Digital Human-Centered Augmentation Technologies

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt von

David Steeven Villa Salazar

M.Sc. Computer Science

München, den 28. Februar 2025

Erstgutachter:	Prof. Dr. Albrecht Schmidt
Zweitgutachter:	Prof. Dr. Suranga Nanayakkara
Drittgutachter:	Prof. Dr. Hideki Koike

Tag der mündlichen Prüfung: 31.07.2025

"The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today."

– J. C. R. Licklider

Abstract

Our senses, cognition, and physical capabilities have long been bound by biological constraints. Not anymore. Augmentation technologies, from wearable haptics to AI-driven cognitive support are expanding human potential in ways once imagined only in science fiction. But with these new capabilities come new questions: *How does augmentation affect decision-making and risk-taking? How do others perceive augmented humans?* This thesis addresses these points by not only studying augmented human behavior and perceptions but also developing functional prototypes to explore augmentation in real-world applications.

Here I explore the design of human augmentation technologies through the dual lenses of autonomy and heteronomy, investigating the interplay between internal and external influences on the augmented human and its decision-making processes.

For autonomy, I examine how augmentation technologies impact individual decision-making and behavior. Specifically, I focus on risk-taking and the sense of agency, demonstrating that the mere presence or even the belief in the presence of augmentation technologies significantly affects human behavior and the attribution of responsibility at both behavioral and physiological levels.

For heteronomy, I explore how external societal factors influence the decision-making of augmented individuals. By examining societal attitudes toward augmented humans across different cultural contexts, this work identifies key factors shaping public perceptions of augmentation technologies and presents a standardized methodology for measuring these attitudes over time. This includes the development and validation of the SHAPE scale, a psychometric tool designed to assess societal acceptance of human augmentation.

Finally, I investigate how individuals experience human augmentations, examining the boundaries between augmentation and learning. In this context, I further explore the potential of Virtual Reality as a platform to prototype and study augmented human experiences, providing insights into how these technologies can be integrated meaningfully into human life.

Beyond these behavioral and societal perspectives, this thesis also contributes to the technical foundations of human augmentation. A central part of this work involves the development of functional augmentation prototypes. These include wearable haptic feedback systems for sensory, cognitive and motor augmentation and interactive VR environments to study augmentation experiences in controlled settings.

The contributions of this thesis span from understanding individual behaviors to societal perceptions and technical artifacts in human augmentation. As augmentation technologies become increasingly plausible, this work provides researchers and practitioners with a human-centered framework for designing augmentation systems that harmoniously integrate into human life.

Zusammenfassung

Unsere Sinne, unsere Kognition und unsere körperlichen Fähigkeiten waren lange Zeit durch biologische Grenzen definiert. Doch das ändert sich. Augmentierungstechnologien – von tragbarer haptischer Rückkopplung bis hin zu KI-gestützter kognitiver Unterstützung – erweitern menschliche Fähigkeiten auf eine Weise, die einst nur in der Science-Fiction vorstellbar war. Doch mit diesen neuen Möglichkeiten entstehen auch neue Fragen: Wie beeinflusst Augmentierung Entscheidungsprozesse und Risikoverhalten? Wie werden augmentierte Menschen von der Gesellschaft wahrgenommen? Diese Arbeit untersucht diese Aspekte nicht nur durch empirische Studien zum Verhalten und zu den Wahrnehmungen augmentierter Menschen, sondern auch durch die Entwicklung funktionaler Prototypen, die Augmentierung in realen Anwendungsszenarien erfahrbar machen.

Diese Dissertation erforscht die Gestaltung von Augmentierungstechnologien aus zwei Perspektiven: Autonomie und Heteronomie. Dabei wird untersucht, wie interne und externe Faktoren die Entscheidungsfindung von augmentierten Individuen beeinflussen.

Im Bereich der Autonomie analysiere ich, wie Augmentierungstechnologien individuelle Entscheidungsprozesse und Verhaltensweisen verändern. Besonders im Fokus stehen Risikobereitschaft und das Gefühl der Handlungskontrolle (Agency). Die Ergebnisse zeigen, dass bereits die bloße Anwesenheit – oder sogar nur der Glaube an die Anwesenheit – von Augmentierungstechnologien das Verhalten und die Zuschreibung von Verantwortung sowohl auf kognitiver als auch auf physiologischer Ebene signifikant beeinflusst.

Im Bereich der Heteronomie untersuche ich, wie gesellschaftliche Faktoren die Entscheidungsprozesse von augmentierten Individuen formen. Durch die Analyse sozialer Einstellungen gegenüber augmentierten Menschen in verschiedenen kulturellen Kontexten identifiziere ich zentrale Faktoren, die öffentliche Wahrnehmungen von Augmentierungstechnologien prägen. Dies beinhaltet auch die Entwicklung und Validierung der SHAPE-Skala, eines psychometrischen Instruments zur Messung der gesellschaftlichen Akzeptanz von Augmentierungstechnologien über die Zeit.

Darüber hinaus untersuchen wir, wie Individuen Augmentierung erleben und wo die Grenzen zwischen Augmentierung und Lernen verlaufen. In diesem Zusammenhang wird insbesondere das Potenzial von Virtual Reality als Plattform zur Prototypisierung und Erforschung augmentierter Erfahrungen erforscht, um zu verstehen, wie diese Technologien sinnvoll in das menschliche Leben integriert werden können.

Neben diesen verhaltenswissenschaftlichen und gesellschaftlichen Perspektiven leistet diese Arbeit auch technologische Beiträge zur Augmentierungsforschung. Ein zentraler Bestandteil ist die Entwicklung funktionaler Augmentierungsprototypen. Dazu gehören tragbare haptische Feedback-Systeme für sensorische, kognitive und motorische Augmentierung sowie interaktive Virtual-Reality-Umgebungen (VR), in denen Augmentierungserfahrungen unter kontrollierten Bedingungen untersucht werden. Durch den Einsatz von VR als Forschungs-

plattform können digitale Simulationen als Testumgebungen für die Entwicklung zukünftiger Augmentierungstechnologien genutzt werden.

Die Erkenntnisse dieser Dissertation reichen von der Analyse individuellen Verhaltens über gesellschaftliche Wahrnehmungen bis hin zur Entwicklung technischer Artefakte in der Augmentierungsforschung. Während Augmentierungstechnologien zunehmend realisierbar werden, bietet diese Arbeit Forschenden und Praktiker:innen ein menschenzentriertes Rahmenwerk zur Gestaltung von Augmentierungssystemen, die sich nahtlos in das menschliche Leben integrieren. Indem sie es Individuen ermöglichen, ihre natürlichen Fähigkeiten zu erweitern, während gleichzeitig eine verantwortungsbewusste Innovation sichergestellt wird, ebnet diese Forschung den Weg für eine Zukunft, in der Technologie das Menschsein bereichert, anstatt es zu ersetzen.

TABLE OF CONTENTS

List Publications During Enrollment

A significant portion of the research presented in this thesis has been peer-reviewed, published, and shared with the broader Human-Computer Interaction (HCI) community. These works have also been presented at various academic venues throughout the course of this thesis. The list below includes all related publications, including those not included in the thesis. Publications that are part of this thesis are marked with specific prefixes (AUT, HET, EXP). A comprehensive record of all publications, projects, and research progress can also be accessed online at PostHCI.com.

Journal Papers

[Aut1] **Villa, Steeven**, Kosch, Thomas, Grelka, Felix, Schmidt, Albrecht, and Welsch, Robin. 'The placebo effect of human augmentation: Anticipating cognitive augmentation increases risk-taking behavior.' In: *Computers in Human Behavior* 146 [2023], p. 107787. DOI: 10.1016/j.chb.2023.107787

[Het2] **Villa, Steeven**, Niess, Jasmin, Schmidt, Albrecht, and Welsch, Robin. 'Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale.' English. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7.3 [2023]. Publisher: ACM, pp. 1–23. DOI: 10.1145/3610915

[Exp5] **Villa, Steeven**, Mayer, Sven, Hartcher-O'Brien, Jess, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Extended Mid-air Ultrasound Haptics for Virtual Reality.' In: *Proceedings of the ACM on Human-Computer Interaction* 6.ISS [2022], 578:500–578:524. DOI: 10.1145/3567731

[Exp4] **Villa, Steeven**, Weiss, Yannick, Hirsch, Niklas, and Wiethoff, Alexander. 'An Examination of Ultrasound Mid-air Haptics for Enhanced Material and Temperature Perception in Virtual Environments.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI [2024]. DOI: 10.1145/3676488

[1] Chiossi, Francesco, Kosch, Thomas, Menghini, Luca, **Villa, Steeven**, and Mayer, Sven. 'SensCon: Embedding Physiological Sensing into Virtual Reality Controllers.' In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI [2023], 223:1–223:32. DOI: 10.1145/3604270

[2] Chiossi, Francesco, Welsch, Robin, **Villa, Steeven**, Chuang, Lewis, and Mayer, Sven. 'Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience.' en. In: *Big Data and Cognitive Computing* 6.2 [2022]. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 55. DOI: 10.3390/bdcc6020055

[3] Schött, Svenja Y., Capsi-Morales, Patricia, **Villa, Steeven**, Butz, Andreas, and Piazza, Cristina. 'Would You Hold My Hand? Exploring External Observers' Perception of Artificial Hands.' en. In: *Multimodal Technologies and Interaction* 7.7 [2023]. Number:

7 Publisher: Multidisciplinary Digital Publishing Institute, p. 71. DOI: 10.3390/mti7070071

[4] Weiss, Yannick, **Villa, Steeven**, Grootjen, Jesse W, Hoppe, Matthias, Kale, Yasin, and Müller, Florian. 'Exploring Redirection and Shifting Techniques to Mask Hand Movements from Shoulder-Surfing Attacks during PIN Authentication in Virtual Reality.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI [2024], 257:1–257:24. DOI: 10.1145/3676502

Full Papers

[Aut2] **Villa, Steeven**, Barth, Lisa, Chiossi, Francesco, Welsch, Robin, and Kosch, Thomas. 'Whose Mind is it Anyway? A Systematic Review and Exploration on Agency in Cognitive Augmentation.' In: *Computers in Human Behavior: Artificial Humans* [2025]

[Aut3] **Villa, Steeven**, Weiss, Yannick, Lu, Mei Yi, Ziarko, Moritz, Schmidt, Albrecht, and Niess, Jasmin. 'Envisioning Futures: How the Modality of AI Recommendations Impacts Conversation Flow in AR-enhanced Dialogue.' In: *Proceedings of the 26th International Conference on Multimodal Interaction*. ICMI '24. Association for Computing Machinery, 2024, 182–193. DOI: 10.1145/3678957.3685731

[Het1] **Villa, Steeven**, Niess, Jasmin, Nakao, Takuro, Lazar, Jonathan, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Understanding Perception of Human Augmentation: A Mixed-Method Study.' en. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 2023, pp. 1–16. DOI: 10.1145/3544548.3581485

[Exp1] **Villa, Steeven**, Krammer, Finn Jacob Eliyah, Weiss, Yannick, Welsch, Robin, and Kosch, Thomas. 'Understanding the Influence of Electrical Muscle Stimulation on Motor Learning: Enhancing Motor Learning or Disrupting Natural Progression?' In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan* [2025]

[Exp2] **Villa, Steeven**, Niess, Jasmin, Eska, Bettina, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Assisting Motor Skill Transfer for Dance StudentsUsing Wearable Feedback.' en. In: *Proceedings of the 2021 ACM International Symposium on Wearable Computers*. ISWC '21. Association for Computing Machinery, 2021, pp. 38–42. DOI: 10.1145/3460421.3478817

[Exp6] **Villa, Steeven**, Ishihara, Kenji, Ziarko, Moritz, Günther, Sebastian, and Müller, Florian. 'Touch It Like It's Hot: A Thermal Feedback Enabled Encountered-type Haptic Display for Virtual Reality.' In: *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2024, pp. 700–709. DOI: 10.1109/ISMAR62088.2024.00085

[Exp7] **Villa, Steeven** and Mayer, Sven. 'Cobity: A Plug-And-Play Toolbox to Deliver Haptics in Virtual Reality.' In: *Proceedings of Mensch und Computer 2022*. MuC '22. ACM, 2022, pp. 78–84. DOI: 10.1145/3543758.3543775

[Exp8] **Villa, Steeven**, Nguyen, Thuy Duong, Tag, Benjamin, Machulla, Tonja-Katrin, Schmidt, Albrecht, and Niess, Jasmin. 'Towards a Haptic Taxonomy of Emotions: Exploring Vibrotactile Stimulation in the Dorsal Region.' In: *Proceedings of the 2023 ACM International Symposium on Wearable Computers*. ISWC '23. Association for Computing Machinery, 2023, pp. 50–54. DOI: 10.1145/3594738.3611363

[5] Shen, Chenxinran, Karpus, Jurgis, Kosch, Thomas, Fernandes, Daniela, Mello, Beatriz, Welsch, Robin, and **Villa, Steeven**. 'The Impact of Asymmetric AI Assistance on Decision-Making in Social Dilemmas: A Study on Human Augmentation in Economic Games.' In: *Augmented Humans 2025*. AHs 2025. Association for Computing Machinery, 2025

[6] Leusmann, Jan, **Villa, Steeven**, Berberoglu, Burak, Wang, Chao, and Mayer, Sven. 'Developing and Validating the Perceived System Curiosity Scale (PSC): Measuring Users' Perceived Curiosity of Systems.' In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan* [2025]

[7] Leusmann, Jan, **Villa, Steeven**, Liang, Thomas, Wang, Chao, Schmidt, Albrecht, and Mayer, Sven. 'An Approach to Elicit Human-Understandable Robot Expressions to Support Human-Robot Interaction.' In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan* [2025]

[8] Chiossi, Francesco, **Villa, Steeven**, Hauser, Melanie, Welsch, Robin, and Chuang, Lewis. 'Design of On-body Tactile Displays to Enhance Situation Awareness in Automated Vehicles.' In: *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. ISSN: 2377-9322. 2022, pp. 1–6. DOI: 10.1109/CIVEMSA53371.2022.9853690

[9] Chiossi, Francesco, Welsch, Robin, **Villa, Steeven**, Chuang, Lewis L., and Mayer, Sven. 'Designing a Physiological Loop for the Adaptation of Virtual Human Characters in a Social VR Scenario.' In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2022, pp. 578–579. DOI: 10.1109/VRW55335.2022.00140

[10] Dietz, Dennis, **Villa, Steeven**, Ziarko, Moritz, Bonfert, Michael, Müller, Florian, and Butz, Andreas. 'Embracer: A Wearable Encountered-Type Haptic Controller for 3 DoF Input and Feedback.' In: *Proceedings of the 2024 ACM International Symposium on Wearable Computers*. ISWC '24. Association for Computing Machinery, 2024, pp. 140–143. DOI: 10.1145/3675095.3676626

[11] Faltaous, Sarah, Hubert, Anna, Karolus, Jakob, **Villa, Steeven**, Kosch, Thomas, and Wozniak, Pawel W. 'EMStriker: Potentials of Enhancing the Training Process of Racket-based

Sports via Electrical Muscle Stimulation.' In: *Proceedings of the Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction*. TEI '22. Association for Computing Machinery, 2022, pp. 1–6. DOI: 10.1145/3490149.3505578

[12] Kloft, Agnes Mercedes, Welsch, Robin, Kosch, Thomas, and **Villa, Steeven**. '"AI enhances our performance, I have no doubt this one will do the same": The Placebo effect is robust to negative descriptions of AI.' in: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '24. Association for Computing Machinery, 2024, pp. 1–24. DOI: 10.1145/3613904.3642633

[13] Pangratz, Elisabeth, Chiossi, Francesco, **Villa, Steeven**, Gramann, Klaus, and Gehrke, Lukas. 'Towards an Implicit Metric of Sensory-Motor Accuracy: Brain Responses to Auditory Prediction Errors in Pianists.' In: *Proceedings of the 15th Conference on Creativity and Cognition*. C&C '23. Association for Computing Machinery, 2023, pp. 129–138. DOI: 10.1145/3591196.3593340

[14] Tannert, Isabel, Schulleri, Katrin H., Michel, Youssef, **Villa, Steeven**, Johannsen, Leif, Hermsdörfer, Joachim, and Lee, Dongheui. 'Immediate Effects of Vibrotactile Biofeedback Instructions on Human Postural Control.' In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. ISSN: 2694-0604. 2021, pp. 7426–7432. DOI: 10.1109/EMBC46164.2021.9630992

[15] Weiss, Yannick, **Villa, Steeven**, Schmidt, Albrecht, Mayer, Sven, and Müller, Florian. 'Using Pseudo-Stiffness to Enrich the Haptic Experience in Virtual Reality.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, 2023, pp. 1–15. DOI: 10.1145/3544548.3581223

Pre-prints

[Aut4] **Villa, Steeven**, Kosch, Thomas, Kloft, Agnes Mercedes, Quinn, Jasper, Kujala, Sari, and Welsch, Robin. 'Inventory of User Expectations for Technology (iExpect).' In: *Computers in Human Behavior Reports* [2025]

[16] Fernandes, Daniela, **Villa, Steeven**, Nicholls, Salla, Haavisto, Otso, Buschek, Daniel, Schmidt, Albrecht, Kosch, Thomas, Shen, Chenxinran, and Welsch, Robin. AI Makes You Smarter, But None The Wiser: The Disconnect Between Performance and Metacognition. arXiv:2409.16708 [cs]. 2024. DOI: 10.48550/arXiv.2409.16708

Late-Breaking Works

[Exp3] **Villa, Steeven**, Neuhaus, Robin, Weiss, Yannick, and Hassenzahl, Marc. 'Exploring Virtual Reality as a Platform for Early-Stage Design for Human Augmentation Technologies.' In: *Proceedings of the International Conference on Mobile and Ubiquitous*

Multimedia. MUM '24. Association for Computing Machinery, 2024, 464–466. DOI: 10.1145/3701571.3703386

[17] schraefel, mc, Andrés, Josh, Tabor, Aaron, Bateman, Scott, Wanyu Liu, Abby, Jones, Mike, Kunze, Kai, Murnane, Elizabeth, and **Villa, Steeven**. 'Body As Starting Point 4: Inbodied Interaction Design for Health Ownership.' In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Association for Computing Machinery, 2021, pp. 1–5. DOI: 10.1145/3411763.3441335

[18] **Villa, Steeven**, Welsch, Robin, Denisova, Alena, and Kosch, Thomas. 'Evaluating Interactive AI: Understanding and Controlling Placebo Effects in Human-AI Interaction.' In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. CHI EA '24. Association for Computing Machinery, 2024, pp. 1–4. DOI: 10.1145/3613905.3636304

Acknowledgements

Although for the reader this is still the beginning, to me this is the conclusion, years of work distilled into these final words. And, more than a simple acknowledgment of gratitude, this is a testament to the countless individuals, joint efforts, and sustained motivation required to bring this work to life. This thesis about human augmentation, the lines you will read following this page, would not have been possible without support of my family, partner, friends, colleagues, and mentors.

I must first and foremost thank **Albrecht Schmidt** for making this journey possible. Your publications and talks inspired my application to this lab and my decision to research *human augmentation*. I am thankful for your ability to spark new ideas in every conversation and for always showcasing your vision of the future. On a personal level, I am deeply grateful for your undertaking the difficult task of leveling the playing field for me as a non-European in a challenging ecosystem. I remain impressed by your patience, energy, and passion for supporting everyone around you."

Second, I must thank **Jasmin Niess** for mentoring me during most of my PhD years, helping me become a better researcher, a better scientist, and a better person. Your mentorship has shaped both this work and who I am in ways I cannot fully express. Thank you for being not just a mentor but a friend who has always guided me and helped me overcome the biggest challenges in this journey. Our meetings and long conversations about philosophy, politics, ethics, and history rank high among the memories I will cherish from this stage of my life.

I would also like to thank **Tonja Machulla** and **Lewis Chuang** for opening the doors of this lab to me and mentoring me in the early stages of my PhD. I am incredibly fortunate that you both saw potential in me and made this opportunity possible. For that, I am grateful.

To my committee, **Prof. Suranga Nanayakkara** and **Prof. Hideki Koike**, I am honored that such distinguished experts reviewed my thesis. Thank you for taking the time, and thank you for contributing actively to the steering of the augmented humans community.

To my mentors and friends, especially **Sven Mayer**: thank you for teaching me a proactive mindset and for motivating me to participate in and better understand our HCI community and conference organization processes. Thank you for making sure that so many of the gears we depend on work properly and improve every day. Thank you for being rigorous and for mentoring junior researchers to be rigorous. Thank you also for being a good friend, for the beers, the trips, the long, highly technical, highly argumentative discussions that could last for hours. Thank you for making me feel comfortable in this lab and country. To **Robin Welsch**, with whom I have explored some of the most exciting research ideas in this thesis: thank you for all the knowledge you have given me. I find it fascinating how every complex concept looks almost trivial after you explain it. Thank you for all the conversations about science, HCI methods, and psychology, and thank you for hosting me and making me part of your lab project. I'm grateful to have someone like you around, giving honest advice and caring about personal growth. Thank you, and thank you, Lynn, for all the nice memories.

To **Thomas Kosch**, you were the first person from the lab I met when I first arrived in this country, just before starting my PhD. Since that day I knew I was with the right people. Since then, we have enjoyed concerts together, drinks until sunrise at Shakira bar, travels, running in the morning in Hawaii, and countless crazy paper ideas.

Thank you to **Francesco Chiossi** for making me feel like this country can feel like home, for all the trips, dinners, drinks, and memories together. I'm lucky we started this journey together. Thank you also to **Luke Haliburton**, your presence always makes the atmosphere feel lighter. Thank you for the nice memories in Mexico and Australia, and for making one of my first long hikes one of the most dangerous ones in the world. Thank you to **Matthias Hoppe** for the never-ending conversations, for helping me tune my ability to detect puns, fly drones, and come up with crazy ideas. Having you at my defense made me feel as motivated as in the early years of my PhD.

To **Carl Oechsner**, for always being understanding and supportive, for being a good friend, complice in adventures, welcoming, honest, caring, and for hosting the absolute best flat parties in Munich. To **Changkun Ou**, for the deep and challenging philosophical questions that have sparked great conversations. To **Matthias Schmidtmeier**, for always being spontaneous, natural, easygoing, and reasonable. Thank you for founding the [redacted] connection and keeping us connected.

To **Yannick Weiss**, for striking the perfect balance between argumentation and reason. I especially enjoy our discussions about methods, our rhetorical debates, and how effortlessly we arrive at the most exciting research ideas together. At some point you almost made me believe that haptic illusions were real haptics. Good try. P.S.: You still owe us a mullet. To **Jan Leusmann**, man, you're next to Yannick because you're both almost always in my office, which I love, but also because discussing with you is always an adventure. No weak argument can survive your sharp thinking. Besides this, you're a great human being. It really has been wonderful to share this time with you both.

To **Jesse Grootjen**, I love that one of us actually got a motorcycle in Germany. Thank you for being there since the beginning. To **Florian Lang**, for the preemptive welcome to Francesco and me since the day of the interview, since then, talking to you is calming or informative, but rarely both. To **Julian Rasch**, for the plants you gifted me. I have many good memories with you, but I'll especially remember our long conversations with the Yokohama sunrise in front of us. To **Maximiliane Windl**, for your energy, for being supportive and inclusive, we are lucky to have you in the lab. To **Henrike Weingartner,** for insisting on inviting me to lunch, for eventually desisting, and then trying again from time to time. Time with you is always great, especially when you passionately narrate your stories. To **Clara Sayffaerth**, for being a calm, understanding person amidst all the craziness around us. To **Teodora Mitrevska**, for bringing happiness back to the lab with your contagious extroversion and drama.

To **Thomas Weber** and **Florian Bemmann**, such wonderful people to share an office with during this journey. Thank you both. Coming to the office always felt like entering a safe, calm, and relaxed space, full of plants and Mochi. Very often there were Haribos and Capybaras

from Thomas, and burritos and great coffee from Florian. You both set the bar really high.

