Participatory ergonomics in industrially developing countries: A literature review. International Journal of Mechanical Engineering Technologies and Applications, vol. 2, no. 1, pp. 53–59.
- Marconi, M. A. and Lakatos, E. M. (2010). Fundamentos de metodologia científica. 7. Ed. – São Paulo: Atlas.
- Mejias-Herrera, S. H. (2018). Las herramientas de intervención ergonómica: consideraciones conceptuales y experiencias prácticas en Cuba y Brazil. Ingeniería Industrial, vol. 39, no. 1, pp. 3–14.
- Nascimento, M. F. (2021). Ergonomia organizacional no home-office: A análise macroergonômica na comunicação interna de equipes de marketing. 2021. 225 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.

- Oliveira, B. R. (2021). Ergonomia organizacional e saúde: condições de trabalho e presenteísmo em enfermagem hospitalar. 2021. 130 f. Mestrado em Enfermagem. Universidade do Estado do Rio de Janeiro, Rio de Janeiro.
- Pires, G. D. (2020). Abordagem da macroergonomia no espaço urbano: fatores que influenciam na percepção de qualidade. 2020. 119 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.
- Sampaio, C. P. and Souza, J. (2012). A Macroergonomia na melhoria das condições de trabalho com ênfase nos aspectos de liderança: estudo de caso com AMT em um restaurante. *Projetica*, vol. 3, no. 1, pp. 79–95.
- Saunders, M.; Lewis, P. and Thornhill, A. (2012). *Research Methods for business students*. 6. Ed. London: Pearson Education.
- Shahnavaz, H.; Helali, F.; Emami, K. H. (2000). Macroergonomics in Industrially Developing Countries-Case Study Iran. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Sage CA: Los Angeles, CA: SAGE Publications.
- Silva, L. D. (2018). Um estudo sobre a viabilidade da abordagem macroergonômica na análise organizacional de startups em Florianópolis-SC. 2018. 143 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.
- Simoni, C. G. and Zerbetto, C. A. A. (2010). Análise macroergonômica em uma empresa da construção civil. *Projetica*, vol. 1, no. 1, pp. 47–67.
- Smith, V. et al. (2011). Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology*, vol. 11, no. 1, p. 15.
- Stanton, N. et al. (2005). *Handbook of human factors and ergonomics methods*. Boca Raton, CRC Press LLC.
- Wojcikiewicz, F. A. (2018). Ergonomia participativa: uma análise macroergonômica para o desenvolvimento da arquitetura organizacional de laboratórios de ensino. 2018. 121 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Design, Universidade do Estado de Santa Catarina, Florianópolis.