Thank you to **Alexander Wiethoff** for the wonderful research and design collaborations. To **Florian Alt**, for always having a smiling face and being welcoming, for the great trip to Brazil, and for recently welcoming me to your group. To **Andreas Butz**, for the pragmatism, calmness, and willingness to always help. To **Johanna Pirker**, for the cool project collaboration and for introducing me to games research. To **Heinrich Hussmann**, for founding this group at LMU and creating the foundation that brought all these people together.

To **Agnes Kloft**, **Nils Meyer-Kahlen**, and **Daniela Fernandes**. You made my time in Finland awesome. Agnes, I'm grateful for the intense Spanglish talks, wonderful collaborations, and warmth. Nils, for showing me Helsinki's coolest bars and the Meta Redmond cafeteria with the best free ice cream I've had. Daniela, for the excellent work, great interactions, and and for hating on my Brazilian Portuguese accent.

To **Rainer Fink**, for always being helpful and supportive, for keeping all the machinery working, I don't know where we would be without you. To **Anja Mebus**, for supporting me in the most important moments, for battling with the migration office to secure me an appointment, and for helping me navigate German bureaucracy, which grows more complex each time. To **Christa Feulner**, for always being ready to help in your calm way. To **Franziska Schwamb**, for always greeting me with warmth and positive energy, and for constantly checking up on my German skills. To **Salsabeel Alkhlyef**, for having dealt with my sometimes, incomplete paperwork and always being friendly.

To all the great students I had throughout this journey, who strongly contributed to the realization of this research program: you have shaped my thinking and approach to both research and practice. I especially want to thank **Moritz Ziarko** for all the technical discussions, future-looking ideas and solutions, and for always being at the forefront of consumer tech—this was a productive journey for us. **Finn Kramer**, for your continuous motivation and curiosity. **Lisa Barth**, for your outstanding analytical skills. **Burak Berberoglu**, for your capability to synthesize broad concepts into actionable tasks. S**arita Sridharan**, for your capability to conduct independent, high-quality research. **Marlene Eder**, for not being afraid to take on challenging projects. **Niklas Hirsch**, for your outstanding productivity. **Kenji Ishihara**, for your capability to discuss, define, and implement projects with minimal instruction. **Carmen Schanderl**, for the entire journey in our scale project, where you also immersed yourself in my Colombian culture. **Christofer Quirin Parstorfer**, for the great discussions and jokes, for exploring the Deutsches Museum with me, and for sharing more than just research interests. To **Nicholas Ornstein**, beyond research, our discussions about Žižek, Fernando González, and other philosophers made hosting you truly enjoyable.

On a similar note, to **Oliver Hein**, for making simple conversations take on a higher level of thinking. I especially remember the time in Abu Dhabi, talking about anarchy, Žižek's concept of ideology, and more, while we played billiards in an open plaza in the center of the city. To **Heiko Drewes**, for thinking differently and asking the provocative questions not everyone makes, or even knows are there.

To **Sarah Voelkel**, **Nađa Terzimehić**, **Fiona Draxler**, and **Linda Hirsch**, for the wonderful conversations and guidance throughout these years. Sarah, thank you also for the invaluable professional coaching. Nađa, for offering your perspective on the immigration experience in Germany. Fiona, for your consistently clear communication. Linda, for your steady support.

To **Bettina Eska**, **Dennis Dietz**, and **Jakob Karolus**: I'm happy I had the chance to meet you all. Bettina, as one of my first students and later a colleague. Dennis, as an early collaborator. Jakob, as an early mentor who introduced me to student supervision in this lab."

To **Rifat Amin**, **Markus Wieland**, **Pavithren "Viren" Pakianathan**, and **Philipp Thalhammer**, for making the lab environment wonderful. Rifat, I still remember the incredible chicken biryani you shared with us. Markus, I am so grateful our paths crossed and that I could even be part of your wedding. V, it's always refreshing to talk to you. Philipp, teaching, researching, sharing experiences with you is always seamless and enjoyable.

To all the wonderful co-authors I had during this time, each collaboration enriched this work and my development as a researcher. While I'm grateful to everyone, I want to particularly acknowledge **Jurgis Karpus**, **Katrin Schulleri**, **Florian Müller**, **Robin Neuhaus**, **Benjamin Tag**, and **Takuro Nakao**.

To the great people I met along the way: **Abdallah El Ali**, for your energizing presence, your intrinsic motivation, and your honesty about research and goals. Since the moment we met at Augmented Humans, every discussion with you has been enormously constructive. To **Kai Kunze**, for the mentorship and involvement in the Augmented Humans community. To **Passant Elagroudy**, for all the advice you've constantly given me as a fellow immigrant. To **Rakesh Patibanda**, I am so lucky to have gained a friend like you during this journey. To **Pascal Knierim**, for the strategic coaching early in my PhD, when I needed it most.

To the friends who have been with me since my master's: **Aimée Sousa Calepso**, for always being supportive and understanding, and for sharing so many experiences during my PhD time. To **Victor Adriel de Jesus Oliveira**, one of my earliest mentors, for all the support coming to Europe and adapting to this new environment. I'm grateful to have you both with me since the beginning.

To my family: To my mom, **Ana Salazar**, for understanding that your son had to travel the world: Brazil, France, Germany, to chase his academic dreams, and for always supporting me along the way. For being proud of me, which has fueled me to continue every day, even in the hardest moments. To my dad, **Carlos Villa** for planting the seed of technology in my head since I was a child, this document is the consequence of you making me curious to understand how all those circuits I tried to fix worked, how computers work. To my sister **Jeimy Andrea Villa**, who was also my sponsor in my first adventures, I literally could not be here without you, thank you for being so caring for the family. People generally like to think their achievements are consequences of their own actions, but my achievements are absolutely, in large part, the consequences of your actions, and your vision. Thank you all for your support.

To my nieces, **Valeria** and **Isabela**, who are like my younger sisters and make me proud every time I talk to them, my day is always happier when I see you. To my cousin **Andrés Agudelo**, for your support during this last mile of my PhD. To my uncle **Gerardo Villa**, for the support at the beginning of my academic career, and for teaching me mechanics.

To my brother from another mother, **Camilo Ruiz**, for being someone I can talk to about everything, for giving me critical feedback when I need it, and for sharing the greatest experiences every time I return to Colombia. Talking with you over some beers, riding motorcycles across the country, having fast-paced discussions, it's almost therapeutic.

Finally, to my wife **Vanessa Jurado**, my best friend, confidant, for encouraging me to apply to this lab, to give it a try, to believe in me more than I did at that time. For helping me bear the insecurities that emerge during a PhD. For supporting me during the long and intense sprints every year. I am so glad that you are such a great scientist and understand the struggles of this work. Thanks for being next to me, now closer than ever, but for so many years more than 10,000 km away, and this never diminishing the intensity of your love.

Todo fue positivo en el proceso, aun lo negativo; aprendimos a vivir.
(...) fue un viaje de aventuras por el conocimiento y la experiencia, azaroso y venturoso;
y en los viajes es real el sueño como la pesadilla.

Thanks

Preface

> *Augmentation technologies are like lenses: they can sharpen, distort, or reveal realities. This thesis examines what happens when we begin to look through these new lenses.*

What started as a naive fascination with providing “superpowers” gradually transformed into a rigorous, sometimes obsessive ideal of defining and correctly framing the terminology surrounding human augmentation. Along this path, I confronted the inherent Human Factors vs. Design Dilemma that lies at the core of HCI and human augmentation research. Yet beyond these academic challenges, I came to realize that science and HCI research surpasses the texts of papers or prototypes. Understanding the phenomena we study also requires understanding the community that studies them.

This community, diverse and collaborative, has shaped me as much as I hope this work might contribute to shaping it. As Donna Haraway once said, “We are all cyborgs now.” Indeed, we stand at the intersection of humans and technology, where augmentation is no longer a dream but a tangible reality demanding reflection and multidisciplinary engagement.

The themes addressed in this thesis require expertise that spans psychology, psychometrics, hardware prototyping, and cognitive modeling. Such breadth of knowledge is impossible for any single individual to possess. To meet the demands of this work, I collaborated with experts across disciplines, brilliant individuals who brought their skills and insights to our shared efforts. As a result, I often use “we” throughout this thesis to acknowledge the collective nature of these findings, conclusions, and reasoning. Where I use “I,” it reflects my personal interpretations, views, or findings, without implying they are shared by my co-authors.

This thesis, then, is more than the culmination of years of research. It is the record of a journey, through concepts, collaborations, and discoveries, that seeks to illuminate the ways technology shapes human experience and, in turn, how humans shape technology.

Let’s Begin

I

CONTEXT

1 INTRODUCTION

"There can be no question of which was the greatest era for culture; the answer has to be today, until it is superseded by tomorrow."

– Steven Pinker.
***Enlightenment Now: The Case for Reason, Science, Humanism, and Progress*. 2010.**

As technology advances, so do human capabilities. AI assistants now enhance our cognitive processes, and augmented reality glasses enrich our perception of the environment. These developments have undeniably moved humanity forward, yet many technologies remain disconnected from human psychology and societal contexts. While earlier issues, such as the rejection of assistive technologies or privacy concerns surrounding devices like Google Glass, were relatively contained, the growing integration of technology into daily life poses bigger challenges. As technology becomes an intrinsic part of the environment and ourselves, its influence on social norms and interpersonal relationships intensifies, potentially leading to social friction or exclusion [19, 20].

To make technological development more psychosocially viable, especially in the field of human augmentation, we need to ask an important question: *Who or what should be the subject of analysis in the* ***future****?*. [1] When Chignell, Hancock, and Takeshita [21] introduced this question nearly thirty years ago—at a time when AI was rule-based, real-time processing was rare, wearable technology was almost non-existent, and connectivity was limited—they likely envisioned 2025 as a distant future. Yet, from the perspective of human augmentation, this question remains unresolved or even overlooked.

Despite this, this question seems more urgent than ever. Generative AI now replicates aspects of human cognition [Aut3][2], augmented reality (AR) enhances sensory experiences [22], and electrical muscle stimulation (EMS) systems improve motor abilities [23]. As these technologies reshape human capabilities, we, the researchers and designers of augmentation technologies must ask: Should our focus lie on the human, the technology, or the dynamic relationship between them?.

This thesis argues that human augmentation must re-embrace a human-centered design approach on two levels: First, in the traditional sense, ensuring technology is intuitive and meets users' needs[3]. But more critically, at a higher level, technology must be socially

[1] Chignell, Hancock, and Takeshita [21] originally posed this question in experimental psychology, but it has grown increasingly relevant for HCI in recent decades.

[2] Across this work, references to papers included in this thesis are marked with the prefix **Aut**, for publications related to Autonomy, **Het**, for Heteronomy, and **Exp**, for Simulation and Experience.

[3] Aligned with usability standards, that emphasize this point, see ISO 9241-210 standard, for example.

acceptable within the augmented human's context and designed in ways that nurture social, psychological, and physical integrity [4]. This is, bringing back the focus/goal from novel technology development to humanity-centered human augmentation. In short, designing augmentations that respect human behaviors, cognition, and social dynamics [Het1, Aut1, Aut2] while enhancing capabilities without disrupting psychological or social harmony.

We can better address design challenges, such as identifying technologies or applications that risk isolating/segregating individuals from their groups, by establishing a clear focus on who is at the center and understanding how augmentation technologies reshape human interactions and analyzing their potential to alter power dynamics within relationships. Thus, this thesis advocates for designing augmentation technologies that empower individuals and integrate seamlessly into human life [24]. While dystopian visions often highlight technology's potential to disrupt society or diminish agency [25], history has shown that thoughtful design can enhance human capabilities and enrich lives [26]. By prioritizing human and humanity needs, researchers can create systems that amplify human potential while preserving what defines us, as envisioned by Engelbart [27, 28].

To bring this vision to life, I propose viewing human augmentation through the lens of autonomy optimization. This perspective considers any bias, unwanted behavior, societal norm violation, or negative attitude introduced by technology as a restriction of autonomy, shifting the focus away from the human and toward the technology. I examine this through human decision-making processes, such as risk-taking and agency, societal factors that shape the perception of augmentation, and building augmentation prototypes that enable people to experience human augmentation. These insights are framed using the dual concepts of autonomy and heteronomy: autonomy encompasses the individual's self-directed decisions and goals, while heteronomy captures external influences that shape those decisions. These concepts, drawn from the Kantian philosophy of ethics, provide a framework for understanding the interplay of internal and external factors that shape the augmented human's everyday life.

RQ1 *How do augmentation technologies influence individual human behavior and decision-making processes?*

RQ2 *What are societal perceptions of augmented humans, and how do these affect their social acceptance and integration?*

RQ3 *How can humans meaningfully experience and interact with augmentation technologies to explore their potential and limitations?*

This work is structured into three main parts: *Autonomy*, *Heteronomy*, and *Experiencing Human Augmentation*, following the current **Part** (**Part I: Context**). In **Part II: Autonomy**, we address **RQ1**, we investigated the influence of beliefs and narratives surrounding human

[4] Reflecting recent proposals, such as Don Norman's Humanity-Centered Design perspective

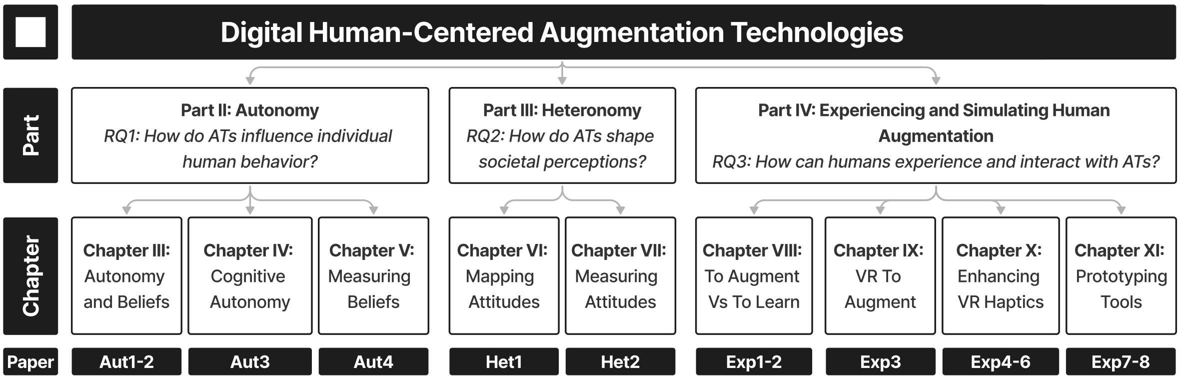


Figure 1.1: Thesis structure overview: Organized into three main parts (**Autonomy (Aut)**, **Heteronomy (Het)**, and **Experiencing Human Augmentation (Exp)**), each addressing a distinct research question (RQ) through dedicated chapters and corresponding papers. Please note that **Part I: Context**, which sets the stage, and **Part V: Ends and Means**, which synthesizes this work, are not included in the diagram.

augmentation on risk-taking and agency through four key publications: [Aut1], [Aut2], [Aut3], and [Aut4]. First, in [Aut1], we examined how the belief of being augmented influences risk-taking behavior using a standardized risk-taking task. Building on this, [Aut2] explored cognitive agency by reviewing the literature for physiological markers and proposing an EEG-based method to measure agency. Then, in [Aut3], we designed and built a prototype that augments face-to-face conversations using AI, analyzed the impact on conversational flow, and extracted insights regarding the sense of agency within that context. Finally, in [Aut4], extending insights from [Aut1], we developed a psychometric tool specifically designed to measure prior expectations regarding the use of human augmentation technologies. Then, after touching on factors that influence the individual's autonomy, I move toward the external layer, which addresses societal factors.

To address **RQ2**, we conducted two major investigations in **Part III: Heteronomy**. In [Het1], we mapped societal attitudes toward augmented humans across four distinct cultures: Colombia, Germany, Japan, and the United States. From this exploration, we identified key factors influencing bystander and observer judgments of individuals using human augmentation technologies. Building on these insights, [Het2] standardized the measurement process by developing a questionnaire based on the factors identified in [Het1] and existing surveys on performance-enhancing technologies. The resulting questionnaire revealed two main factors—*Perceived Autonomy* and *Social Threat*—which closely align with the concepts of *Autonomy* and *Heteronomy* explored in this thesis. Finally, after exploring autonomy and heteronomy within the context of human augmentation, I showcase the experience of human augmentation and potential ways to simulate human augmentation in VR.

To address **RQ3**, we conducted a series of investigations in **Part IV: Experiencing Human Augmentation**. We begin by discussing the distinction between *Augmentation* and *Learning*, illustrated through a practical use case of Electrical Muscle Stimulation for motor learning in [Exp1, Exp2]. In [Exp3], we report findings from a field study with students developing human augmentation experiences in VR, where the lack of haptic feedback emerged as a major obstacle to fully experiencing human augmentation. Building on this, [Exp4], [Exp5], and [Exp6] explore methods to enhance sensory perception in VR: [Exp4] investigates ultrasound haptics for rendering material properties, [Exp5] presents a room-scale ultrasonic rendering system using a robotic arm, and [Exp6] explores the use of robotic arms to simulate temperature sensations.

Additionally, we present two tools for facilitating human augmentation experiences. [Exp7] is a plugin that enables rapid development of haptic experiences in VR using serial robots, while [Exp8] is a haptic vest designed to render tactile sensations on the back, demonstrated through a use case involving emotion rendering.

To conclude, in **Part V: Ends and Means**, I synthesize the key insights gained from this series of investigations and discuss their broader implications for the field of human augmentation. I also outline essential considerations for researching, developing, and communicating augmentation technologies effectively.

1.1 Research Approach

This section outlines the key contributions of this work, structured to reflect the range of our research across empirical studies, methods, and artifact development.

Many existing approaches to human augmentation emphasize technical performance while overlooking human and contextual complexities. This thesis focuses on the human aspects that shape autonomy and heteronomy in augmented experiences. Yet, since state-of-the-art human augmentation research is closely tied to the artifacts and prototypes that enable these experiences, this thesis also includes and reports a series of technical implementations that support the study of these concepts. To do so, I drew from multiple disciplines, incorporating methodologies from psychology (e.g., standardized decision-making tasks), physiology (e.g., psychophysical methods for sensory perception), engineering (e.g., electronics, robotics), and cognitive sciences (e.g., EEG to examine physiological factors).

This work employs a combination of methods, including *Mixed Methods* approaches where applicable **[Het1, Aut3]**, predominantly *quantitative* methods **[Het2, Exp1, Exp4, Exp5, Exp6, Exp8, Aut1, Aut2, Aut4]**, and complemented by *qualitative* methods **[Exp2, Exp3]**. The contributions are classified based on Wobbrock and Kientz [29] taxonomy for HCI research.

1.1.1 Empirical Research Contributions

The core of the research presented in this thesis falls under the category of empirical research, as we adopt an empiricist approach to HCI. This perspective emphasizes experimentation and inductive reasoning to generate insights. Throughout this thesis, our empirical contributions span a variety of methods and study designs, as outlined below.

Controlled Studies:

Many of our investigations required controlled environments to ensure reliable operationalization of variables. For example, in **[Exp1, Aut1, Aut3]**, we relied on established standardized tasks from the literature. Specifically, in **[Aut1]**, we employed the Columbia Card Task (CCT) to measure risk-taking behavior, a widely used instrument in psychology. In **[Aut3]**, we utilized a Discourse Completion Task (DCT) to analyze how augmentation influences verbal articulation in face-to-face conversations. Similarly, in **[Exp1]**, we applied a Mirror Drawing Task (MDT) to assess motor skills under augmented conditions.

For studies focusing on sensory perception, we adapted psychometric methods to suit the needs of our experiments. For instance, we employed adjusted versions of well-known approaches like the two-alternative forced-choice and staircase methods in **[Exp4, Exp5]** to obtain measurements of sensory thresholds and perceptual experiences.

Field Studies:

In **[Exp3]**, we conducted a field study to explore how VR developer trainees design and develop human augmentation technologies over the course of a semester-long program. By analyzing the outcomes and interviewing a subset of participants, we gained insights into their design processes, challenges, and overall experiences.

Physiological Measurements:

Several of our studies incorporated physiological data to understand human responses to augmentation technologies better. In **[Exp4, Aut1, Aut2]**, we used EEG measurements to capture physiological responses to specific stimuli. These measurements provided objective markers of cognitive and sensory processes. In detail in **[Aut1]**, we performed an Event-Related Potential (ERP) analysis and studied specifically the P300 component in the central area of the head, while in **[Aut2]** we performed a spectral analysis across the alpha and beta bands. In **[Exp4]** we again performed an ERP analysis but focused on mismatch negativity in the N3 component.

Data Modelling:

To derive nuanced insights from empirical data, we employed data modeling across multiple studies. In **[Het2, Exp1, Exp5, Aut1, Aut4]**, we applied tailored models to fit the data and analyze underlying phenomena. For instance, in **[Aut1]**, we used a Bayesian model to account for censored data in the CCT analysis. In **[Exp1]**, we fitted a logarithmic decay model to examine the learning effects of Electrical Muscle Stimulation (EMS). Structural equation modeling was employed in **[Het2, Aut4]** to validate scales and subscales in the developed questionnaires. Additionally, in **[Exp5]**, psychometric models were applied to analyze tactile sensitivity and detection thresholds.

Interviews:

Qualitative methods played a central role in several of our investigations, allowing us to uncover insights that would have been difficult to quantify. In **[Het1, Exp3, Aut3]**, we conducted interviews to explore underlying factors influencing human experiences with augmentation technologies. Through thematic analysis and coding, we identified recurring themes, motivations, and challenges, providing a richer understanding of the human aspects of augmentation [30].

1.1.2 Artifact Contributions

To make the experience of human augmentation more tangible, we developed several prototypes that allowed individuals to experience augmentation rather than merely imagine it. These prototypes played a crucial role in demonstrating sensory augmentation and enhancing user understanding of the possibilities and limitations of augmentation technologies.

Prototypes:

In **[Aut3]**, **[Exp2]** and **[Exp8]**, we created prototypes that augment participants sensorily, enabling them to experience augmentation effects directly. This approach was particularly significant in **Part IV**, where we focused on creating richer and more immersive VR augmentation experiences **[Exp5, Exp6, Exp7]**. To achieve this, we designed and implemented custom components, such as specialized mounts **[Exp5]** and end-effectors **[Exp6]**, tailored to specific experimental setups. These hardware solutions facilitated the exploration of novel sensory interactions, including haptic feedback, material properties, and temperature simulation, advancing the practical development of augmentation systems. Furthermore, in **[Exp3]**, we explore multiple prototypes of human augmentations in VR.

Multi-system integration:

We implemented multi-system integrations in several studies to achieve real-time interactivity and functionality. For instance, in **[Exp1]**, we integrated the Falcon delta robot with the Unity Game Engine and the Let Your Body Move toolkit from [31], enabling seamless real-time interactions, **[Aut3]** required to integrate the Meta Quest Pro headset with the Microsoft Azure and OpenAI APIs. Similarly, the setups described in **[Exp4, Exp5, Exp6, Exp7]** required integrating the Unity Engine with additional hardware, such as the Kinova serial robot, the Ultrahaptics mid-air ultrasonic haptic system, and the Leap Motion hand-tracking system. These integrations allowed us to explore complex interactions, combining robotic systems, mid-air haptics, and real-time hand tracking to deliver richer, more immersive augmentation experiences.

1.1.3 Methodological Contributions

Throughout our exploration of human augmentation, we frequently encountered gaps in the availability of appropriate tools and methods to address our research questions. To overcome these challenges, we developed two standardized questionnaires and two tools to support prototyping efforts, contributing to the advancement of methods in HCI and human augmentation.

Scale Development:

HCI, and particularly human augmentation, remain emerging fields where tools and methodologies have often been borrowed from other disciplines. Recently, there has been a growing trend toward developing domain-specific tools tailored to the unique needs of HCI research. In alignment with this shift, we contributed to human augmentation methodology by developing two standardized questionnaires:

- In **[Het2]**, we introduced a questionnaire to measure societal attitudes toward users of augmentation technologies, enabling a systematic understanding of public perceptions across diverse contexts.
- In **[Aut4]**, we developed a questionnaire designed to measure expectations of technology use prior to interaction, filling a methodological gap in assessing user preconceptions and biases.

Tools:

To facilitate the prototyping of human augmentation systems, we developed two tools that streamline and enhance research workflows: In **[Exp7]**, we contributed Cobity, a plug-and-play plugin for the Unity3D game engine. This tool enables researchers to quickly prototype VR experiences involving Kinova Cobots. As of today, Cobity has been successfully utilized in studies such as **[Exp5, Exp6]**. In **[Exp8]**, we developed a prototype capable of rendering tactile sensations on the user's back, which we showcased through an emotion-rendering use case. This tool has since been used in additional studies, including [8, 14].

1.1.4 Survey Contributions

In **[Aut2]**, we systematically examine studies at the intersection of augmentation and agency, identify critical research gaps, and propose methods for measuring agency in the context of cognitive augmentation. This survey supports the development of more robust frameworks for understanding and evaluating agency in human augmentation.

1.2 Research Context

This thesis is based on research conducted at the Chair for Human-Centered Ubiquitous Media and the Media Informatics Group at LMU Munich over a period of approximately five years. My primary supervisor was Prof. Albrecht Schmidt, head of the research group. During this time, I collaborated with researchers and project partners across multiple institutions, engaging in studies, co-organizing events, and co-authoring papers.

1.2.1 HIVE-Lab

From the beginning of my PhD until the end of 2021, the HIVE-Lab was my primary source of funding and support. Within this context, I developed many of the VR and Haptics papers presented in this thesis. The HIVE-Lab research network established two complementary "Living Labs" in Düsseldorf and Munich to facilitate applied research and innovation. These labs supported research and development projects through technical expertise, such as selecting appropriate algorithms, while enabling the evaluation of technical innovations in both real and simulated everyday environments.

In addition to technical support, the HIVE-Lab created a publicly accessible knowledge base to promote the transfer of scientific know-how to economic and social domains. This knowledge base was disseminated through public events and open-access publications. The lab also offered design, ethical, and social consulting services to its research, economic, and societal partners.

1.2.2 AMPLIFY

Following my involvement with the HIVE-Lab, I joined the AMPLIFY project, a program that strongly aligns with the core themes of this dissertation and was a significant motivation for me to join the research group. AMPLIFY explores how digital technologies can augment, amplify, and enhance human perception.

While technical sensor systems have advanced to outperform human senses in specific areas, such as cameras with higher spatial and temporal resolution and broader spectral ranges, human perception encompasses far more than mere sensory input. AMPLIFY seeks to seamlessly integrate technologies that extend human perception and cognition without causing information overload. These technologies aim to enrich how individuals see, hear, and feel, ultimately shaping their experiences and cognitive processes.

This project explores such advancements on multiple levels: technical, conceptual, individual, and societal. By augmenting and amplifying human perception, AMPLIFY opens new possibilities for human-environment interaction, with transformative potential for human experience and understanding.

1.3 Thesis Outline

This thesis follows a hierarchical structure organized into parts, chapters, and papers, with papers representing the smallest unit. In most cases, one paper corresponds to a single chapter. However, when a concept from one paper is closely related to another, two or even three papers have been grouped into a single chapter. The thesis is divided into five parts (For details on the parts II to IV see Figure 1.1).

Part I: Context provides the necessary introduction and background to understand the main content and conclusions of the thesis. It consists of two chapters: *Introduction* and *Related Work*. **Part II: Autonomy** contains three chapters. The first chapter, *Autonomy and Beliefs,* explores decision-making and autonomy in the context of human augmentation. The second chapter, *Cognitive Autonomy,* investigates cognitive augmentation during conversations and its relationship to autonomy. The third chapter, *Measuring Beliefs,* introduces a standardized questionnaire designed to measure users' expectations of technologies prior to their use. **Part III: Heteronomy** consists of two chapters. The first chapter, *Mapping Attitudes,* presents a multi-cultural, mixed-methods study that maps attitudes toward augmented humans. The following chapter, *Measuring Attitudes,* proposes a standardized approach to measure these attitudes systematically. **Part IV: Experiencing and Simulating Human Augmentation** includes three chapters. The first chapter, *To Augment vs. To Learn,* examines the distinction between learning and augmentation in the case of motor learning, specifically using EMS. The second chapter, *VR to Augment,* explores virtual reality as a tool for augmentation and introduces techniques to enrich VR environments within this context. The final chapter of this part, *Prototyping Tools,* reports on two tools developed to prototype augmentations. **Part V: Ends and Means** comprises three chapters. Here, the implications of the research are discussed, alongside its limitations, future work, and conclusions.

2

BACKGROUND AND DEFINITIONS

"I learned very early the difference between knowing the name of something and knowing something."

– Richard P. Feynman

This chapter outlines the foundational concepts underlying this thesis, focusing on clarifying the conceptual scope of the thesis. It begins with an overview of human augmentation as a field and a concept, followed by a discussion of autonomy and heteronomy, which are informed by Kant's philosophy and adapted to the context of human augmentation.

2.1 Human Augmentation as a Field and as a Concept

The concept of enhancing human potential is not new. For centuries, humankind has explored methods to augment natural abilities using various techniques and technologies, such as cognitive enhancements [32, 33] and genetic modifications [34]. Broadly, the practice of improving human skills or capabilities, irrespective of the approach, is referred to as *human enhancement*. When these enhancements are achieved through *digital technologies*, the term human augmentation is used [35]. Which, since 1962, has been the vision initially introduced by Engelbart [27].

The origins of human augmentation can be traced to the development of assistive technologies designed to compensate for impairments [24, 36]. For example, initial sensory augmentations emerged to address sensory deficits, but their adoption by non-impaired individuals led to improved capabilities, such as enhanced hearing [37] and vision [38]. Similarly, motor augmentation technologies, such as exoskeletons, originally intended to restore mobility, now enable users to achieve capabilities beyond natural human limits [39].

Still, the research field emerging around the term *human augmentation* is relatively young, and a shared understanding of its core concepts is still in flux [35, 36]. As a result, definitions of human augmentation often remain broad [36], making it challenging to delineate what constitutes human augmentation and what does not. For instance, consider a lab technician using an electron microscope. This scenario satisfies some of the cited conditions (e.g., overcoming biological limits)—their vision is significantly enhanced, and modern microscopes employ Artificial Intelligence (AI) for image processing. Yet, most people would not view this as a prototypical example of an *augmented human*.

Nevertheless, as research on human augmentation aspires to form a research agenda, some consensus is emerging regarding its foundational aspects [40]. Human augmentation generally involves a close relationship between humans and technology to achieve goals unattainable independently [27, 28, 41]. Importantly, the agency over the task is typically localized within the human; the technological component is often conceptualized as subordinate, not as a partner [42, 43]. In other words, technology enhances the human but does not gain task-level agency, even when it incorporates a degree of intelligence.

This distinction is particularly relevant to this thesis, as it contrasts with alternative visions of human-computer relationships where agency is more distributed [44]. For instance, in Licklider [45] concept of human-computer symbiosis, both human and machine share control over a task, each complementing the other's strengths and mitigating their respective weaknesses. A similar perspective underlies human-computer integration, introduced by Farooq and Grudin [25] and later refined by Mueller et al. [46], which builds on the idea of symbiosis but emphasizes sensory fusion—seamless communication between humans and computers through integrated sensors. In both cases, the technological system is not merely a tool but an active collaborator in decision-making and execution.

Depending on the type of technology used and the function it performs (input, output, or processing), human augmentation can be categorized into three primary types: *sensory*, *motor*, and *cognitive* augmentations [35, 47]. Although Some taxonomies also include *social augmentation* as an additional category [36][1]. Under this framework, the distinction between human augmentation and overlapping concepts, such as human enhancement, becomes clearer: a genetically enhanced human is not an augmented human, whereas a technologically (digital technologies) enhanced human is. Extending human capabilities is used not only to overcome impairments but also to enhance users so that they can unlock new potential for interacting with their environment.

> **Definition – Human augmentation:** Human augmentation is the discipline that seeks to enhance human performance and skills using near-body digital technologies that mediate the interaction of the individual with the self or the world. Such augmentation technologies do not assume control over the task that the augmented human performs but instead serve as the user's subordinate. [Het1]

Consequently, throughout this work, a user of digital technologies designed for human enhancement is referred to as an **Augmented Human**. This term encompasses individuals augmented by any of the types of augmentations mentioned earlier (Cognitive, Motor, Sensory) and depicted in Figure 2.1. These augmentation categories are further detailed as follows:

[1] I adopt the taxonomy proposed by Raisamo et al. [35], which classifies human augmentation into cognitive, sensory, and motor categories. This framework aligns with neuroscientific models of brain function, as social interactions rely on higher-order cognitive processes [48]. By categorizing social functions within cognitive augmentation, this taxonomy remains consistent with neuroscience [49].

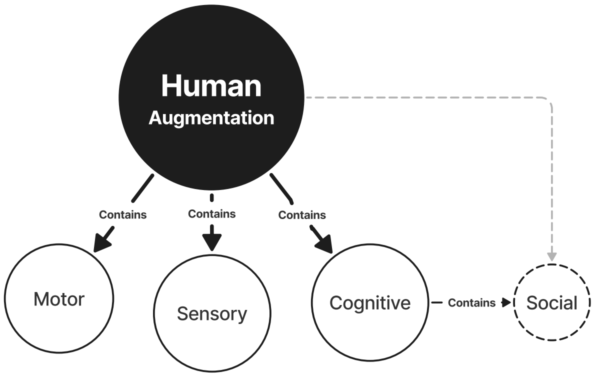


Figure 2.1: A conceptual diagram of human augmentation, depicting the relationship between motor, sensory, cognitive, and social components. The core concept contains motor, sensory, and cognitive augmentations, while the cognitive aspect also connects to social augmentation.

2.1.1 Motor Augmentation

Motor augmentation in Human-Computer Interaction (HCI) explores the integration of technology to modify and augment human physical actions [50]. This includes devices such as exoskeletons [51, 52, 53], robotic manipulators [3], which provide users access to motor capabilities beyond their biological limits or aid in recovering lost functions. Initially developed to compensate for restricted mobility [39], motor augmentation technologies now enable diverse forms of interaction, from multitasking with additional robotic limbs [43] to refining motor control in precision tasks by using Electrical Muscle Stimulation (EMS) in the back of the hand [54].

Exoskeletons, for instance, have been shown to enhance physical strength and allow individuals to carry heavy loads more efficiently [55], altering the perceived weight of objects during interaction [56]. Applications of motor augmentation span multiple domains, including, among others, healthcare, where these systems support rehabilitation and assistance [57], manufacturing [58], where they improve ergonomic performance [59], and creative industries, where augmented physical capabilities enable new modes of expression and innovation.

❯ The Case of Electrical Muscle Stimulation in Motor Augmentation

EMS involves delivering electrical impulses through skin electrodes to stimulate muscle contractions [60, 61]. EMS was designed to aid motor function recovery by mimicking the brain's natural signals to trigger muscle activity. EMS sends controlled electrical impulses through electrodes on the skin, replicating the nervous system's signals to initiate muscle contractions. These impulses activate motor neurons, causing the muscle fibers to contract. By adjusting the intensity, frequency, and duration of the impulses, EMS can be tailored to strengthen muscles, improve endurance, enhance recovery, or reduce tension. This process bypasses the brain's direct control, simulating voluntary muscle contractions [62]. Different intensities and frequencies have different effects on the perceived stimulation. Moreno-Aranda and Seireg [63] showed that high-frequency alternating current signals trigger powerful muscle contractions with minimal discomfort, for example, when treating paraplegic or quadriplegic patients. In this context, EMS expanded to rehabilitation and medical settings [64] or support fitness training [65]. In rehabilitation, EMS helps prevent muscle atrophy and promotes muscle re-education, while athletes use it to enhance strength and endurance [66]. Recently, EMS has gained significant traction in HCI research to provide feedback and steer the physical movements of users [31].

This case is explored in detail in Section chapter 8 in the context of motor learning.

2.1.2 Sensory Augmentation

Sensory augmentation in HCI explores how technologies can modify, extend, or substitute sensory input, enabling new ways of perceiving and interacting with the world. These innovations include devices that enhance existing senses, such as integrating thermal imaging into vision and those that provide sensory substitution, such as translating auditory information into tactile feedback for individuals with hearing impairments. In line with early developments in this domain often focused on compensating for sensory deficits, as seen in systems designed to restore hearing [37] or vision [38].

More recent advancements aim to expand sensory experiences for broader applications. For instance, Abdelrahman et al. [67] demonstrated how augmented reality headsets combined with thermal cameras could enable users to perceive the infrared spectrum naturally, a capability particularly beneficial for safety-critical tasks like firefighting in high-temperature environments [22]. Similarly, augmented sensory systems provide real-time information about the surrounding environment through visual overlays or vibrotactile cues, enhancing situational awareness and decision-making, for example, by providing augmented information about the surroundings [8, 68]. By altering how sensory information is acquired and interpreted, these technologies redefine the scope of human perception in both everyday and specialized contexts.

❯ The Case of Emotions in Sensory Augmentation

Emotions, a multifaceted form of information, have garnered attention in HCI and human augmentation [69, 70] due to their relevance in various applications, including online communication [71, 72, 73], storytelling [74], and supporting neurodivergent individuals [75].

Communicating emotions through vibrotactile encoding as a form of sensory augmentation presents unique challenges because emotions are not a uniform category of information [76]. Their psychophysiological effects vary significantly [77] and depend on factors such as the body location of the stimulation [78]. Researchers have explored different encoding approaches, targeting areas like the wrist [79], forearm [80], hands [81, 82], and back [78, 83, 84, 85]. Common methods for designing these patterns involve combining vibrotactile dimensions and gathering user ratings for valence and arousal [79, 83, 86, 87], identifying perceived body locations for emotions [88], mapping facial expressions to vibration intensities [82], or creating patterns through user input [85, 89]. However, these strategies often rely on analogies to vision or are tailored to individual users, limiting their generalizability.

In Chapter chapter 11, I explore the use of a vibrotactile vest to encode emotional information through stimulation applied to the back.

2.1.3 Cognitive Augmentation

Cognitive augmentation explores the use of digital technologies to enhance, extend, or complement human mental processes. Building on early research into tools for memory, reasoning, and problem-solving [16], the field now spans diverse applications, including support for attention [90], decision-making [Aut1, 12, 16], and creative workflows. Advances in AI have led to systems offering dynamic [91] and context-aware assistance [Aut3], transforming how people engage with information and carry out cognitive tasks. Examples include memory aids like *SenseCam* [92], which externalizes episodic memory for individuals with amnesia, and wearable interfaces such as *AlterEgo* [93], enabling silent, seamless communication and information retrieval. These innovations demonstrate the deepening integration of technology into human cognition. Devices such as life-logging systems extend human memory by preserving experiences with greater detail and longevity than natural memory capacities [94, 95, 96, 97].

❯ The Case of Conversations in Cognitive Augmentation

Enhanced conversation exemplifies cognitive augmentation, as it demands the ability to perceive, process, and articulate information under time-critical conditions. I explore these dynamics further in Chapter chapter 4. Building on this foundation, research has explored how technology can integrate with in-person interactions to enhance communication. For instance, Zisk and Dalton [98] introduced the concept of "dual-purpose speech," where conversational context informs system interactions. This approach enables the system to capture and adapt to speech contexts that are not directly addressed to it, enriching its functionality. Contextual information has also been leveraged to support users in diverse ways. Kane and Morris [99] used object context to suggest relevant words for individuals with Amyotrophic Lateral Sclerosis (ALS), while Vargas, Dai, and Moffatt [100] developed context-specific suggestions for users with Autism Spectrum Disorder (ASD) using images.

Another important focus of research examines how different modalities of information delivery influence technology-assisted conversations. Ofek, Iqbal, and Strauss [101] and Cai et al. [91] investigated visual and auditory modalities, revealing that users often prefer concise visual information delivered during conversational pauses. However, preferences for auditory feedback show considerable variability across individuals [91, 101].

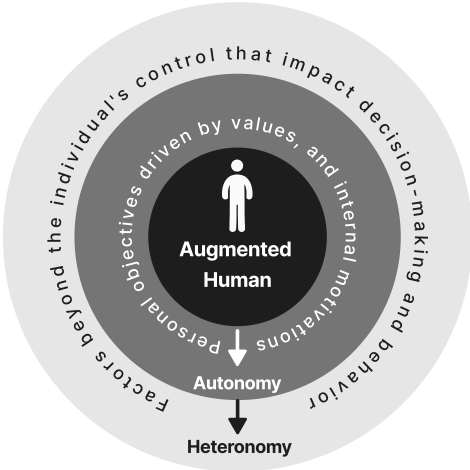


Figure 2.2: Layers of Influences in Human Augmentation: The relationship between external influences, personal goals, and the autonomy of an augmented human, while also acknowledging the presence of heteronomy as external dependencies.

2.2 The Human as Subject of Study Within Human Augmentation Research

Examples such as the societal rejection of Google Glass [102], user hesitance toward cochlear implants [103], and the unintended over-reliance on ABS technology [104], which increased accidents [105], illustrate the consequences of overlooking societal and fundamental human factors during technology development and deployment.

Human augmentation poses a distinct challenge. Unlike tools that can be set aside, many future applications—such as implants—will become integral to how individuals perceive and interact with the world [Het1, 35]. These technologies will not merely supplement human abilities but redefine them [41], raising fundamental questions about autonomy, a concept central to Kantian philosophy. Kant distinguishes autonomy from heteronomy, emphasizing self-governance as essential to moral agency. This perspective highlights the need to ensure that augmentation technologies support, rather than undermine, an individual's capacity for self-directed decision-making.

To navigate this challenge, I propose centering human autonomy as the guiding principle for human augmentation research. This involves designing technologies that expand human capabilities while ensuring users retain agency over their decisions. In cases where autonomy is inherently limited, these limitations must be explicitly acknowledged and carefully managed.In the following, I describe the concepts of Autonomy, Heteronomy, and related subconcepts addressed in this thesis:

2.2.1 Autonomy

The concept of autonomy has been a focal point in philosophy, political theory, and, more recently, HCI. Rooted in the Greek *autos* (self) and *nomos* (law), autonomy initially denoted self-governance within a political framework [106]. Kant's moral philosophy later redefined autonomy as the capacity for rational self-determination [107], emphasizing the role of reason in prescribing laws to oneself [106, 108]. This understanding positions autonomy not merely as independence but as an engagement with rationality and moral self-regulation within a socio-material context, where designers, alongside other societal actors, shape the conditions under which autonomy is exercised [108]. Such philosophical origins resonate strongly in HCI, where the notion of autonomy is often reinterpreted through the lens of user experience, decision-making, and control in human-technology interactions [109].

In HCI, autonomy frequently intersects with concepts of agency, where users act in accordance with their goals and values, supported—or constrained—by technological design [109]. The distinction between autonomy (self-governance) and agency (self-causation) is often blurred, as technologies increasingly mediate human action at multiple levels. For instance, frameworks such as Self-Determination Theory [110, 111] emphasize autonomy as a dynamic interplay between individual identity and social context [110]. Similarly, work on Human-Computer Integration highlights the tight coupling of user and technology [46], raising questions about how such integration impacts autonomy at various timescales, from immediate interactions to long-term identity formation.

Definition – Autonomy: Autonomy is the capacity of an individual to self-govern by making decisions and acting in alignment with their goals, values, and rationality. It involves the ability to evaluate options and exercise control over actions.

Desicion Making

Decision-making emerges as a critical dimension in understanding autonomy within HCI, as it reflects the capacity of individuals to act in alignment with their goals. Technologies can either enhance or undermine this capacity by influencing how users evaluate options, navigate constraints, and exercise control over their actions. For instance, design choices that obscure critical information or overly automate tasks may erode decision-making autonomy [112], while transparent and adaptive systems can empower users to make informed choices [113]. Exploring autonomy through the lens of decision-making thus provides a pathway to understanding the broader implications of human-technology interactions, particularly in how they shape users' ability to govern their lives and achieve meaningful outcomes.

Augmentation Technologies, in particular, mediate how individuals engage with their surroundings [35, 114]. By altering the user's perception of their environment and its affordances, augmentation technologies can lead to unexpected behavior or heightened risk-taking [115]. For instance, Low and Chan [116] demonstrated that excessive reliance on SCUBA-diving systems—a widely used form of augmentation—increases the likelihood of risky behavior. Similarly, Borenstein, Wagner, and Howard [117] found that individuals often trust exoskeletons in high-risk scenarios without prior system knowledge, advocating their use in contexts for which they were not designed.

This tendency to over-rely on augmentation technologies can have serious consequences, particularly when users misjudge the system's operational state [118]. For example, acting as if an exoskeleton is providing support, such as when lifting heavy loads, without verifying its active assistance, can lead to physical harm [119]. Consequently, the expectation of enhanced abilities through augmentation may unintentionally increase risk-taking during decision-making.

➜ The Case of Risk-Taking Decision-making involves selecting one option from a set of alternatives [120]. This process can be driven by calculated assessments of risks and benefits, or by emotional responses and intuitive judgments about the options. The former is referred to as "cold" decision-making, while the latter is known as "hot" decision-making [121]. Risk-taking, in particular, is the act of choosing an uncertain course of action [122]. Prior research suggests that individuals who engage in risky behavior often do so because they perceive the potential benefits of an action to outweigh its possible consequences [123]. However, risk-taking is not always rational, as it is influenced by cognitive biases tied to personality traits, age [124], group-specific tendencies [125, 126], and self-assessments such as perceived skill levels, which are associated with higher risk-taking [127].

Patterns of decision-making under uncertainty can be effectively studied using behavioral tasks in laboratory settings [121]. Commonly employed tasks include the Iowa Gambling Task [128], the Balloon Analogue Risk Task [129], and the Columbia Card Task (CCT), which has both "cold" and "hot" variants [130]. Although these tasks are abstract in nature, they demonstrate strong external validity [121].

These tasks also exhibit internal validity, particularly in their ability to measure physiological correlates such as electro-dermal activity, heart rate, functional near-infrared spectroscopy (fNIRS), and EEG responses. For example, Holper and Murphy [131] found that participants displayed increased electro-dermal activity and fNIRS responses, along with a decreased heart rate, during the hot version of the CCT compared to the cold version. They proposed that electro-dermal activity and fNIRS together provide an effective means of studying hemodynamic and affective responses.

In Chapter chapter 3, we implement and adapt the hot version of the CCT to assess affective risk-taking and analyze event-related potentials recorded in the EEG when participants believed that they were cognitively augmented.

➜ The Case of Sense of Agency in Cognitive Augmentation

The sense of agency refers to the **experience** of controlling one's actions and their consequences [132, 133], enabling individuals to recognize themselves as the agents of their behavior [134].

Both HCI and cognitive neuroscience seek to understand how people experience agency [109, 133]. To this end, various measures of agency have been developed. Direct measures typically involve explicit self-report questionnaires, such as the *Sense of Agency Scale* [135], where participants rate their subjective sense of agency using Likert scales. In addition, customized scales are often employed. Indirect measures of agency have also been proposed; for instance, Haggard, Clark, and Kalogeras [136] introduced the "intentional binding effect" as an implicit measure. This effect refers to the subjective compression of time between a voluntary action and its sensory consequence.

In neuroscience, researchers explore objective and physiological correlates of the sense of agency. Using brain imaging techniques, distributed neural networks associated with agency have been identified [137]. Neurophysiological indicators, such as EEG data, have been found to correlate with the sense of agency, allowing for its continuous real-time measurement [138, 139]. Specifically, Kang et al. [139] demonstrated that EEG alpha band activity, a key neural metric for information processing [140], serves as a primary neural oscillation linked to the sense of agency, with the anterior frontal lobe playing a crucial role in generating this sense. Further research has shown that increased alpha and beta EEG power in the parieto-occipital regions correlates with changes in the sense of agency during hand movements with delayed visual feedback [141].

The beta frequency, like the alpha frequency, is particularly important for understanding the neural basis of agency. Modulation of local population activity and changes in functional connectivity mediated by the beta band suggests a network-level perspective on the agency. For example, Buchholz et al. [142] found that during the belief of agency, the primary motor cortex (M1) exhibited stronger beta-band functional connectivity with the inferior parietal lobe and right middle temporal gyrus, suggesting a neural network that supports agency based on varying causal beliefs.

Neurotechnologies, such as Brain-Computer Interfaces (BCIs), can intervene between a user's intention and action, potentially altering the sense of agency [143]. For instance, Haselager [144] found that BCIs might disrupt the sense of agency, leading users to question their role in actions. They concluded that raising awareness of this issue could guide the development of solutions to mitigate its effects.

In chapter chapter 3, we explore how the feeling of being augmented impacted the physiological markers of agency in the alpha and beta bands.

Definition – Agency in Human Augmentation: Agency in human augmentation is an individual's sense of control and purposeful engagement with technological tools that enhance cognitive, motor, or sensory abilities. It means actively choosing, directing, and managing technological interventions that support personal cognitive functions like memory and decision-making, motor functions such as walking, and sensory functions such as seeing, and hearing, while ensuring the person remains the primary driver of their enhanced capabilities [109]

2.2.2 Heteronomy

In Kant's philosophy, heteronomy refers to situations where actions are influenced by external factors or authorities rather than being guided by a person's own reasoning. This contrasts with autonomy, which means acting based on one's own rational principles [145]. Stensson and Jansson [146] pointed out that terms like autonomy and intelligence are often applied to technologies without clear scientific backing. They suggested revisiting Kant's ideas on autonomy and heteronomy to better understand how human and technological roles should be defined in design contexts [147].

Kant described autonomy as the ability to reason and act independently, without being persuaded by external pressures or personal biases[2]. He described it as *"the principle of autonomy of the will, in contrast with every other which I accordingly reckon as heteronomy"* [148]. Autonomy focuses on the rights and self-imposed responsibilities of someone capable of self-directed action, while heteronomy involves being influenced by external rules, prior knowledge, or cognitive biases [146].

In HCI, heteronomy can be seen as the external factors that shape how users make decisions. Augmentation technologies might include social norms, bystander opinions, privacy concerns, or other outside influences. These factors might cause users to act differently than they would if they were fully autonomous—making decisions entirely free of external constraints or the effects of the augmentation itself, coherently to how human augmentation is defined and understood in this work.

Definition – Heteronomy: Heteronomy, in the context of HCI, refers to the influence of external factors—such as social norms, authority, prior knowledge, or biases—on a user's decision-making or behavior. It exists on a continuum, with autonomy at one end, representing self-directed actions guided by rational principles, and heteronomy at the other, where external pressures or constraints dominate. In augmentation technologies, heteronomy can involve influences like privacy concerns, societal expectations, or bystander attitudes that shape how users interact and make decisions.

Society's Influence on Emerging Technologies

An individual, as part of society, might be influenced by societal beliefs, concerns, and biases, making the generalized attitudes of society play a role as a heteronomous factor in the individual's autonomy/ Society's perception of technology has the power to shape its trajectory. Negative perceptions can inhibit widespread acceptance and adoption, and the

[2]This idea is rooted in his Categorical Imperative, a key concept in his moral philosophy.

fear of social stigma can prevent early adopters from embracing new technologies [20]. This can result in a vicious cycle in which the lack of early adopters leads to additional negative perceptions of a specific technology, which in turn discourages potential adopters [149]. This is a critical challenge for emerging technologies, as preconceived biases can cloud the objective evaluation of their advantages and disadvantages, thereby limiting their potential for positive impact [150]. Emerging technologies such as AI [151], Robotics [152], and Human Augmentation [68], are particularly susceptible to this issue partly due to their extensive coverage in science fiction literature, movies, the news, and social media [153] as well as because of their closeness to or their resemblance of the human body [154]. These negative attitudes can be attributed to social validation, perceived aesthetics, intrinsic motivation, and technology-related stigma [155]. Collectively, these factors influence an individual's perception in terms of social acceptability and, ultimately, their willingness to adopt a new technology. To illustrate this, users often refuse to accept assistive technologies [156], even though many such tools have been shown to effectively compensate for users' hearing [157], sight [158], and movement [159] impairments. Since technology acceptability has been recognized as a key concern in HCI [149], various instruments for measuring public opinion on technological innovations have been developed. For instance, measurement scales based on the technology acceptance model [160], The WEAR scale [161], and, more recently, the creepy technology scale [162], amongst others. Note that these scales often focus on the technology itself and not on the person using it, thus may not apply to cases of human-computer integration, where the lines between technology and the user blur. As users are increasingly likely to experience augmented humans in their everyday lives, it is increasingly important for HCI to investigate attitudes toward augmented humans as a heteronomous factor in augmented human psychology. To address this, I examine two key cases: the societal attitudes towards human augmentation and how to ecologically measure social attitudes toward performance-enhancing technologies, including human augmentation, detailed as follows

❯ The Case of Society and Human Augmentation

Recent years have seen increased interest and research into the question of social acceptability's influence in shaping the evolution of technology [163, 164, 165]. Koelle, Ananthanarayan, and Boll [19] posited that a human-machine interface is sociably acceptable if its existence or the user's interactions with it are congruent with the user's self-image and external image, or positively affect them. In response to these demands, we intend to identify the aspects that influence the perception of human augmentation technology users. Social psychology, neuroscience, and ethics research have identified a core group of dimensions relevant to the assessment of and experience with human enhancement (Enhancing humans in a broader sense, not necessarily using digital technologies), and these are repeatedly featured in research. For example, Fitz et al. [166] reported that safety, pressure, fairness, and authenticity are the dimensions that modulate public attitudes toward human enhancement. In detail, they described *safety* as the analysis of risk and benefits of cognitive enhancement for the individual (this dimension is also addressed by Scheske et al. [167]). The *pressure* dimension is defined as the social pressure to have augmentations (similarly defined by Dubjevic et al. [168]). The *fairness* dimension is reported instead as the sentiment of distributed justice, balance, and feelings of cheating [169]. Finally, the *authenticity* dimension was expressed as the impact of the enhancement on the individual's character and worthiness of achievement [170]. In addition, Conrad et al. [171] demonstrated that people are more open to others using enhancements than using them themselves. Though these dimensions are consistently reported in the human enhancement field, it is not clear if they are transferable to human augmentation technologies. Therefore, an understanding of how people assess augmented humans from the HCI perspective is required. We aim to fill this gap by conducting the first study on the perception of augmented humans across a diverse sample.

Moreover, how the enhancements are communicated to the public influences the attitudes toward human enhancement. Evidence shows that the terminology used in the discourse about human enhancement impacts the acceptance and attitudes of these technologies [171]. As an illustration, using the word "fuel" instead of "steroids" evokes less negative attitudes. Therefore, how augmentations are articulated can impact society's attitudes toward AHs. To take this aspect into account, we have worded both our survey and our interview protocol as neutrally as possible.

Recent work in HCI reported that the level of integration of augmentation in the body plays a role in its acceptance. Specifically, Rousi et al. [172] studied cognitive enhancement from the body's perspective; they refer to these levels of integration as Endo (in-body), Exo (wearable/embodied), and External (environment). Rousi et al.'s work addressed emotional attitudes toward human augmentation technologies, wearable devices, smart clothing, smart glasses, and what the authors refer to as cognitive enhancement games. For example, they found that people are less willing to use a brain or eye implant than smart glasses or smart textiles, suggesting that the integration level impacts augmentation acceptance.

While prior research has explored public opinions on enhancements, detailed insights into attitudes toward technology-enhanced humans (augmented humans) remain limited. This work addresses this gap through a mixed-method approach, examining the factors shaping attitudes toward augmented humans and the perception of different types of augmentations through these factors.

In Chapter chapter 6, we analyze societal attitudes toward augmented humans, identifying key factors that influence judgments about users of human augmentation technologies.

❯ The Case of Measuring Social Attitudes Towards Performance-Enhancing Technologies

Psychology and medicine have extensively studied attitudes toward performance-enhancing technologies that do not rely on digital computation, such as pharmaceuticals, prosthetics, and biochemical interventions, and found that these attitudes vary significantly depending on the context [173]. For instance, the use of performance-enhancing drugs in sports has been a source of disagreement for many years. It has been reported that society, at different layers, has markedly different attitudes towards enhancing supplements, depending on their social affiliations to specific groups [174]. Dijkstra and Schuijff [175] found a widespread attitude of mild disapproval to strong disapproval of using enhancement technologies for applications other than medical treatment. Moreover, their findings suggest that the acceptability of enhancement use is dependent on the motivation behind it, with socially motivated enhancements being perceived more positively than those used for personal gain. The debate around doping in sports has greatly contributed to the search for strategies and tools to measure the attitudes toward performance-enhancing technologies and create an understanding of their impact, reflected in the creation of tools like the Performance Enhancement Attitudes Scale (PEAS) [176]. Yet a big part of the efforts to measure attitudes towards performance enhancement from the medical, psychological, and sport science domains have been directed towards predicting doping behavior by connecting the attitudes and the chances that it can be correlated to the doping behavior intention and use [177].

Attitudes toward digital technology enhancements present distinct challenges compared to non-digital enhancements. For instance, while the use of anabolic steroids is widely regarded as punishable behavior [176], technologies such as exoskeletons are often framed as necessary in specific contexts [178]. More complex scenarios arise when advancements challenge conventional norms, such as cases where amputee athletes outperform their non-amputee peers [179]. These examples highlight the nuanced and evolving perspectives on human augmentation technologies, underscoring the need for robust tools to systematically assess attitudes toward these innovations.

In Chapter chapter 7, we introduce and validate a standardized scale designed to ecologically measure attitudes toward performance-enhancing technologies.

2.3 Moving Towards Digital Human-Centered Augmentations

In human augmentation practice, research, deployment, and use, multiple stakeholders are involved, including designers, companies, society, bystanders, and users. Consequently, after addressing Chignell, Hancock, and Takeshita [21]'s question of 'who' is the subject of research, it becomes essential to ask 'whose' values shape the development of human augmentation technologies. Recognizing humans as central to this field introduces additional complexity: should these technologies primarily serve individual users or broader societal needs? Although many augmentation systems aim to integrate seamlessly into individuals' lives, their design can be influenced by biases that emphasize technological features at the expense of human needs. These biases risk distorting decision-making, compromising autonomy, and neglecting cultural context, which may lead to unintended consequences for those who rely on these systems.

This tension highlights the need for a foundational premise to guide the development of human-centered augmentation technologies. I argue that these systems must prioritize individual autonomy while remaining sensitive to external influences, such as societal and cultural contexts. This dual focus ensures that technologies empower users without imposing values that undermine their agency. The guiding principle can be articulated as follows:

> *"Digital Human-centered augmentation technologies must prioritize individual autonomy while recognizing the broader societal and cultural influences shaping decision-making."*

This principle rests on two interconnected pillars: *Individual Autonomy:* Technologies must enhance users' capacity to make decisions aligned with their own values, free from coercion or undue external influence. *Heteronomy Awareness:* Technologies must account for societal, cultural, and contextual factors influencing decisions, ensuring these external influences support rather than undermine individual agency. Autonomy serves as a cornerstone for ethical technology design, drawing on frameworks such as Kantian autonomy and self-determination theory. By respecting autonomy, technologies empower users to make decisions grounded in their own goals, values, and contexts. For instance, an augmentation designed for healthcare should enable patients to choose enhancements that align with their personal needs and beliefs rather than impose a predetermined path.

However, decision-making under augmentation is rarely isolated from external influences. Societal norms, cultural principles, and even inherent biases in system design often shape choices. Researchers must also recognize that users bring their own biases to interactions with these technologies, which can influence outcomes in ways that are complex and difficult to predict. The interplay between these factors must be studied to ensure that technologies do not inadvertently constrain or distort user decisions.

Bringing the Kantian conceptualization of Autonomy and Heteronomy to human-centered human augmentation research provides a clear optimization criterion: augmented human autonomy. This framework places user autonomy at the core while also highlighting external factors that may influence an augmented human's decision-making. Augmentation technologies function like lenses—they can sharpen, distort, or reveal different aspects of reality. This thesis explores what happens when we adopt these new lenses and how to ensure they enhance user goals (autonomy) rather than impose external influences beyond the user's control (heteronomy).

II

Augmented Human Autonomy

3

DESICION-MAKING, BELIEFS, AND AGENCY

"In order to design a future of positive change, we must first become expert at changing our minds."

– Jacque Fresco.

This chapter is based on the following publications:

The Placebo Effect of Human Augmentation: Anticipating Cognitive Augmentation Increases Risk-Taking Behavior - Steeven Villa, Thomas Kosch, Felix Grelka, Albrecht Schmidt, Robin Welsch - *In Computers in Human Behavior, Volume 146, September 2023, Elsevier.*

Whose Mind is it Anyway? A Systematic Review and Exploration on Agency in Cognitive Augmentation - Steeven Villa, Lisa Barth, Francesco Chiossi, Robin Welsch, Thomas Kosch - *Computers in Human Behavior: Artificial Humans (UR), Elsevier.*

I propose autonomy as the primary driver for human-centered human augmentations, understanding autonomy as the capacity of the individual to make decisions and act aligned with their goals (see definition in the previous chapter). A central part of autonomy is decision-making. thus, in this chapter, we investigate decision-making processes, specifically in the context of risk-taking and the subjective experience of being augmented. A central focus is placed on understanding how the belief in augmentation influences decisions, particularly those involving risk. To explore this, we employ the Columbia Card Task (CCT), a standardized operationalization of risk-taking, and complement it with EEG data collection to capture neural correlates of decision-making. Furthermore, we extend this investigation to the concept of cognitive agency. We describe how EEG data can be leveraged to measure cognitive agency within the same experimental framework. This approach is grounded in a review of relevant literature, which serves to establish the theoretical foundation and methodological validity of our measurements.

Decision-Making and Beliefs

We operationalized risk-taking through the Columbia Card Task (CCT), the participants were given a revised version of the Hot CCT (which measures affective, emotion-driven decision-making) two times, one per condition. In the augmentation condition, participants were told to be supported by a cognitive augmentation. In the no-augmentation condition, participants were told the augmentation system was turned off and that any beneficial effects did not exist. We anticipated that being cognitively augmented affected the participant's risk-taking behavior during the CCT.

We used an EEG as a placebo augmentation technology and informed participants that the EEG was a Brain-Computer Interface (BCI) playing an inaudible sound that was proven to improve the ability to process information and thus perform better in the CCT. However, the setup was identical for both conditions, the system was not functional, and no sound was played. Note that our study was designed to assess the placebo effect of an augmentation technology and that we did not focus on BCIs or the efficacy of augmentation of cognitive capacities per se, only on the propensity of placebo effects of augmentation technologies to increase risk-taking.

We adapted one aspect of the CCT for our study. Typically, in each round of the CCT, a participant is presented with a set of cards face down. Behind these are loss or win cards representing the given amount the player can win or lose. For the purpose of our study, participants were briefly presented with the cards face-up. Afterward, the cards were put face down again and shuffled using an animation. Participants were led to believe by the verbal description that the augmentation would support them in tracking the cards on the screen. Participants did not know that win or loss cards were rendered at each draw and that the whole game was rigged.

Note that medical research on pain-alleviating placebos has varied a large set of contextual variables that can modulate the placebo effect in size. Wager and Atlas [180] provides a taxonomy of different contextual cues affecting placebo effects. These cues include the treatment cues (e.g., the novelty of the treatment), the place (e.g., a medical lab), the social situation (e.g., the experimenter wearing a white coat) and verbal suggestion (e.g., describing the mechanism closely); while we have taken these contextual variables into account, our study should resemble a user study for augmentation technologies as closely as possible.

To reiterate, we state the following research question:

RQ: Can anticipation of being augmented be induced by verbal description and can this increase risk-taking behavior?

We investigated the following hypotheses to answer this research question:

H1: A verbal description of an augmentation technology results in an increase in performance expectations

H2: A verbal description of an augmentation technology results in an increase in performance judgments after interaction

H3: Performance expectations improvement induced by a placebo-treatment increases risk-taking behavior.

H4: Performance expectations induced by a placebo-treatment affect processing of risk-related information.

3.1 Method

In the following, we motivate and document our methodological choices in realizing the study.

3.1.1 Participants

We recruited participants through the university's mailing lists and communication channels. To prevent study participants from detecting the placebo condition (verbal description of the augmentation system), we refrained from recruiting individuals with prior knowledge of EEG or human augmentation systems. We recruited a total of thirty participants (N = 30), one of whom did not consent to the use of their data following the experiment, and two were excluded due to poor data quality (no data was recorded concerning their expectancy ratings). There were a total of twenty-seven participants (N = 27, Male = 17, Female = 10, 0 non-binary, 0 participants did not disclose or self-specified a gender) with an average age of 29 years (M = 29.13, SD = 9.51) and a reported technical competence of (M = 4.76, SD = 1.43) in a Likert scale from 1 to 7. Participants were compensated 5 euro/30 min for their involvement.

3.1.2 Experimental design

We conducted a within-subjects lab study with four variables of interest, each with two levels. In detail, the independent variables were: (1) Verbal description, referred to as DESCRIPTION (The setup is augmenting participant cognitive skills is referred to as AUGMENTATION condition whereas the setup is not augmenting the participant's cognitive skills is referred to as NO-AUGMENTATION condition), (2) Number of loss cards (one loss card vs. three loss cards) compared to the total number of cards referred to as LOSS CARDS, (3) Value of win cards (10 points vs. 30 points) referred to as WIN AMOUNT, and, (4) Value of loss cards (250 points vs. 750 points) referred to as LOSS AMOUNT. The order of presentation of the verbal description was counter-balanced, while the CCT-related variables (LOSS CARDS,WIN AMOUNT,LOSS AMOUNT) were randomized.

3.1.3 Stimulus

Verbal description:

We compared the influence of two verbal descriptions regarding human augmentation. We did this by manipulating the system description (i.e., AUGMENTATION condition or NO-

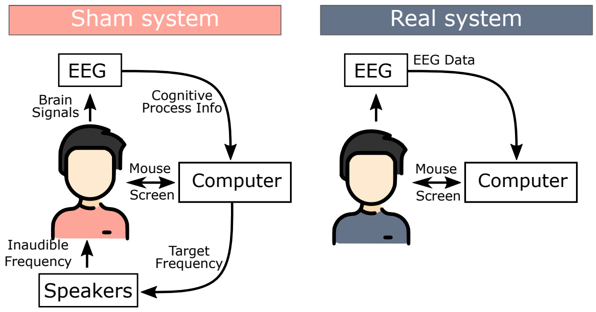


Figure 3.1: Stimulus: Verbal description; we told participants the EEG was a BCI system that modulated an inaudible sound that improves their information processing and RCCT performance. In reality, the system played no sound and was used to record data only.

AUGMENTATION condition). The participants were informed about the assigned condition before conducting the RCCT (see Figure 3.1).

During the AUGMENTATION condition, participants were informed that the BCI was analyzing their brain waves to emit an inaudible brain-stimulating sound to boost visual processing, allowing the participants to recognize the win cards more precisely. A coherent explanation of how the system works was provided to the participants. We stated that we used binaural sounds [181], which are administered through inaudible frequencies [182], are proven to have a positive impact on cognitive functions (e.g., mitigating Alzheimer's symptoms [183]). In the no-augmentation condition, participants were informed that during this condition, the augmentation device would not be active; therefore, their performance would be determined solely by their ability to visualize the cards shifting, identify the winning cards, and play the game. This condition serves as a control condition in our experiment.

Columbia Card Task Related Variables

According to [130], the risk assessment of participants in the Columbia card task is influenced by three variables: the value of win and loss cards, and the number of loss cards in the deck. We used literature-informed values for the CCT, namely, 10 and 30 for win cards, 250 and 750 for loss cards, and 1 and 3 for the number of loss cards. The value of win cards is added to the participant's total round score upon flipping a win card. In the same way, the value of loss cards is subtracted from the participant's total score upon flipping a loss card. The number of loss cards in the deck is the number of cards that can lead to a point deduction out of the total 27 cards present in the deck.

Procedure

The participants' assignment to the starting condition (augmentation or no-augmentation) was counterbalanced. Participants were supplied with an explanation of the study's design, as well as data protection and comprehensive information. The participants were then requested to grant informed consent to participate in the study in accordance with the Declaration of Helsinki and to continue with the demographics and technical competency evaluation.

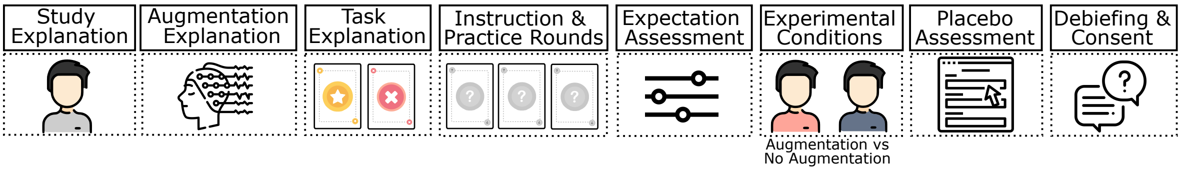


Figure 3.2: Study flow diagram: We conducted a within-subjects study. We induced a placebo effect by changing system descriptions. Participants took the Revised Hot Columbia Card Task (RCCT) twice to measure risk-taking. Participants in the AUGMENTATION condition were told that they would be helped by a cognitive augmentation. In the NO-AUGMENTATION condition, participants were told the augmentation system was off and no benefits existed. Finally, we informed them about the actual purpose of the study.

We collected the participants' age, occupation, identity, and gender information, as well as seven-point Likert scale ratings of their technical competence.

The researcher then described in full the notion of human augmentation, cognitive augmentation, and the apparatus. Colzato et al., Møller and Pedersen, Clements-Cortés et al. [181, 182, 183] works were specifically cited as evidence that the outlined augmentation is functional. However, the augmentation used in our study was a placebo. It was non-functional and did not improve the participants' cognitive abilities. See Figure 3.2 for an overview.

The induction of placebos adheres to a typical medical research process, see [184]. Participants received a stimulus consisting of an augmentation system that reportedly enhances human skills and a verbal description. The system was presented as a functional EEG-based human augmentation system that analyzes electric potentials in the brain and boosts performance by playing inaudible sounds to improve cognitive skills, even though no sound is actually created. This design integrates past studies demonstrating that the sound of musical compositions may improve performance through placebo effects [185] and that EEG caps can be utilized as a placebo [186]. The function of the augmentation device in the RCCT was presented as enabling participants to follow the movement of the quickly shuffling cards so they could determine the location of the loss cards, see Figure 3.2.

The following is an excerpt of the explanation provided to the participants:

> *We tune the audio to high and low frequencies that cannot be actively perceived to minimize listener fatigue and distraction from the sounds. For this purpose, the hearing threshold, loudness at which sounds are just heard, is measured. An artificial intelligence (AI) evaluates brain activity during the experiment and dynamically adjusts the binaural tones accordingly. The resulting feedback cycle ensures that the AI optimally adjusts the signal for maximum augmentation and thus maximum performance. In this study, we now want to evaluate whether the system enhances performance and compare this to a control condition without cognitive augmentation by AI.*

After describing the augmentation to the participants, we questioned them on their comprehension of the experiment, the augmentation, and its informed purpose (see supplementary material: `https://osf.io/gex4t/?view_only=4ed5d5ee9eab40069b908651b6d96a24`). This included three questions: *What are the two conditions you will test in this study?*, *How does the augmentation work?*, and, *What are the measured metrics used for?*. Each item had three possible response options, but only one was correct. All participants included in the study answered these questions correctly.

The RCCT was explained once the experimenter checked that the individual understood these points. Participants were informed that their remuneration would depend on their success in each game condition. Thus, they would receive 2.50 euro at the beginning of each card game. The worst scenario would result in 0 euros, while the best outcome would result in 10 euros. They were informed that the actual payout amount would be determined by the number of points obtained at the completion of each condition. At the end of the experiment, all participants were compensated with 5 euro per half hour.

The participants then played two rounds of guided instruction to familiarize themselves with the task. The system guided them through the first round by displaying win and loss cards, and the second round instructed them on how to use the rest of the interface (see Figure 3.3). After the two instruction rounds, we had the participants play two practice rounds, one of which was intentionally manipulated to demonstrate the risk of flipping loss cards. Following this explanation and prior to the actual experiment, we did an assessment of performance expectations prior to the RCCT.

Participants underwent a standard auditory threshold detection task across different frequency bands. Thresholds were not of interest in the study but were used to strengthen the placebo system's narrative of the verbal description. Then, depending on the condition (i.e., augmentation or no-augmentation), the participant either receives a pop-up stating that the augmentation is inactive and the game begins, or they are presented with a loading screen where they must wait two minutes until the system allegedly begins generating the inaudible sounds to augment them. After this delay, a message would appear confirming that the augmentation is now active, and then participants would finally be able to play the game. Throughout each condition of the RCCT, we recorded the number of cards flipped and the type of cards flipped. We simultaneously collected EEG data. The conditions were counterbalanced to avoid order effects.

Then we assessed task load and game experience after each condition. After completing both conditions, we measured participant judgments of improvement. Once participants had completed all questionnaires, we examined the usability of the augmentation technology, and, finally, debriefed them on the details of the experiments. Then, we measured user judgment of improvement and how they persisted after interaction. After debriefing participants we asked participants if they consented to the use of the collected data once they were fully informed regarding the purpose of the study. The experimenter did not know what their decision was and their decision did not affect their compensation.

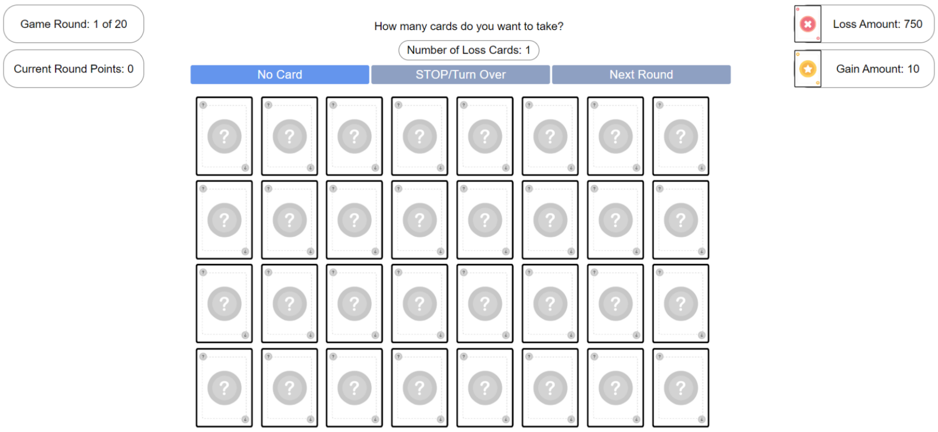


Figure 3.3: RCCT interface: The interface was a deck of cards with five indicators: current round, number of points, loss and win card values, and number of loss cards. The interface permits players to skip and stop rounds.

3.1.4 Measures

Assessment of judgments of performance

We measured user judgments of performance and how they persisted after interaction. For performance expectations (judgments prior interaction) we used three questions: First, a seven-point Likert item with anchors 1: Strongly disagree, and, 7 Strongly agree compared the expected performance between both conditions: *"I think I will do better in the augmentation condition as compared to the no augmentation condition"*. Then, two slider questions from zero to 930 (the theoretically possible maximum points if the game was not rigged) asking participants the expected number of points in each condition *"How many points do you think you will get in the no augmentation condition in the game?"*, and *"How many points do you think you will get in the augmentation condition in the game?"*. For judgments of improvement after interaction, we asked participants to rate Table 3.1 Likert items with anchors 1: Strongly disagree, and, 7 Strongly agree, after completing both conditions.

Risk-taking behavior

We applied the Columbia card task [130, 187] (hot version) to assess risk-taking behavior. In the CCT, risk-taking is operationalized by the total number of cards flipped in a round by the participants under a set of factors that modulate the risk of flipping a loss card.

There are a predefined number of loss cards (1 vs. 3) in the deck, each of which is equipped with a win (10 vs. 30) and loss point (250 vs. 750). Participants are instructed to flip as many cards as they dare to.

To maintain the task's credibility and prevent participants from simply flipping every card in the deck, seven rounds of each condition game are loss rigged, thus predetermined to result in a loss. These rounds are selected at random. The Columbia card task has been shown to correlate with affective decision-making [121, 130], risk behavior in adolescents [188], and other experimental measures of risk [187].

Two variants of the CCT exist (i.e., cold and hot) depending on the ability to interact with the cards in the game. In the hot version, the player must make incremental judgments (i.e., turn over one card at a time) and receive feedback after each decision. In the cold version, the player chooses the number of cards to turn over for the trial. We used the hot version of the CCT for two reasons. First, it measures bias in affective decision-making likely relevant for the use of augmentation under risk. Second, it could be perceived as less random as it allows participants to choose cards individually. The location of loss cards is not known to participants. This means that in the hot version of the CCT participants can pick cards from arbitrary locations until they encounter a loss-card while in the cold version the algorithm turns over cards sequentially from the beginning. To reiterate, in the cold CCT participants choose the number of cards to turn over while in the hot version of the CCT the participant chooses to flip individual cards.

The verbal description stimulus (DESCRIPTION) goal is to induce the participant's belief to be knowledgeable of the location of loss cards, i.e. the verbal description suggested they have an advantage in selecting the cards due to their enhanced information processing abilities. To allow for participants to know the location of loss cards and thus have an advantage in the game, we adapted the original CCT by showing the location of loss cards briefly using a card flipping animation and then shuffling the cards, referred in this manuscript as RCCT.

In the original CCT, the participant has no visibility of the win and loss cards, so the task depends on the participant's willingness to take risks based on the aforementioned factors (i.e., number of loss cards, amount of gain, amount of loss) that are displayed in the interface. For our narrative, however, we required a skill-based task that is subsequently executed more effectively due to cognitive enhancement. In detail, we implemented two changes (see Figure 3.4) to the CCT: Each round begins with the deck facing up (one second) so that the player can identify the winning and losing cards, and then the deck is flipped over and shuffled. We repeated the shuffling process five times. The cards are shuffled at an extremely rapid rate. One card could relocate from one side to the other in less than 480ms and its trajectory was shuffled five times before each round, preventing participants from determining the actual location of the cards. The last shuffle lasted 100ms to ensure it was not possible to follow the card location. Thereby preserving the element of risk in the actual task. As in the original CCT the location of loss cards was pre-determined, most rounds were rigged to be win rounds (13 of 20 rounds; 7 rigged-loss rounds). Thus, only the last or last three cards were loss cards.

Note also that, participants had to decide whether to flip over a card on a given location. Therefore, the augmentation that facilitated the processing of location information was described as giving them a relative advantage in the task. Note that implementing the same routine in the cold CCT would only yield an advantage to participants if they knew all locations of loss cards. Therefore, the hot CCT is better suited for our study.

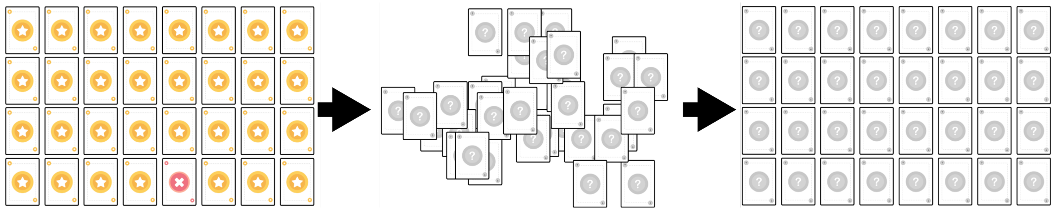


Figure 3.4: Revised Columbia Card Task: Each round begins with the deck facing up (one second) so the player can identify the winning and losing cards. The deck is then flipped over and shuffled at an extremely rapid rate and relocated in less than 300ms five times before each round, preventing participants from determining the actual location of the cards and preserving the element of risk in the actual task.

EEG recordings

Groot and Van Strien [187] showed that feedback evaluation following risky decision-making in the CCT was linked with feedback-related negativity (FRN) and a P300 in the EEG, where smaller FRN differences were associated with greater risk-taking and, impulsivity, with a decreased loss sensitivity, while smaller P300 differences were most strongly associated with greater reward responsiveness. Therefore we operationalize the processing of risk-related information through the Feedback Related Negativity and P300 in the EEG. For the recording of the EEG, we used an R-Net 64 channel EEG with a wireless amplifier (LiveAmp, Brain Products, Germany) and the corresponding recording software (Brain Vision Recorder) for electrode impedance calibration and the Brain Products LSL Streamer for signal streaming (R-Net, Brain Products, Germany). Electrodes were electrically connected to the scalp using a saline solution. The impedance of the electrodes was kept below 50kΩ (below the manufacturer's recommendations of 100kΩ). We utilized an average reference and a 500 Hz sampling rate to record the data. We have recorded data from 32 electrodes (**see Figure 8.4**).

Task load

We distributed a NASA-TLX task load [189] questionnaire to compare potential variations in task load generated by the stimulus. It is a widely used subjective assessment tool for evaluating task load. It measures task load by assessing six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Participants rate each dimension on a scale from 0 to 100, with higher scores indicating higher levels of task load. The NASA-TLX has been extensively validated and is considered a reliable and valid tool for measuring task load in various contexts.

3.1.5 Apparatus

We used Lab Streaming Layer (LSL) for time-series data acquisition. It was used for networking, time synchronization, and centralized data recording of the EEG streams and the RCCT annotations. We based our RCCT on a web-based CCT experiment provided by The

Experiment Factory Sochat [190] (Stanford, CA). The task was carried out using Microsoft Edge on a Windows (Windows 10 Version 21H2) desktop computer (HP Z1 G6) with an i7 (i7-10700) processor, 16GB of RAM and a screen size of 27 inches with a refresh rate of 60 Hz. Additionally, the web-based experiment was modified to transmit time annotations with the information of each button pressed (i.e., card flip or next round) to the lab streaming layer network and synchronize with EEG data. The participants used the mouse to select the cards in the RCCT. They were positioned in front of the screen, which was calibrated to their eyesight level. The distance between the participant's forehead and the screen was roughly 75 cm (29,5 inch).

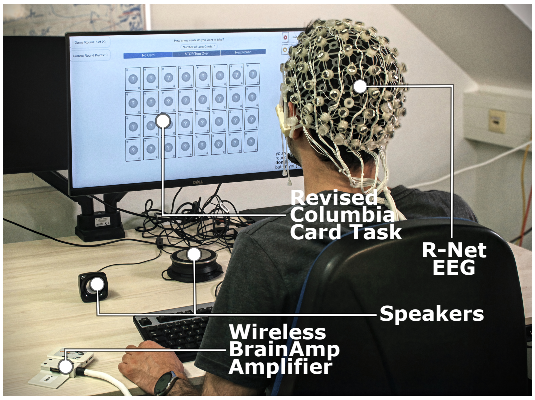


Figure 3.5: Microsoft Edge was used on a Windows (Windows 10 Version 21H2) desktop computer (HP Z1 G6) with an i7 (i7-10700) processor, 16GB of RAM, and a 27-inch screen with a refresh rate of 60Hz to complete the task. We used an R-Net 64-channel EEG with a wireless amplifier.

3.1.6 Data analysis

EEG Data Processing

To analyze the recorded data, we used the Python MNE library. The data was high pass filtered at 1 Hz and low pass filtered at 15 Hz [191, 192]. The data was then re-referenced to the average of all channels, which included the original reference electrode FCz. We applied a notch filter to remove the 50 Hz power-line noise. Then, we sliced the epochs into blocks of -0.3 ms and 0.7 ms, where 0.0 ms denotes the onset of the stimulus. We use the time between -0.3 ms and 0.0 ms as a baseline for the measured stimulus signal. We detected and rejected epochs likely to contain noise using the Autoreject library [193]. We automatically detected the local maximum around 300ms and 450ms to extract the P300 amplitudes for each epoch according to previous work [187].

Bayesian Data Analysis and Inference

For this paper, we use a Bayesian approach to data analysis. We used Bayesian linear mixed models (BLMM). The Bayesian approach has been taken up lately [194, 195, 196, 197, 198] as it presents several advantages to classical statistics. Kay, Nelson, and Hekler [194] explain the advantages of Bayesian statistics in technological contexts that are also relevant to our study. These are in particular: 1. The ability to use prior knowledge and learn from data. 2. To inform on the size of the placebo effect with a given level of precision; 3. It allows for the estimation of effects in small n-studies, 4. The approach enables readers to evaluate the effect size, which can also be close to zero, rather than the mere effect existence.

This Bayesian approach to modeling the CCT is frequently used [199, 200]. Following Weller et al. [199], we used censoring to model incomplete data distributions (e.g., rigged-loss trials of the CCT). The mean and standard deviation of the data distribution are reported without these censored trials. For a tutorial on Bayesian statistics, a description of the common workflow using brms, and reporting guidelines, see [201, 202, 203, 204]. Most importantly, the existence and the non-existence of a placebo effect are likewise important. The Bayesian approach to statistical inference allows us to measure the placebo effect and the non-existence of placebo effects on the measures.

Here, we use Bayesian parameter estimation, which allows us to estimate parameter values of effect sizes and quantify the uncertainty regarding these estimates based on the information in our data and the priors applied. We used brms [204], a wrapper for the STAN-sampler [205]. For statistical inference, we used R [206] along with packages for preprocessing [207, 208, 209, 210, 211], modeling [212, 213, 214, 215, 216, 217] and post-processing [218, 219, 220, 221, 222, 223, 224, 225, 226] the data. We computed 4 Hamilton-Monte-Carlo chains with 40000 iterations each and 10% warm-up samples. Trace plots of the Markov-chain Monte-Carlo permutations were inspected for divergent transitions. All Rubin-Gelman statistics [227] were well below 1.1, for effective sampling size.

We compare possible models with approximate leave-one-out cross-validation (LOOCV) [228]. This procedure allows us to compare information criteria across models. Relatively smaller LOOCV values indicate a better fit of the model to the data. The best model is then selected and parameters are further analyzed. For these, p_b was computed by calculating the relative proportion of posterior samples being zero or opposite to the median. This metric has similar properties to the classical p-value [229, 230, 231] but quantifies the proportion of probability that the effect is zero or opposite given the data observed. Note that this is the reverse of the classical approach to inferential statistics, where one measures the probability of the data given the null hypothesis with respect to the test statistic. Effects were considered meaningful when there was a particularly low probability (p_b <= 2.5%) of the effect being zero or the opposite. In addition to the median of the parameter, we calculated the High-Density Interval (HDI) at 95% of the posterior distribution for all parameters, which indicates the possible range of effects given the data, alongside the median of the respective parameter. Simple mean comparisons were done on standardized outcome variables. Therefore, all $\tilde{b}$ represent an effect size in terms of deviations of the standard deviations from the mean (corresponding to Cohen's D for simple effects of categorical predictors with two levels). For models on factorial designs, our analysis of the behavioral and physiological data, we calculated δ_t, which can be interpreted quite similar to Cohen's d and is based on standardizing the population-level effects on the varying-effects and residual variance [232, 233]. We explored the effect of different weakly informative priors on the data. None affected statistical inference. We also provide classical tests resembling Bayesian analysis for each step of inference and ordinal regression analysis for Likert-type questions in the supplementary material `https://osf.io/gex4t/?view_only=4ed5d5ee9eab40069b908651b6d96a24`.

For simple mean comparisons, priors were chosen to resemble only weakly informative priors when standardized with a prior on the standardized mean difference of (*M* = 0, *SD* = 1) and thus encompass positive and negative small to large effect sizes, d_z $HDI_{95\%}$ = [-.1.96, 1.96], centered at zero on the standardized outcome, for the intercept and the residual a t-distributed prior (*df* = 3, M = 0, SD = 1) was used and we specified a student-link function (ν following a γ distribution with p = .1, b = 2) to resemble the commonly used t-test with pooled variances.

3.2 Findings

We first report on the belief of participants that the system augmented them. Then, we analyze user judgment of improvement before and after the stimulus. Followed by modeling risk-taking behavior as a function of verbal description and judgment of improvement [184]. We follow this up with an analysis of feedback-related negativity in response to loss cards [187] for the EEG signal.

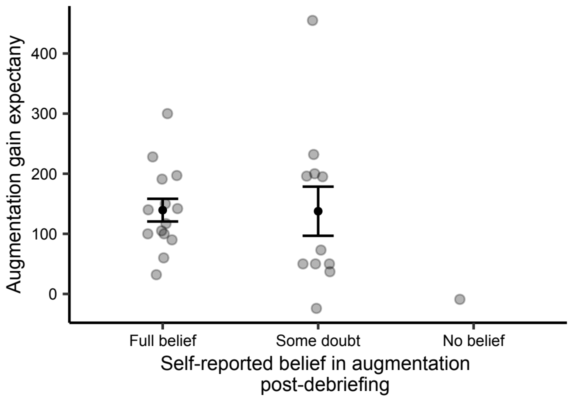


Figure 3.6: Mean expected augmentation gain in points for the RCCT with individual data points for each subject as a function of self-reported belief in the augmentation after debriefing. Error bars denote ±1 standard error of the mean.

3.2.1 Manipulation Check

After the experiment and debriefing participants about the deception and sham treatment, we asked them to indicate whether they believed in the functionality augmentation system or suspected that they were deceived. Only one out of 27 participants (3.70%) indicated that they did not believe in the system's capabilities. Eleven out of 27 (40.74%) participants reported some minor suspicion of the system's functionality (e.g., *P2: I believed that augmentation takes place, but that it really helps was skeptical. I was aware that the difference was more influenced by sequence, fatigue, and other factors."*). The majority of participants, 14 (51.85%), fully believed in the augmentation technology's effect. One participant did not disclose whether they believed in the DESCRIPTION or not (3.70%).

3.2.2 Impact of Verbal Description on Performance Expectations and Judgments of Performance (H1 & H2)

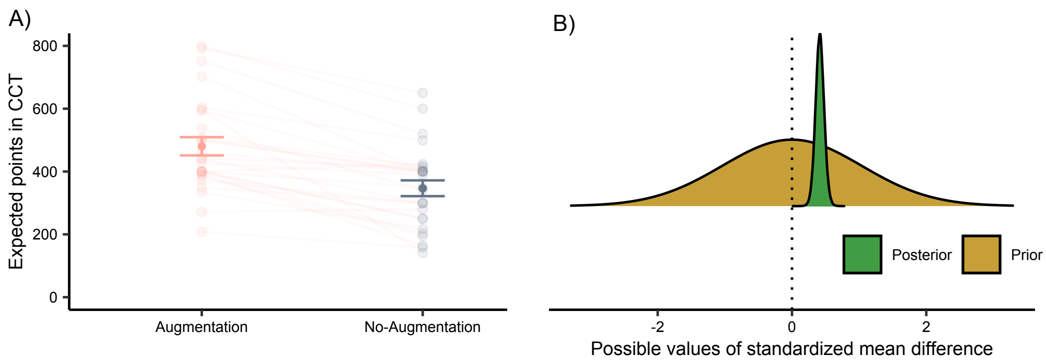


Figure 3.7: A: Mean expected points in the RCCT with connected individual mean values. Error bars denote ±1 standard error of the mean. B: Prior and posterior density plots. Prior samples are in beige, and posterior samples are in green. The relative density increase from prior to posterior shows how the data has informed the model. No posterior samples lie opposite of zero, indicating that the effect is unlikely to be opposite, or zero.

After the description of the experiment, the task, and the system being used but before interaction, we asked participants to indicate how many points they thought they would score with and without the augmentation on a scale ranging from 0 to 930 points. Participants indicated that for the AUGMENTATION condition (M = 480.30, SD = 150.91), they will score more points as compared to the NO-AUGMENTATION condition (M = 346.70, SD = 130.43). This difference could be distinguished from zero , $\tilde{b}_{\text{std}}$= 0.41 [0.29, 0.54], p_b = 0%, see Figure 3.7B. Figure 3.7A shows the mean for each condition and the substantial variation in participants. While some estimated their gain to be small, others considered it quite substantial. Therefore, hinting at the notion that the placebo effect is subject to high levels of individual variation, which is in line with Kosch et al. [184].

To inspect whether this variation corresponds to participants' reported judgment of improvement in the AUGMENTATION condition after use, we plotted the difference in expected points, further referred to as relative AUGMENTATION EXPECTANCY (expected points for AUGMENTATION - expected points for NO-AUGMENTATION) as a function of indication of belief on the system (manipulation check). One can see that while on average, there is no substantial difference between the full-belief group and the group of participants that reported some doubt $\tilde{b}_{\text{full/doubt}}$= -0.04 [-0.47, 0.42], p_b = 56.54%, (see Figure 3.6), the variation is larger in the group that reported some doubt; however, the difference in variance between groups was not distinguishable from zero, $\tilde{\sigma}_{\text{full/doubt}}$= 0.31 [-0.01, 0.65], p_b = 3.12%. Also noteworthy is that some participants that voiced minor doubts after the experiment were expecting no gain in points through the augmentation. This is also the case for the one participant who reported that they did not believe in the system at all after the experiment, see Figure 3.6.

As it could be argued that participants' lack of familiarity with the game mechanics could have impacted our results, we also asked participants to indicate their agreement to *"I think I will do better in the augmentation condition as compared to the no augmentation condition"* on a 7-point Likert scale with anchors 1:Strongly Disagree, and, 7: Strongly Agree. On average, participants reported agreeing with the item with M = 4.81 (SD = 1.47). We tested this mean against an expected value of 3 (which would indicate neither to agree or disagree with the statement), resembling a one-sample t-test. Here, we used a normally-distributed prior on the intercept centered at zero with a SD that was two times the standard deviation of the observed variable again with a studentized link-function (ν following a γ distribution with p = .1, b = 2) for the residuals. The sigma prior resembled the mean-comparison model and to allow for more variation a t-distributed prior (df = 3, M = 0, SD = 1).. The difference between the mean and the expected value of 3 was distinguishable from zero, $\tilde{b}_{\text{std}}$= 1.24 [0.83, 1.61], p_b = 0.00% . We also asked whether they still believed this after interaction with the system and experiencing the NO-AUGMENTATION condition. On average participants still believed in the augmentation, M = 4.44, SD = 1.67, $\tilde{b}_{\text{std}}$= 0.86 [0.46, 1.24], p_b = 0%) and when comparing their response before and after interaction there was no distinguishable reduction in confidence $\tilde{b}_{\text{std}}$= -0.12 [-0.37, 0.14], p_b = 17.29%. The $HDI_{95\%}$ was centered around zero with a maximum effect of 0.37 SD on the outcome variable. Therefore, the belief of superior performance for the AUGMENTATION condition was sustained after the interaction, which generated the placebo effect [184].

This placebo effect is also exemplified in the post-experimental questionnaire (see Table 3.1). We found that participants, on average, judged the augmentation system to facilitate task completion and improve performance and cognitive abilities. This has also prompted participants to conclude that this augmentation has potential for future development, again see Table 3.1.

3.2.3 Influence of Performance Expectations on Risk-taking Behavior (H3)

Participants each played 40 rounds of the game. These were sampled from combinations of 2 (1 vs. 3 LOSS CARDS) $\times$ 2 (250 vs. 750 points LOSS AMOUNT) $\times$ 2 (10 vs. 30 points WIN AMOUNT) for each condition of DESCRIPTION. Note that the mixed model approach, we use for analysis does not require equal distribution of trials across experimental variations. For 27 participants, this resulted in 1080 data points that indicated risk-taking as the number of cards turned over in the RCCT. We used censoring for rigged loss rounds to model the whole game in line with Weller et al. [199]. Censoring takes into account that the number of cards in loss rounds only represents a minimum but otherwise unknown estimate of the number of cards the participant would have turned over.

Priors and Model Selection

For multilevel-data and trial-based modelling of the RCCT, we applied normally distributed priors (M = 0, SD = 10) on all population-level effects, with Cholesky priors on the unstructured (residual) correlation (η = 2), and a t-distributed prior (df = 3, M = 0, SD = 5) on the intercept, sigma and the variance, with a normally-distributed prior on the intercept parameters (M = 20, SD = 10). Two-way interactions in our model were followed up by posterior predictive plots, which serve a similar purpose as post-hoc comparisons in classical statistical inference. We used effect-coding on categorical variables (e.g., 1, -1).

We modeled the effect of the stimulus using a varying intercept for every participant to account for the repeated-measures structure of the data in the mixed model. To allow for individual variation of effects in participants, we added cross-varying slopes for interaction terms for LOSS AMOUNT, WIN AMOUNT, and LOSS CARDS for every subject. The varying intercepts and varying slopes for each participant serve the purpose of normalization and thus control for systematic individual differences in the dependent variable (e.g., individual differences in loss aversion). All population-level effects of LOSS CARDS, LOSS AMOUNT and WIN AMOUNT, were matched with an interaction term of DESCRIPTION and AUGMENTATION EXPECTANCY (See supplementary material for the full model specification. We compared a null model that only estimated the intercept and the mean (LOO = 4732.99) with a model that accounted for LOSS CARDS, loss amount and win amount with population-level effects and varying-level effects (LOO = 4269.19) similar to Weller et al. [199] with the LOOCV information criterion and then subsequently added main-effects and fully crossed interaction terms for the DESCRIPTION (LOO = 4218.09) and AUGMENTATION EXPECTANCY (LOO = 4225.04). We selected the most complex model with both DESCRIPTION and AUGMENTATION EXPECTANCY as it allows us to quantify the effect of individual AUGMENTATION EXPECTANCY while providing

Table 3.1: Items were answered on a 7-point likert scale(1 - strongly disagree; 7 - strongly agree). We tested against an indecisive value of 3. Effects that are distinguishable from zero are marked with *. We did not test the SUS against a hypothesized value.

Item/scale	M	SD	$\tilde{b}_{\text{std}}$	$HDI_{95\%}$	p_b
*The game was easy to play.**	3.93	1.86	0.47	[0.07, 0.86]	1.07%
*The cognitive augmentation has made the task easier.**	3.74	1.40	0.50	[0.12, 0.88]	0.59%
The cognitive augmentation has made the task more enjoyable.	3.56	1.93	0.28	[-0.12, 0.68]	8.24%
The cognitive augmentation has made me more confident.	3.67	1.86	0.35	[-0.05, 0.74]	4.35%
The cognitive augmentation has made me more efficient.	3.52	1.50	0.33	[-0.06, 0.72]	4.63%
*The cognitive augmentation has improved my performance.**	4.11	1.55	0.70	[0.31, 1.09]	0.06%
*The cognitive augmentation has improved my cognitive abilities.**	3.93	1.54	0.59	[0.21, 0.96]	0.20%
*The cognitive augmentation in this game has a lot of potential for future development.**	4.37	1.55	0.85	[0.45, 1.24]	0.00%
System usability scale	56.94	11.34	-	-	-

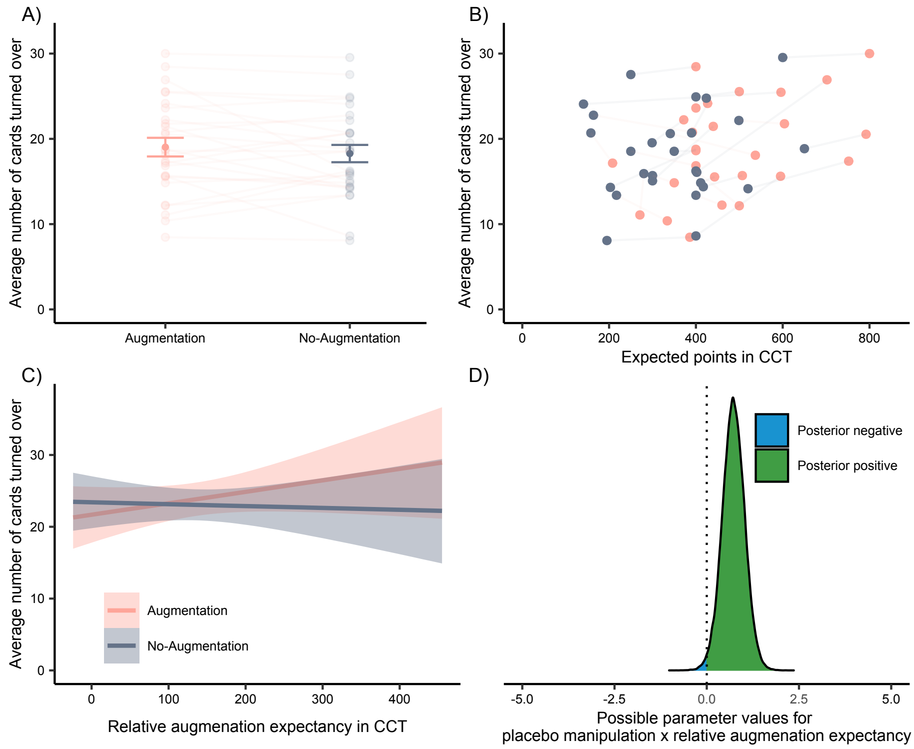


Figure 3.8: A: Average number of cards turned over in the RCCT with connected individual mean values. Error bars denote a +/-1 standard error of the mean. B: Average number of cards turned over in the RCCT for each participant as a function of expected points in the RCCT. C: Predicted average number of cards turned over in the RCCT by our model as a relative augmentation gain (Augmentation-No augmentation). D: Posterior density plot. The blue indicates the proportion of posterior samples opposite to the median and thus is a visual representation of the posterior p-value. It quantifies the proportion of probability that the effect is zero or opposite given the data observed. The smaller the blue areas in comparison to the green areas are, the more reliable is the estimation of the effect. We omitted to display the prior distribution as it would appear flat given the wide SD when it is, in fact, normally distributed.

a fit indistinguishable from the more parsimonious model. For the sake of brevity, we will analyze the posterior only for this final model.

Posterior Distribution Analysis

As is typically the case for the CCT, our model could show that participants considered the number of LOSS CARDS when making their decision, $\tilde{b}_{\text{loss cards}}$= 3.90 [3.12, 4.69], p_b = 0.00%, $\tilde{\delta}_{\text{b}}$= 0.78 [0.61, 0.94]. They turned over relatively fewer cards (M = 14.49, SD = 6.35) when there were three LOSS CARDS in the deck as compared to the conditions when there was one LOSS CARDS in the deck (M = 22.61, SD = 4.73). There was also an effect of LOSS AMOUNT, $\tilde{b}_{\text{loss amount}}$=0.60 [0.19, 1.01], p_b = 0.35%, $\tilde{\delta}_{\text{b}}$= 0.12 [0.04, 0.20]. With more cards turned in games with 250 points loss possibility (M = 19.28, SD = 4.83) as compared to 750 points losses (M =

17.96, SD = 5.81). The WIN AMOUNT did not affect participant's decision to turn over cards, $\tilde{b}_{\text{win amount}}$= 0.12 [-0.42, 0.64], p_b = 31.53%, $\tilde{\delta}_{\text{b}}$= 0.02 [-0.08, 0.13]. Therefore, our data of the RCCT is in line with other psychological studies using the CCT[199, 200].

We did not find any direct effect of the DESCRIPTION on risk-taking, $\tilde{b}_{\text{description}}$= 0.42 [-0.19, 1.01] , p_b = 8.07%, $\tilde{\delta}_{\text{b}}$= 0.08 [-0.04, 0.20]. The HDI indicates that any difference between conditions is smaller than 1 and can therefore be neglected. This lack of a substantial effect, was probably due to the high level of variation in the placebo effect, see Figure 3.8A. However, we found that relative AUGMENTATION EXPECTANCY (see Figure 3.8B), increased the number of cards chosen in the AUGMENTATION condition, $\tilde{b}_{\text{description} \times \text{augmentation expectancy}}$= 0.72 [0.12, 1.32], p_b = 1.04%, $\tilde{\delta}_{\text{b}}$= 0.15 [0.02, 0.26], see also Figure 3.8D. The more participants expected to gain from the AUGMENTATION in the game, the more risks they took when expecting to be augmented, see also Figure Figure 3.8C. The direct placebo effect term, as well as the interaction effect of DESCRIPTION × AUGMENTATION EXPECTANCY, were not qualified by any interaction with the factors LOSS AMOUNT, WIN AMOUNT, or LOSS CARDS, all effects centered around zero with p_b > 15.94%. The Bayesian analysis can thus show that relative AUGMENTATION EXPECTANCY is a necessary condition for risk-taking during interaction.

3.2.4 Task Load

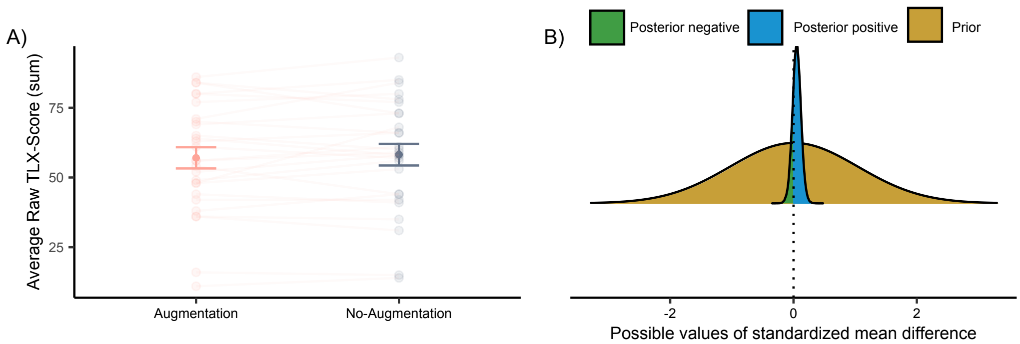


Figure 3.9: A: Average TLX sum score with connected individual mean values. Error bars denote ±1 standard error of the mean. B: Prior and posterior density plots. Prior samples in beige, and posterior samples in green. The relative density increase from prior to posterior shows how the data has informed the model. Posterior samples are centered at zero, indicating that the effect is likely to be small, or zero.

We compared the average NASA TLX Raw sum score across DESCRIPTION. There was no significant difference between conditions, $\tilde{b}_{\text{std}}$= -0.03 [-0.09, 0.04], p_b = 21.96%. Looking closely at the posterior distribution of the mean difference (Figure 3.9) and taking into account the $HDI_{95\%}$, it is highly unlikely that the AUGMENTATION condition produced any kind of increased subjective workload in the TLX. The $HDI_{95\%}$ indicates that any difference would be smaller than around 1/10 of a point on the sum-score. We can follow that the effect of the DESCRIPTION on the TLX is negligible and not distinguishable from a null-effect. We also found no effect on any of the TLX-subscales, all p_b > 5.23%.

3.2.5 Influence of Performance Expectations on Processing Risk-related Information (H4)

Priors and Model selection

One participant had to be discarded from the dataset due to corrupted data in the recordings. Data from 26 participants were left for the EEG data analysis. We modeled the EEG separately regarding the amplitude of the FRN and the P300.

For multilevel-data and average-based analysis of the P300 and FRN amplitudes in the EEG, we applied normally-distributed priors (*M* = 0, *SD* = 10) on all population-level effects and varying-level effects, and normally-distributed prior (*M* = 0, *SD* = 20) on the intercept. σ was modeled with a t-distributed prior (*df* = 3, M = 0, SD = 5) and the student-link function with ν following a γ distribution with p = .1, b = 2. Two-way interactions in our model were followed up by posterior predictive plots, which serve a similar purpose as post-hoc comparisons in classical statistical inference. We used effect-coding on categorical variables (e.g., 1, -1).

To allow for individual variation of win/loss card effects in subjects, we added a varying slope for every subject. The population-level effects of DESCRIPTION, AUGMENTATION EXPECTANCY and win/loss cards were fully crossed (For the full model specification, see supplementary material). As event-related EEG data is prone to outliers, we used a student link function (The deviation of normality was due to the heavy tails of the distribution. For a histogram and a Shapiro-Wilk test). For model selection, we compared a null model that only estimated the intercept, varying slopes, and the mean (LOOFRN = 512.26, LOO$P300$ = 616.88) with a model that accounted for win/loss cards as population-level effect (LOOFRN = 508.06, LOO$P300$ = 611.16), and then subsequently added main-effects and fully crossed interaction terms for the DESCRIPTION (LOOFRN = 515.95, LOO$P300$ = 618.24) and AUGMENTATION EXPECTANCY (LOOFRN = 521.53, LOO$P300$ = 608.11). For the FRN, the best fit was the NULL model. The LOO information criteria, therefore, suggest that none of the modeled population-level effects had any influence on the amplitude of the FRN. For the P300, the most complex model with all population-level effects had the best fit to the data. We will thus only analyze the posterior of this P300 model.

Posterior distribution analysis

DESCRIPTION affected the strength of the P300, $\tilde{b}_{\text{description}}$= -0.76 [-1.38, -0.08], p_b = 1.23% , $\tilde{\delta_{\text{b}}} = -.24$ [-.47, -0.02]. Participants had higher P300 amplitudes in the NO-AUGMENTATION condition M = -0.19, SD = 5.42 as compared to the AUGMENTATION condition M = -0.46, SD = 5.24. We also found a distinguishable effect of AUGMENTATION EXPECTANCY, $\tilde{b}_{\text{expectancy}}$= 1.42 [0.41, 2.38], p_b = 0.53%, $\tilde{\delta}_{\text{b}}$= 0.46 [0.12, 0.82]. With every 10 points of relative AUGMENTATION EXPECTANCY, the amplitude of the P300 increases by about 0.14. There was no distinguishable main effect of win/loss cards on the P300 amplitude $\tilde{b}_{\text{win/loss}}$ = -1.05 [-2.13, 0.05] , p_b =2.97%.

The DESCRIPTION $\times$ win/loss trials interaction was distinguishable from zero, $\tilde{b}_{\text{description} \times \text{win/loss}}$= -0.64 [-1.23, -0.01], p_b = 2.31%, $\tilde{\delta}_{\text{b}}$= -0.21 [-0.42, 0.00], for posterior predictive plot see Figure 3.10A. Likewise, a relative AUGMENTATION EXPECTANCY $\times$ win/loss trials interaction was distinguishable from zero, $\tilde{b}_{\text{expectancy} \times \text{win/loss}}$= 1.31 [0.18, 2.39], p_b = 1.30%,$\tilde{\delta}_{\text{b}}$= 0.43 [0.05, 0.81] (see Figure 3.10B), as well as a DESCRIPTION $\times$ relative AUGMENTATION EXPECTANCY interaction, $\tilde{b}_{\text{description} \times \text{expectancy}}$ = -0.90 [-1.45, -0.21], p_b = 0.85%, $\tilde{\delta}_{\text{b}}$= -0.31 [-0.51, -0.09] (see Figure 3.11A).

Note that these two-way interactions were driven by a the three-way interaction, $\tilde{b}_{\text{description} \times \text{expectancy} \times \text{win/loss}}$ = -0.97 [-1.48, -0.32] , p_b =0.42% . To grasp the model estimates and the interaction effects, we compare the raw data to the model predictions Figure 3.12. One can see that the P300 only increased with AUGMENTATION EXPECTANCY for the NO-AUGMENTATION condition in loss trials; for win cards and loss cards in the NO-AUGMENTATION condition, this correlation was not present. We can thus follow that heightened AUGMENTATION EXPECTANCY is associated with a decreased P300 response for loss trials.

3.3 Discussion

Our study investigated the placebo effect of augmentation technologies and their consequences for risk-taking. We replicated prior research on placebo effects in technology evaluation inducing expectations with a verbal description of an augmentation technology (H1) and after using the sham augmentation technology, participants maintained their judgment of improvement (H2). Consequently, using augmentation technologies results in an inherent perception of improvement in the subject, a placebo effect. While we have not found a direct effect of the placebo on risk-taking, our Bayesian analysis demonstrates that an expectation of improvement is required for increased risk-taking when being told to be augmented (H3). The P300, which typically occurs in the RCCT for loss trials [187], was lowered when anticipating support from the augmentation compared to the NO-AUGMENTATION condition (H4).

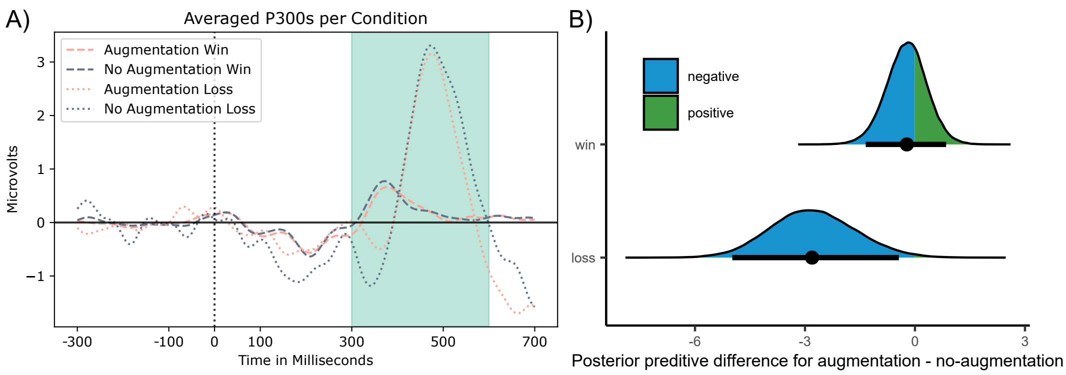


Figure 3.10: A: ERP averaged across the central region (Fz, Cz, Pz, Oz). There is a significant decrease in the P300 amplitudes for DESCRIPTION between loss/win -trials for the AUGMENTATION condition. B: Posterior predictive plot for the DESCRIPTION $\times$ win/loss trials interaction.

3.3.1 Experiencing Benefits from a Sham augmentation technology

The placebo effect of augmentation technologies extends previous studies on placebo effects that focused on improvement after treatment in medical research and psychology [234, 235, 236, 237, 238] but also in technology evaluation [184]. In our study, a mere expectation of improvement changed the user's risk-taking, and their expectation of improvement was sustained after use. Particularly interesting is that, in contrast to Kosch et al. [184], not only the joint performance with the assistance system was increased, but the users' very own capabilities were expected to be improved. Mapping our results onto theories of human-computer integration [46], our study can assert that perceived human-system capabilities may be judged in the absence of probing system functionality. In this domain of research, our methodology of employing a placebo augmentation technology could be used to study how human–system integration affects the users' decision-making. Note, however, that for the use of placebo for research purposes, the mechanisms [184] and contextual variables [239] in the placebo effect of augmentation technologies need to be examined more closely.

3.3.2 Taking Risks with Augmentation Technologies

Augmentation technologies are mediators of interaction with the real world. Our findings indicate that a belief of being augmented, in conjunction with the user's expectations regarding the augmentation technology's performance, is sufficient to modify the user's risk-taking behavior. This must be examined from two standpoints. Firstly, it could be that users pose a risk to themselves. Secondly, the user could engage in risky behavior and endanger those around him. This may be exaggerated in situations where enhancements support in interacting with environments that pose conditions that can not be met with the users' capabilities alone, but only when augmented, e.g., [22, 67, 117].

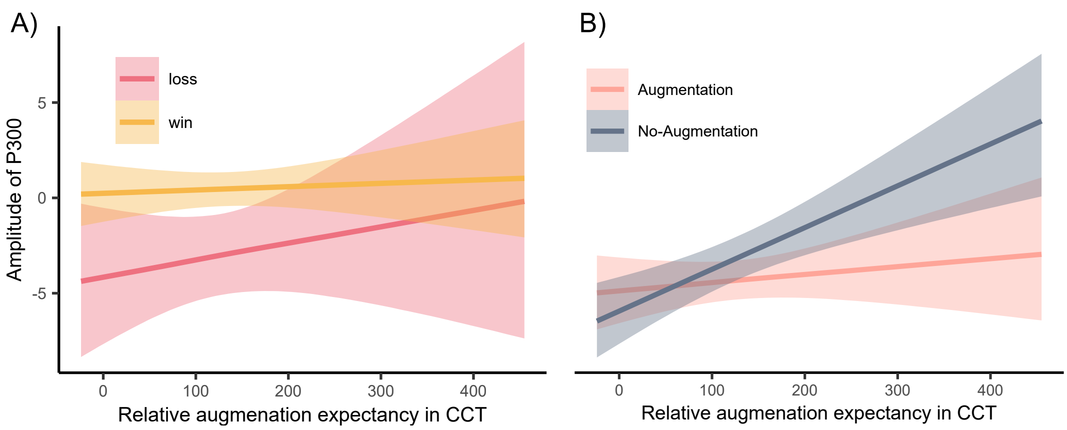


Figure 3.11: A: Posterior predictive plot for the win/loss trials × relative AUGMENTATION EXPECTANCY interaction. B: Posterior predictive plot for the DESCRIPTION × relative AUGMENTATION EXPECTANCY interaction.

Our findings suggest that in these situations, decision-making will be biased in favor of riskier options that match the subjective capabilities of augmentation rather than the objective capabilities of the augmentation technology user [117]. An immediate possibility to prevent placebo effects from promoting risky decision-making would be to support the user in building appropriate mental models about the augmentation technology, e.g. by training them to know about the constraints and limitations of the augmentation technology. A more advanced strategy would be to support the user in an appropriate control. Here, one could give feedback to the user that human-system capabilities are not enough to meet the user's expectations and therefore foster risk-averse decisions. For this, users' expectations in a given context could be measured verbally (i.e., by polling expectations), extracted from simulated behavior as in the RCCT, or based on physiological sensing (e.g., comparing the amplitude of the P300 for expected and non-expected events). These levels of information could be integrated and presented to the user in an open-loop system. In a closed-loop system, the level of support could be mapped onto expectations in low-risk situations to calibrate the user's mental model, e.g., less support by an exoskeleton when carrying an object that is not too heavy for the user without augmentation. Overall, our study can highlight that decision-

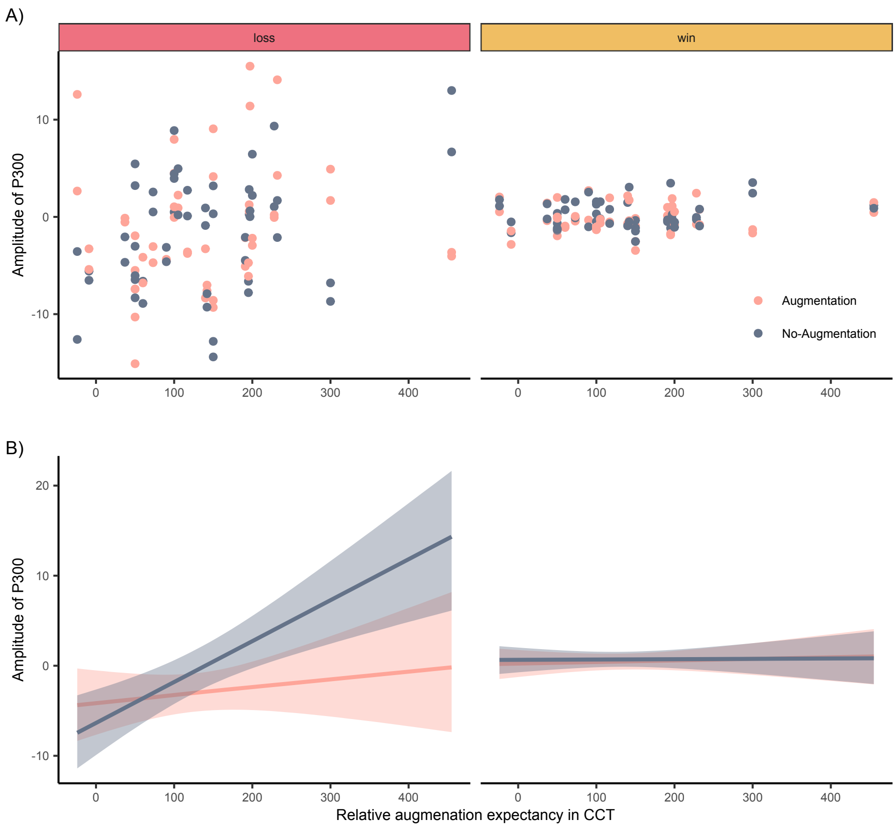


Figure 3.12: A: Average P300 for each participant as a function of AUGMENTATION EXPECTANCY and win/loss cards. B: Predicted P300 by our model contrasting AUGMENTATION EXPECTANCY (Augmentation- No augmentation) and win/loss cards.

making under uncertainty needs to be taken into account when designing augmentation technologies, irrespective of the actual human-augmentation technology capabilities.

3.3.3 P300 as a correlate of Risk Processing for Augmentation Technologies

We observed greater P300 amplitudes in the absence of augmentation than in the presence of augmentation for loss-trials. In the context of our study, a reduction in the P300 for loss trials when in the AUGMENTATION condition as compared to the NO-AUGMENTATION condition could have two concurring explanations. First, Gray et al. [240] postulate that in decision-making contexts, the P300 indexes the self-relevance of events. Concerning our study, a reduced P300 for the AUGMENTATION condition could index that non-functional human-augmentation technology interactions are processed as less self-relevant. Secondly, one can argue that this was only due to a difference in brain-related potentials caused by perceived ambiguity in decision-making. Previous research by Wang et al. [241] shows that the P300 amplitude is attenuated in ambiguous situations of risk-taking and less attenuated when there is less ambiguity concerning outcomes. In our study, the AUGMENTATION condition represents less ambiguity as compared to the NO-AUGMENTATION condition because participants subjectively experienced more control over the outcomes of their decisions, i.e., an advantage in knowing where the loss cards are. However, looking closely at our results, the reduced P300 was only found in loss trials and not in win trials and not only as a main effect. Thus, it is likely that self-relevance, as posed by Gray et al. [240] can explain the pattern in our data. Information about loss trials was not preferably processed as self-relevant when being augmented.

3.3.4 Effects of Augmentation Technologies on Information Processing in Augmented Individuals

While previous research has suggested that Augmentation Technologies may impact self-perception and behavior [46], empirical evidence has been lacking until now; Our results show a notable change in P300 amplitude based on expectancy of augmentation, which may be explained by self-relevance; this finding raises questions about how people process information in tasks performed with augmentation technology support. This highlights the significance of developing more effective augmentation technologies, given their potential impact on decision-making, as well as the importance of further investigating decision-making when using augmentation technologies.

3.3.5 Implications for Motor and Sensory augmentations

Expectations regarding the perception of external events are known as stimulus expectancy, whereas expectations regarding our own involuntary reactions to events are known as response expectancy [242]. An example of a response expectancy would be the belief that a sugar pill will improve response time. In contrast, a placebo that improves target detection concentration could be considered a stimulus anticipation. While both expectancy mechanisms to placebo-effects have been studied in the medical domain, it has not yet been determined how these mechanisms contribute to the evaluation of AI augmentation technologies.

For example, Kosch et al. [184] employed a response expectancy framing technique, informing participants that the task would be easier to complete. However, augmentation technologies can also generate stimulus expectancy. A placebo in sensory augmentation would be considered a stimulus expectancy, whereas a placebo in motor and cognitive augmentation would be a response expectancy. While response expectancies are considered more stable and robust in producing placebo effects, stimulus expectancies rely on the ambiguity of the stimulus [242]. As placebo effects for stimulus expectancy can be modulated by stimulus ambiguity and are typically weaker than response expectancies in terms of the placebo effect, future research should investigate whether the likelihood of placebo effects varies between augmentation approaches, i.e., sensory, motor, and cognitive.

3.3.6 Generalizability to other Technological Contexts

Our study has examined the contextual factors related to cognitive augmentation technologies, which are emerging and highly anticipated technologies. Due to the limited understanding of this technology and the external narratives surrounding it, users may develop high expectations of its capabilities [243]. Similarly, over-hyped technologies such as AI have been found to induce placebo effects and affect user performance. Hence, it can be argued that expectations of technologies are central in the judgment of their performance, thus emphasizing the significance of user perception of the technology over its form factor (which was embodied in our study but was desktop-based in [184]. Thus, researchers should consider controlling for users' expectations of the technologies under investigation to prevent potential biases in evaluations and alterations in user behavior. This, for example, implies that models such as the Technology Acceptance Model [244] must account for user expectations.

Currently, the placebo literature in Medicine and Psychology emphasizes the role of physical artifacts (e.g. pills) or psychological treatments [245, 246]. However, placebo effects can be found for game elements, e.g., power-ups, [247, 248], control modules in user interfaces [249], or when being supported by AI [184]. Therefore, our study supports the hypothesis that verbal descriptions of digital-technological artifacts can serve as a placebo.

Does the Belief of Being Augmented Also Impacts the Sense of Agency?

We performed a systematic literature review following the guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) [250]. Then, we re-

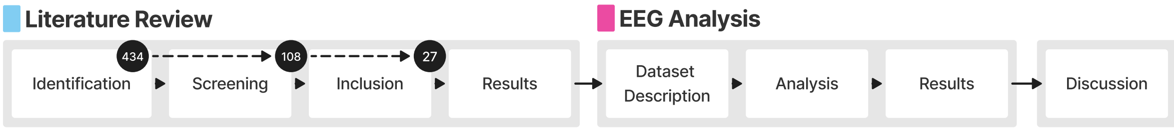


Figure 3.13: Methodology Overview. We conducted a literature review to identify research gaps regarding cognitive agency. Afterward, we conducted a user study to investigate the subjectively perceived impact on cognitive user agency when interacting with an AI-driven assistant.

analyzed the study presented in chapter 3 to understand the impact with and without perceived AI assistance [Aut1] on cognitive agency. Although previous work showed that EEG is a validated and reliable measure for studying the sense of agency [139, 141, 251], EEG has not yet been extensively evaluated in the context of cognitive augmentation. Thus, the analysis investigates how EEG metrics, such as alpha and beta band activity, could form an alternative metric for quantifying agency in the context of interactive systems. The following sections describe the review's search strategy and selection process, re-analysis, and conclusions.

3.4 Search Strategy

The ACM Digital Library and IEEE Xplore electronic databases were searched for primary HCI studies on cognitive augmentation, which also measure or consider the sense of agency. On the ACM Digital Library, the content type "Research Articles" was selected. On IEEE Xplore, publications were filtered for "Journals" and "Conferences". Records from 2003 until 2023 were considered.

To allow for the reproducibility of search results, the exact search queries can be found in the in Table 3.2. After testing various keywords and combinations, we ultimately selected one specific search query that combined the terms "sense of agency" with "cognitive augmentation", "augmented cognition", or "human augmentation", allowing us to obtain articles that use the established terms. Additionally, since human augmentation is not a well-defined term yet, we decided also to include a more descriptive query (see Query 2 in Table 3.2) which combined: (i) "sense of agency", (ii) "augment*" or several synonyms, (iii) "cogniti*" or "brain", and (iv) "human-computer interaction", aiming to capture all relevant HCI articles, even if they do not mention the term "augmentation". The "human-computer interaction" keyword was added to filter out records that lacked relevance to HCI research, as, without this addition, the volume of records was not manageable within our time constraints.

Through this search process, a total 445 records were identified from the ACM Digital Library (n = 378) and IEEE Xplore (n = 67). Eight duplicate records were removed, which were caused by utilizing the two different search queries. Three records had to be removed because they were not research articles. In total, this resulted in 434 publications to be screened.

Table 3.2: Databases, search queries and number of identified records

Database	Search query	Records
ACM Digital Library	*Query 1:* [All: "sense of agency"] AND [[All: "cognitive augmentation"] OR [All: "augmented cognition"] OR [All: "human augmentation"]]	10
	Query 2: [All: "sense of agency"] AND [[All: augment*] OR [All: improve*] OR [All: extend*] OR [All: enhance*]] AND [[All: cogniti*] OR [All: brain]] AND [All: "human-computer interaction"]	368
IEEE Xplore	*Query 1:* ("Full Text & Metadata":"sense of agency") AND (("Full Text & Metadata":"human augmentation") OR ("Full Text & Metadata":"cognitive augmentation") OR ("Full Text & Metadata":"augmented cognition"))	3
	Query 2: ("Full Text & Metadata":"sense of agency") AND (("Full Text & Metadata":augment*) OR ("Full Text & Metadata":improve*) OR ("Full Text & Metadata":extend*) OR ("Full Text & Metadata":enhance*)) AND (("Full Text & Metadata":cogniti*) OR ("Full Text & Metadata":brain)) AND ("Full Text & Metadata":"human-computer interaction)	64
		445

3.4.1 Selection Criteria and Process

We checked all identified publications (n = 434) against the following initial inclusion criteria:

1. Peer-reviewed, original work (excluding literature reviews)
2. Written in English
3. Sense of agency measured or considered
4. Focus on cognitive augmentation (i.e., augmentation technologies that aim to enhance cognition)

We did not restrict the selection to a specific type of research method, allowing both quantitative and qualitative research. In the reference management and knowledge organization program Citavi [252], version 6.7, we imported all found records and documented the respective database and search query. Then, one author checked all publications for eligibility one by one whilst recording the eligibility decision and reason for exclusion, where applicable.

The screening process, which is visualized in Figure 3.14, entailed three phases, Identification, Screening, and Inclusion. The first screening phase encompassed reading all titles and abstracts, which excluded 326 records. If it was unclear whether a publication met all of the inclusion criteria based on the title and the abstract, we included it in the second screening phase.

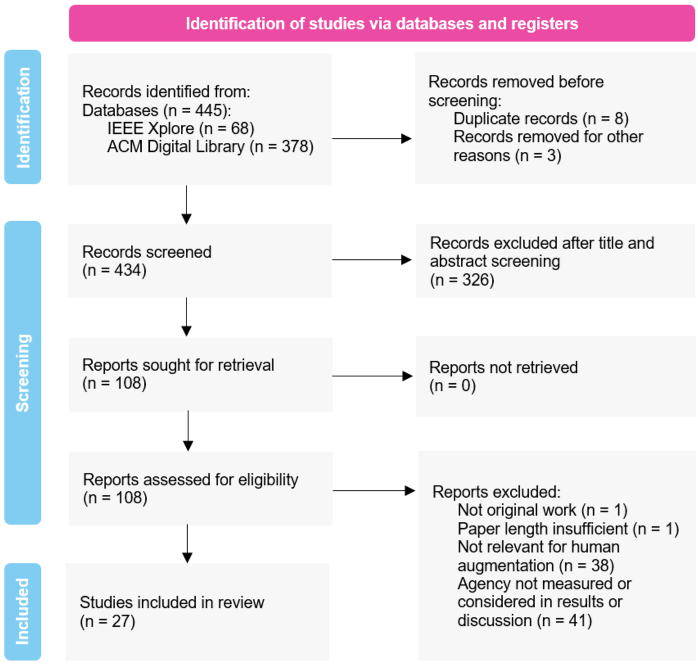


Figure 3.14: PRISMA flow diagram of the literature review process. Adapted from [250].

For the remaining 108 publications, we retrieved full texts and examined them in detail in the second screening phase. The article was discarded when it became clear that at least one of the inclusion criteria described above was not met. However, following the described process, we found zero publications that met all initial inclusion criteria, as none of the studies specifically focused on cognitive augmentation.

Consequently, we extended the inclusion criteria number four to be: “Focus on human augmentation or related domains from which insights on agency can potentially be transferred to cognitive augmentation“. In a third screening phase, we again examined the 108 publications from the previous screening phase, checking their eligibility based on the modified inclusion criteria. As a result, we included 27 works in this review (see exclusion reasons in Figure 3.14).

3.5 Results

In the following, we report the analysis of the 27 reviewed studies. We grouped them into cognitive, motor, and sensory augmentation topics based on the framework by Raisamo et al. [35]. Further studies that did not fall under these three categories, however, address agency in a context related to human augmentation, are presented in the *Agency in related domains* section.

3.5.1 Agency in Motor Augmentation

Seven articles were identified that examine agency related to motor augmentation [23, 253, 254, 255, 256, 257].

Kasahara, Nishida, and Lopes [23] developed a preemptive force-feedback system using electrical muscle stimulation (EMS) to enhance reaction time while maintaining the user's sense of agency. In a user study (n=12) consisting of a tapping task, they found that EMS actuation 160 ms after a visual target improved reaction time by 80 ms while preserving agency. Participants felt in control even when movements were EMS-induced. A second study (n=12) confirmed that this preemptive action approach with optimal timing yielded faster reaction times than participants' own reaction times and a higher sense of agency than traditional EMS methods, though voluntary actions still provided the highest sense of agency [23]. Their work demonstrates that carefully timed EMS can improve reaction speed without fully compromising agency.

Kasahara et al. [254] extended their preemptive action studies with another reaction time experiment where participants (n=17) tapped in response to an LED flash under three EMS conditions. Faster reaction times were retained post-EMS removal only if trained in the agency-EMS condition with optimal timing of preemptive action (40 ms before their natural reaction), but not after the fast-EMS (synced with the LED) or late-EMS condition (after natural reaction). This suggests that preserving agency also increases the effectiveness of motor adaptation after EMS training.

Yet, Kasahara, Nishida, and Lopes [23] only considered scenarios of congruent situations when an alignment between user-driven and machine-driven touch exists. Given this limitation, Tajima et al. [255] expanded upon this work by comparing assistive-touch and adversarial-touch in a force-feedback EMS study. They found that participants reported a higher sense of agency for favorable outcomes (assistive-touch) as compared to unfavorable outcomes (adversarial-touch) [255]. This suggests that the level of perceived agency is affected by an outcome bias [255]. Tajima et al. [255] also created the "agency-assistance trade-off matrix" (see Figure 3.15), depicting design implications for haptic systems using actuators. Joint success (i.e., user and EMS correct) preserves some sense of agency even with faster computer-driven touch, replicating previous studies [23, 254]. Forced success (i.e., user incorrect, EMS correct) involves corrective haptic assistance where the user involuntarily performs the correct action due to faster computer-driven touch, which hinders agency. Forced failure (i.e., user correct, EMS incorrect) should be prevented as it results in an incongruent as well as false outcome, whilst also diminishing agency. Joint failure (i.e., user and EMS incorrect) results in a negative outcome but could be useful for adversarial touch, with the system taking blame for failures when incorrect user-driven touch was predicted.

Shahu, Wintersberger, and Michahelles [256] examined EMS acceptance in four scenarios (motor learning, virtual reality, media player, and road safety), where one of the investigated factors was controllability (i.e., "a user's capacity to control a situation (sometimes referred

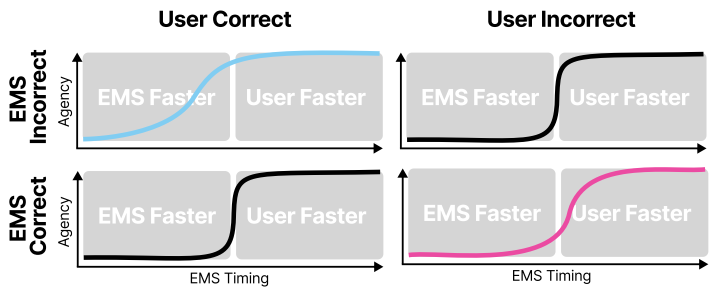


Figure 3.15: Agency-assistance trade-off matrix. Adapted from [255].

to as 'sense of agency')" [256]). An online survey (n=113) and interviews (n=5) showed the highest acceptance for VR-EMS and the lowest for road safety, where loss of control was a negative factor. They recommend EMS systems maintain user control and offer alternatives to preserve the sense of agency.

Finally, Coyle et al. [253] studied the impact of assistance levels (none/mild/medium/high) on agency in a machine-assisted point-and-click task with a gravity algorithm enhancing users' mouse movement (n=27). They found the highest sense of agency in no-assistance and mild-assistance conditions, with no significant difference between these two levels, indicating that computer assistance up to a certain level can still allow a high sense of agency. However, medium and high assistance significantly reduced agency, with no significant difference between them. These findings suggest that assistance can lead to a loss of agency once a certain threshold is reached, which is relevant for human augmentation applications.

So far, most identified studies on motor augmentation have focused on EMS technology. Venot et al. [257] investigated a multimodal BCI, examining how the timing of motor imagery tasks affected performance. Their BCI integrated eye-tracking to enhance the overall sense of agency, which participants reported via a questionnaire. However, the authors did not report or discuss the agency results.

In summary, these motor augmentation studies emphasize that the level of assistance and timing of the intervention are relevant factors for maintaining a sense of agency.

3.5.2 Agency in Sensory Augmentation

Research by Zolyomi and Snyder [258] relates to sensory augmentation due to its focus on vision-enhancing technology. It is centered on understanding the social implications of digitally enhanced vision based on a head-mounted assistive device for low vision. In interviews (n=13) with long-term users, the authors observed that users' desire to experiment with the assistive technology and their perception of its value was influenced by users' overall sense of agency in life [258]. For future research on agency in cognitive augmentation, this suggests that considering a more holistic view of agency in life, rather than just during specific machine-assisted tasks, may lead to valuable insights.

3.5.3 Agency in Related Domains

In the following, we analyze the role of agency in domains related to human augmentation, grouped by (i) neurotechnology, (ii) human-computer integration, (iii) VR and AR, and (iv) AI and machine learning applications.

3.5.4 Neurotechnology

Neuroscience technology is commonly used to enhance human cognitive abilities [259, 260, 261]. Therefore, studies that do not investigate neurotechnology specifically for human augmentation but consider the role of agency are relevant here, as insights can likely be transferred to the design of cognitive augmentation technologies.

A mixed-method study by Martinez et al. [262] examined the ethical concerns for neurotechnology in the future workplace. Participants had generally positive attitudes towards future brain-scanning technologies, especially those boosting concentration (an example of cognitive augmentation). However, interviews (n=10) revealed concerns about "trust and agency", particularly regarding devices that alter emotional states, which the authors suggest threaten users' sense of agency [262].

In the context of BCIs for stroke rehabilitation, low BCI performance can decrease agency [263]. Hougaard et al. [263] found that fabricated input (i.e., preprogrammed positive feedback) increased perceived agency and reduced frustration in users in (i) a surrogate BCI based on eye blinks [263], (ii) a surrogate BCI study with stroke patients (n=13) [264], and (iii) real motor imagery tasks in an online BCI study (n=16, healthy) [264].

Staying in the context of BCIs, Mercado-García et al. [265] investigated whether the design approach (traditional Graz BCI vs. user-centered design including a VR CAVE system) impacts the modulation of EEG brain signals in a motor imagery task. They found that user-centered design enhanced the brain activity modulation. The authors suggest that natural interactions that resemble BCI-users' daily life activities offers them "a real sense of agency" [265], for which they promote user-centered design as a promising alternative to traditional BCI design. However, Mercado-García et al. [265] did not measure users' agency, hence the benefit on agency is an assumption that needs further investigation.

3.5.5 Human-Computer Integration

According to Raisamo et al. [266], human-computer integration, which uses computing resources and AI to support and work together with a human, is closely connected to cognitive augmentation. Hence, insights on the role of agency in human-computer integration can likely be transferred to cognitive augmentation research. Mueller et al. [267] explored shared agency of bodily control in intertwined human-computer integration. In one case study, "EduExo," an exoskeleton with an electromyography sensor supports arm movement (i.e., motor augmentation). The user can access all system data on a laptop, making the machine's agency transparent. The authors present a framework with two key dimensions of intertwined systems (awareness and alignment of the machine's agency) and four system roles (angel, butler, influencer, adversary) [267]. Their framework supports designing cognitive augmentation systems with shared agency.

3.5.6 VR and AR

Human augmentation builds upon and draws elements from the fields of VR and AR [35]. Therefore, examining the sense of agency in those contexts can be insightful for designing cognitive augmentation technologies.

Several identified VR studies, which consider the sense of agency, focus on embodiment [268, 269, 270, 271, 272, 273]. It was shown that embodying a body-matched virtual avatar as opposed to a virtual object (a color-matched box at the place of the body) increased participants' sense of agency during a cognitive task (n=11) [268]. Furthermore, avatar hand realism affected agency, with lower agency scores found for abstract hands compared to iconic or realistic hands [269]. In the same study, body continuity, i.e., whether the virtual hands and arms were disconnected or connected, showed no significant effect on perceived agency [269].

A VR study (n=33) on the emotional effects of the full-body ownership illusion demonstrated that movement synchrony between virtual and real body led to increased emotional valence as well as increased sense of agency compared to a pre-recorded movement condition[270].

Another VR study (n=24), in which motor tasks were performed, found that participants' sense of agency was higher in a virtual hand condition compared to a physical keyboard condition (without virtual representation) [271]. However, there was no significant difference in agency between virtual hands and virtual controllers [271].

In VR, it is also possible to embody multiple bodies simultaneously. In a "Parallel Embodiment" system developed by Takada et al. [272], users play ping-pong while simultaneously controlling two robot arms. In a survey (n=142), users reported high ratings for the sense of agency over both robot arms despite visuomotor incongruences. However, some users suggested the robot arm itself had agency, with comments such as "the robot arm is moving on its own" [272]. Takada et al. [272] conclude that agency may be influenced by users' prior knowledge of another agent's presence.

Furthermore, Miura et al. [273] found that participants perceived agency over four virtual bodies in parallel when controlling them simultaneously. The results from one of their self-reported agency items suggest that the sense of agency might have decreased with more bodies. However, as this was inconsistent with results from another agency item, the effect remains open for future investigations.

In the context of object translation in handheld AR, Sun et al. [274] observed a higher subjective and objective agency in one degree of freedom compared to three degrees of freedom. Sun et al. [275] also revealed a negative association between mental workload and agency in head-mounted AR.

3.5.7 AI and Machine Learning Applications

AI methods are an essential part of human augmentation [35], including memory augmentation [276]. For instance, AI assistants can efficiently carry out various tasks and make decisions for the user on their behalf. The model by Raisamo et al. [35] for wearable augmentation proposes that AI is an enabling technology specifically for cognitive augmentation. Hence, the role of agency in AI and machine learning (ML) applications will be considered in the following since insights apply to cognitive augmentation.

Xu et al. [277] investigated explainable AI (XAI) in AR, proposing a XAI design framework based on a survey and expert workshops. To provide user agency, they recommend always making AI explanations accessible and offering detailed explanations upon request.

In AI-mediated social interactions, Wang et al. [278] developed an AI agent for online learning platforms assisting users in building social connections. An interview study (n=26) revealed that students were concerned about losing agency over the connections they build . The authors suggest that a balance is required between a sufficient amount of pressure to ensure successful interactions mediated by AI and preserving users' agency with whom to start a conversation. This is consistent with the previously described findings from motor augmentation [253] (see subsection 3.5.1), which already highlighted a required trade-off between assistance and agency.

In healthcare, Thieme et al. [279] developed an AI application to predict treatment outcomes in human-supported, internet-delivered Cognitive Behavioral Therapy (iCBT) for depression and anxiety. They found that AI design could affect clinical supporters' sense of agency and recommend that AI should inform the care rather than interfere with medical assessments, keeping the supporter in charge of examining patients' individual circumstances and potential reasons for the AI prediction outcome.

Sali et al. [280] explored natural language understanding (NLU) in games, finding that players reported more agency when the NLU interface provided pauses and prompts (as opposed to free-form text entry and reactive pauses), even if it limited their actions and free will. This suggests that guided actions can enhance the sense of agency.

Moreover, Sun et al. [281] found that users of automated machine learning systems actively exercise agency to overcome challenges in customizability, transparency, and privacy by employing workaround strategies.

Finally, Ahmad et al. [282] studied how tangible control and feedback mechanisms affect users' sense of agency in smart voice assistants. Their qualitative analysis highlighted the importance of total control over the devices and easy-to-use control mechanisms. The authors recommend designing future voice assistants with tangible hardware controls (e.g., physical buttons to mute the microphone), aiming to increase users' agency over their privacy. Although not focused on cognitive augmentation, these findings can apply to cognitive augmentation technologies using voice assistants, e.g., as memory extenders.

3.6 Agency in Cognitive Augmentation: A Research Gap

The systematic literature search identified no study that directly aimed at enhancing human cognitive capabilities (such as memory, problem-solving, attention, cognitive overload, etc.) using digital technology while also measuring or discussing the sense of agency. This research gap will be addressed in the discussion (see section 3.10).

One identified study closely related to cognitive augmentation was conducted by Semertzidis et al. [283], who developed Dozer, a closed-loop wearable beanie that accelerates sleep onset through auditory and electrical brain stimulation. After an EEG detects drowsiness, the user's brain is stimulated through transcranial alternating current stimulation (tACS), and speakers play pink noise for sleep enhancement. In an in-the-wild study (n=11), the authors identified "closed-loop neurocentric agency" [283] as a user experience theme related to bodily agency. They found that: (i) Participants demonstrated high agency over the system despite feeling disconnected due to a lack of feedback, process understanding, and system familiarity. (ii) Knowledge of the system's function is important for experiencing agency. (iii) Bodily-integrated systems can provide a high sense of agency without explicit user inputs (i.e., initiating causal influence over the system). Nevertheless, participants reported a diminished sense of body ownership and high awareness of the system's hardware, which compromised Dozer's ability to effectively promote sleep onset.

Whilst aiming to accelerate sleep onset does not directly enhance cognitive capabilities, Dozer can be considered a cognitive augmentation technology in a broader sense. It targets cognitive processes involved in sleep regulation through a closed-loop wearable utilizing EEG, auditory, and electrical brain stimulation, thereby augmenting cognitive aspects related to sleep onset. Overall, this study implies that the user's understanding of the augmentation technology's functionality may be relevant for maintaining a sense of agency. Yet, the specific field of sense of agency in cognitive augmentation remains largely unexplored.

3.7 Mapping the Literature on Agency on Cognitive Augmentation

To understand the role of agency in cognitive augmentation technologies, we performed a systematic review of two decades of original HCI research following the PRISMA guidelines. We reported our analysis of 27 reviewed studies regarding their role of agency, categorized by cognitive, motor, and sensory augmentation, as well as four domains related to augmentation: (i) neurotechnology, (ii) human-computer integration, (iii) VR and AR, and (iv) AI and ML applications.

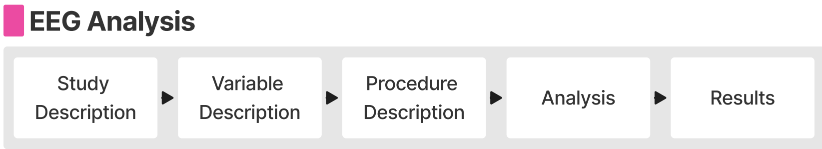


Figure 3.16: EEG analysis process overview.

We systematically identified a research gap regarding the role of agency in cognitive augmentation technology. One potential explanation is that publications investigating agency in cognitive augmentation technologies exist but were not found due to the chosen keywords or the limitation of our search to two databases[1]. Another potential explanation is that the HCI community has not yet investigated this topic. Whilst agency in cognitive augmentation appears underexplored, several relevant studies were found in motor augmentation. This might be because the sense of agency, by definition, relates to feeling in control over one's *actions* [132], i.e., having motor control. Consequently, the topic of agency may hold greater prominence in motor augmentation than in cognitive augmentation. However, low agency over one's cognitive performance can have various negative consequences, such as shifting responsibility to the augmentation technology [284] and taking higher risks [Aut1].

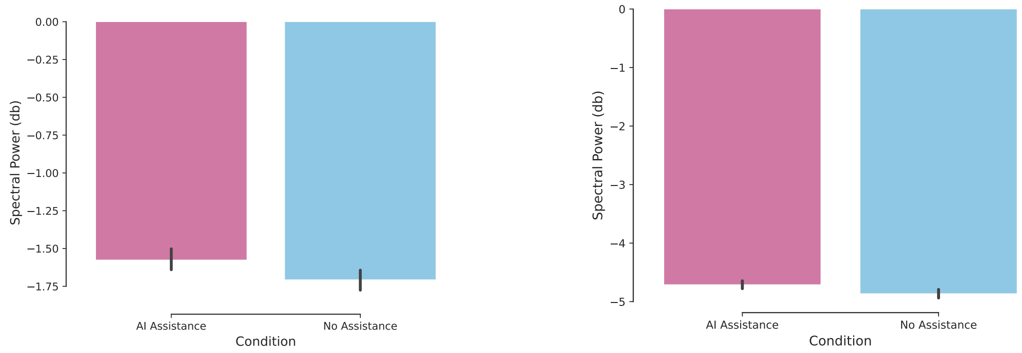
3.8 Exploring Neuroagency: Assessing the Impact of Perceived AI Support on Sense of Agency

While the impact of technology on perceived control has attracted attention in HCI [44, 109, 133, 253], quantifying and understanding perceived cognitive agency remains a significant research gap. Our literature review shows that neural activity is linked to the perception of agency. For example, studies have successfully linked alpha and beta band oscillations in the EEG to agency perception during motor actions, suggesting that these frequencies serve as quantifiable metrics [139]. Importantly, research further demonstrates that these neural markers correlate with the subjective sense of agency [285]. This section describes the data analysis of an EEG experiment that evaluated the perceived cognitive agency. In the following, we evaluate the feasibility of EEG data for assessing cognitive agency.

3.8.1 Measures

Previous work by Kang et al. [139] has shown that Agency can be operationalized through alpha (8 - 12 Hz) and beta (12 - 15Hz) band oscillations during motor tasks. In this analysis, we further test if this operationalization holds for cognitive agency; therefore, we study cognitive agency by evaluating the spectral behavior of the alpha and beta bands for objectively assessing perceived agency in real-time.

[1] ACM Digital Library and IEEE Xplore.



(a) Alpha Band (b) Beta Band

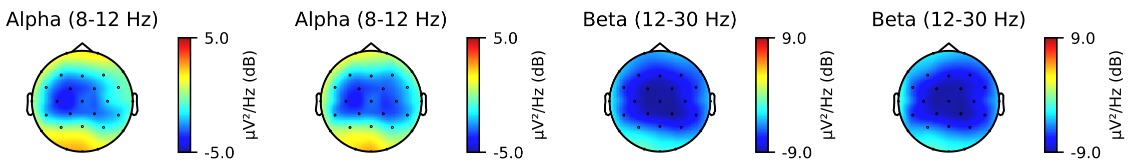
Figure 3.17: Spectral Power for Alpha and Beta Bands: There is a decrease in spectral power for both frequency bands in the No-Assistance condition, suggesting a higher sense of agency.

3.9 Results

This section details our data analysis of the author's experimental study. For questionnaire items, we employed independent-sample t-tests to establish statistically significant differences between the AI-assisted and no-assistance conditions. Conversely, we utilized generalized linear models (GLMs) to analyze the Spectral Analysis values. This choice facilitated the dissociation of the effects arising from the card type and those due to the experimental condition (AI-Assistance vs. No-Assistance). In the following, we report the EEG spectral analysis.

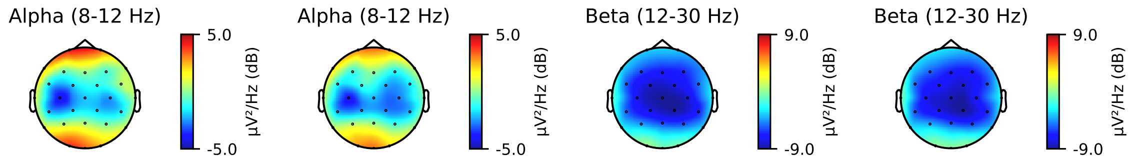
3.9.1 Spectral Analysis

We initially assessed the spectral power at each electrode. The spectral power of each 1-second epoch was determined using the short-time Fourier transform (STFT) with a 512-point sliding Hanning window and 50% overlap. Subsequently, the results were averaged across trials for each condition and electrode. The averaged spectral power for each condition and electrode was then accumulated over two frequency bands: alpha (8 - 12 Hz) and low-beta



(a) AI-Assistance. (b) No-Assistance. (c) AI-Assistance. (d) No-Assistance

Figure 3.18: Spectral Power Topographical Maps displaying Alpha and Beta frequencies under conditions with AI-Assistance and without AI-Assistance in the scenario of Win Cards.



(a) AI-Assistance. (b) No-Assistance. (c) AI-Assistance. (d) No-Assistance.

Figure 3.19: Spectral Power Topographical Maps displaying Alpha and Beta frequencies under conditions with AI-Assistance and without AI-Assistance in the scenario of Loss Cards.

(12 - 15 Hz). We selected 19 electrodes (FP1, FP2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, Pz, P3, P4, T5, T6, O1, O2) based on previous work [139]. For further statistical analysis, the spectral power for each condition was computed in decibels and plotted on a topographical map (see Figure 3.18 and Figure 3.19).

In analyzing the effects of assistance condition and card outcome on Spectral Power, we employed a Generalized Linear Model (GLM) with a Gaussian family and an identity link function. The model for the alpha band revealed a statistically significant negative association between the presence of NO-ASSISTANCE and Spectral Power, with a coefficient of -.13 (SE = .050, $z = -2.63$, p = .008), indicating that Spectral Power decreases by .13 units when participants were informed that assistance is not provided, holding other factors constant. Similarly, the outcome of 'Win Card' was significantly associated with a decrease in Spectral Power by 1.38 units (SE = .17, $z = -7.99$, $p < .001$), compared to the baseline condition of 'Loss'.

For the beta band, the model identified a statistically significant negative association between NO-ASSISTANCE and Spectral Power, with an estimated coefficient of $-.15$ (SE = .052, $z = -2.98$, p = .003), suggesting that Spectral Power decreases by .15 units when participants were informed that assistance is not provided. Conversely, the 'Win Card' outcome was significantly associated with a reduction in Spectral Power, evidenced by a coefficient of $-.95$ (SE = .18, $z = -5.30$, $p < .001$), indicating a substantial decrease in Spectral Power associated with winning outcomes.

3.10 Discussion

Our work makes two key contributions to agency on cognitive augmentation research. First, we reviewed agency in this domain, identifying central themes and research gaps. Second, we present empirical evidence on the impact of perceived AI assistance on user agency, utilizing neurophysiological markers to measure this effect based on the data collected in chapter 3. In this section, we discuss and articulate the main findings of this paper; in detail, we discuss mapping neurophysiological markers from agency in motor tasks for cognitive augmentation. Then, we discuss the impact of cognitive augmentation on agency; then, we discuss insights and future directions. For a general overview of antecedents of agency in HCI, see [109], which goes beyond our study.

3.10.1 Quantifying Sense of Agency in Cognitive Augmentations

This study on cognitive augmentation reveals how AI assistance affects users' sense of agency through its electrophysiological underpinnings. Our findings reveal that participants experienced a significant decrease in the sense of agency when led to believe in the presence of AI assistance, as evidenced by shared variations in alpha and low-beta EEG power in the parieto-occipital regions. This suggests that the mere belief in AI assistance can effectively lead to a diminished sense of agency in the context of cognitive augmentations at a neurophysiological level.

Compared to a baseline condition (No-Assistance), this outcome was observed through alterations in alpha and beta brain wave activities. Such findings highlight how perceived AI support can evoke different patterns in brain frequency that can allow discrimination across agency states. The decrease in the sense of agency aligns with the results reported by Bu-Omer et al. [141], providing a neurophysiological basis for understanding how external cues and perceived technological interventions can modulate individuals' sense of control over their actions. This is in line with the call for disambiguating agency definitions in HCI as outlined in Bennett et al. [109].

The findings also reveal that winning outcomes lead to a significant decrease in spectral power compared to losing outcomes across both frequency bands. This could indicate that successful outcomes, particularly in the context of the CCT, elicit a neurophysiological response associated with reduced cognitive load or decreased need for further action adjustment, as supported by Chen, Chaudhary, and Li [286]. This outcome delineates distinct brain activities during the anticipation and outcome phases of win and loss scenarios.

Winning outcomes can enhance the perception of having effectively influenced an event, i.e., "successful" agency. In situations where individuals believe AI is assisting them, this success can further shape their sense of control. Essentially, positive outcomes from tasks, when combined with the notion of AI support, refine how people perceive their influence over outcomes, reinforcing their sense of agency.

The new results expand on the role of alpha and beta on the sense of agency, as we introduced a new paradigm that was not explored before, while still replicating results from previous work [136, 141, 287]. For instance, Buchholz et al. [142] shows that the belief of agency itself changes the dynamics within sensorimotor networks, particularly highlighting how the beta band is modulated by one's causal belief about the origin of actions. This suggests that the decrease in beta frequency we observed might be not only a direct response to the AI assistance but also a reflection of the participants' altered belief systems regarding the control over their actions. Haggard, Clark, and Kalogeras [136] provides a overview of the sense of agency's underlying brain mechanisms, emphasizing the role of predictive processes and the contribution of alpha and beta waves in the sense of agency. This supports our findings from a theoretical standpoint, suggesting that the alterations in brain frequencies observed in our AI-assisted condition may reflect disruptions in the participants' prediction about their outcomes.

3.10.2 Individuals Who Believe Being Cognitively Augmented Present Reduced Agency

The observed effect in alpha and beta frequencies, indicative of a diminished sense of agency under perceived AI-Assistance, highlights a central consideration for HCI design: the need to maintain or enhance a user's sense of control and autonomy when interacting with intelligent systems. This consideration becomes especially pertinent as we explore the integration of neurophysiological markers of a sense of agency as an input for interaction, such as in BCIs. In rehabilitation, for instance, adaptive protocols could significantly improve recovery outcomes by aligning therapeutic activities with the patient's specific neural patterns, thereby reinforcing their sense of agency. Similarly, in educational technologies, learning experiences that adapt to the student's sense of agency could make education more engaging.

3.10.3 Neurophysiological Insights into Agency with Cognitive Augmentation

Our EEG findings demonstrate decreased alpha and beta spectral power with perceived AI-Assistance and upon winning, elucidate the neurophysiological facets of agency within cognitive augmentation. This reveals how AI perceptions and outcomes affect users' neurophysiological states, influencing their sense of agency. These insights inform the design of augmentation technologies that enhance abilities while preserving autonomy. We argue that considering users' neurophysiological reactions to aid and feedback, should aim for empowering rather than overpowering user experiences. AI development and design of AI should consider how it affects users' sense of agency at a neurophysiological level. By considering this, Designers can create AI systems and applications that are more aligned with users' sense of agency and better control over the AI outcomes, ultimately leading to technologies that are both more accepted by users.

3.10.4 Quantifying the Sense of Agency in Real-Time

In our analysis, we utilized an EEG-based spectral analysis and based on literature [139, 141, 251], we demonstrated that the alpha and beta EEG metrics initially developed to measure sense of agency in a motor action context also can measure the sense of agency in Human-AI interaction. Further supporting this, Freeman et al. [288] showed that EEG indices using alpha and beta bands could be employed in adaptive automation systems, where the system dynamically switched between manual and automatic modes based on changes in user engagement measured through these EEG metrics. Similarly, Prinzel III et al. [289] developed a bio-cybernetic system that utilized an EEG index based on beta and alpha bands to modulate operator engagement in real time, demonstrating the feasibility of using these metrics for adaptive automation in cognitive tasks. This opens up new alternatives for HCI researchers to integrate real-time agency measurements in their experimental designs, with the advancement of EEG devices.

3.11 Implications and Future Research Directions

Based on the knowledge gained from our systematic review and analysis, we now draw future research directions to further investigate agency's role in cognitive augmentation.

3.11.1 Find the Right Balance Between Augmentation and Agency

In the context of machine-assisted point-and-click tasks, it was possible to assist users up to a certain level (mild assistance) without harming their sense of agency, but a rapid drop of agency occurred once more assistance was provided [253]. A good balance between the level of AI assistance and agency preservation was also identified as a crucial factor in the context of building AI-mediated social connections [278].

This suggests that the level of intervention may fundamentally impact agency, a finding that can likely be transferred to cognitive augmentation. **We recommend that designers of cognitive augmentation technologies find the right balance between the augmentation level and the perceived agency level for their specific system.** This will likely depend on the specific use case, how important a high sense of agency is for the user in the given context, and how a lower level of augmentation would affect the user and their surrounding. Hence, future research should investigate this trade-off between the level of augmentation and agency for different types of cognitive augmentation technologies in various contexts. If a distinct threshold exists after which agency is lost, it is important to identify it so that designers can make informed decisions.

Finding this right balance, hence leaving the user in control to some extent, may also increase the acceptance rate of the augmentation technology [256].

Moreover, **we suggest considering human-computer incongruent situations when designing cognitive augmentation technologies [255]. In particular, prevent forced failures by design, e.g., a surgeon using an augmentation technology in medical care. Avoid forced successes when agency preservation is highly important, but allow it in training and safety-critical scenarios.**

3.11.2 Optimize Intervention Timing

The level of assistance and the exact timing of the intervention must be considered [23, 254]. For motor augmentation, it was found that early preemptive EMS actuation decreased the sense of agency [23]. Identifying the optimal timing for preemption resulted in faster reaction times whilst preserving the user's agency to some extent [23], even after removing the EMS device [254].

Likely, the intervention timing is also highly relevant in maintaining agency over cognitive augmentation. Designers should carefully consider the exact moment the cognitive augmentation technology assists the user. In future studies, this timing should be manipulated in controlled experiments to investigate whether a sweet spot exists in which users' cognitive performance can be enhanced whilst also experiencing a high sense of agency.

3.11.3 Examine agency in different types of cognitive enhancement

People appear to be most excited about brain-scanning devices that boost concentration; however, when emotional states are altered, they have trust and agency concerns, as this may threaten their sense of agency [262]. This indicates that the level of agency and amount of concern may depend on the cognitive ability that the system augments. **We recommend that future research examines which types of cognitive abilities can be enhanced whilst maintaining users' agency.**

3.11.4 Reduce Cognitive Load

In head-mounted AR, Sun et al. [275] revealed a negative association between mental workload and the sense of agency. Consequently, cognitive augmentation in AR, which aims to reduce users' cognitive load, may lead to a win-win situation of decreased mental burden whilst increasing agency at the same time. Although measuring cognitive load has been the focus of HCI and user experience research since decades [290], further research is needed to establish the causal relationship between cognitive load and agency. Nonetheless, **we recommend that designers aim to minimize cognitive workload.**

3.11.5 Explain the System's Functionality

Explanations of AI output were found to be crucial for fostering agency in AI systems [277]. Similarly, users of a bodily-integrated sleep wearable expressed a desire to comprehend the system's functioning to attain a high sense of agency [283]. Therefore, **in the design of cognitive augmentation technologies, consider providing adequate explanations to enable users to understand the system's operations and to foster a relationship of competency and trust [283] between users and the system.**

3.11.6 Resolve Performance Issues in Neurotechnology

BCI studies conducted by Hougaard et al., Hougaard et al. [263, 264] indicate that when individuals' sense of agency is already reduced due to low performance of the BCI, computer assistance in the form of preprogrammed, fabricated input can improve agency. It appears that negative consequences (here, reduced agency) of low technology performance are mitigated through increased reliance on computer assistance in the form of fabricated input. Moving forward, **it is crucial to resolve the underlying performance issues (the cause of the problem) to enhance agency in BCIs, which may eliminate the need for fabricated input.**

3.11.7 Use Objective Measures of Agency

As no gold standard on measuring agency exists yet, we also looked at which agency measures were utilized in the reviewed studies. To assess agency, the majority of identified studies relied on subjective self-reports by asking participants to respond to an agency questionnaire [257, 274, 275], commonly a single-item 7-point Likert scale question [23, 254, 255, 281]. In some of the included VR studies, the agency measure was included in a self-report questionnaire concerning embodiment [268, 269, 272, 273], ownership illusion [270], or general VR experience [271].

Several of the studies included did not directly measure the sense of agency. Instead, agency was identified as a theme based on qualitative study results [256, 258, 262, 267, 277, 278, 279, 280]. However, four studies did not specifically measure the sense of agency [265]. Instead, some assessed user's perceived control [263, 264, 282].

Only two studies utilized objective measures of agency. Coyle et al. [253] assessed intentional binding as an implicit measure of the sense of agency, using the Libet clock method in one experiment and interval estimation in the other. Finally, only one study used physiological measures of agency based on EEG data, applying spectral power analysis and brain activity analysis [274].

The advantage of objective agency measures is that they do not rely on introspection and subjective reports, whereas self-reported agency can be biased and confounded [137]. In future studies, we recommend enhancing objectivity by combining subjective measures such as questionnaires with objective, physiological measures of agency [137], for example, using EEG data [138, 274], as our results have shown that state-of-the-art EEG correlates for sense of agency apply for the context of cognitive augmentation. **Therefore, we recommend that researchers include neurophysiological measures when assessing the sense of agency during cognitive augmentation.**

4

COGNITIVE AUTONOMY

This chapter is based on the following publication:

Envisioning Futures: How the Modality of AI Recommendations Impacts Conversation Flow in AR-enhanced Dialogue - Steeven Villa, Yannick Weiss, Karin Lu, Moritz Ziarko, Albrecht Schmidt, Jasmin Niess - *In International Conference On Multimodal Interaction (ICMI '24). Association for Computing Machinery.*

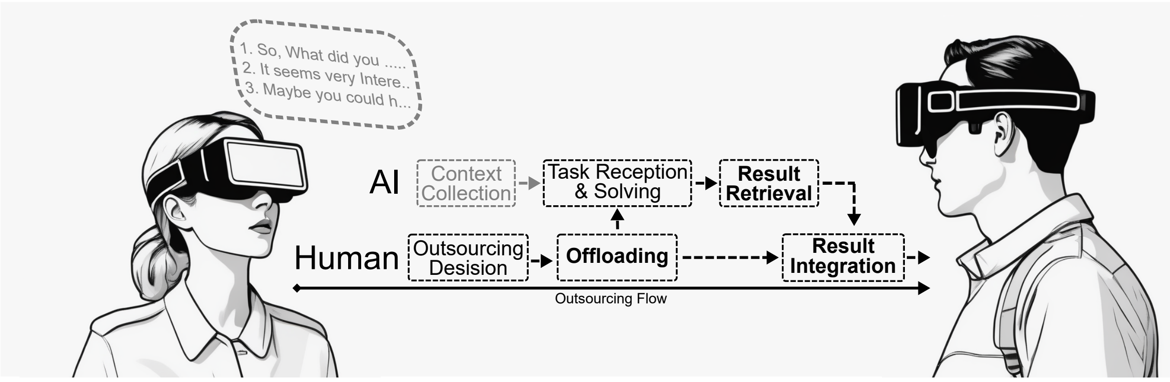


Figure 4.1: Interface for visual modality retrieval: AI suggestions are rendered the near-peripheral vision of participants, the UI element on the right is a placeholder for pictures in future implementations.

After examining decision-making and agency in cognitive augmentations in the previous chapter, we now explore the impact of functional cognitive augmentation on behavior during dialogue—where autonomy and social context intersect. Notably, the mere belief of being augmented already influences behavior; here, we investigate the effects of functional augmentations. Augmenting human cognition with AI in augmented reality presents a compelling case, as it integrates social context, cognition, and rapid interaction. Using a fully functional prototype, we conducted an in-depth evaluation of its impact on conversational dynamics and participant behavior.

4.1 System Specification & Apparatus

We developed a system that supports users in conversational situations. The system continuously collects contextual information about the conversation via an omnidirectional microphone. When the user requires support, they can trigger to *offload* the current context of the conversation and start processing the information and *retrieving* the processed information. The system was tuned to suggest ways to continue the conversation based on the current state of the conversation.

4.1.1 Implementation

The hardware configuration of the system comprises a Video-see-through Augmented Reality (AR) device (Meta Quest Pro by Meta, Menlo Park, The USA) for visual rendering and inputs, non-occluding earpods (Sony Linkbuds, Sony, Tokyo, Japan) for auditory rendering, and an omnidirectional microphone (Senheiser SP20, Senheiser, Wedemark, Germany) for capturing audio input from all individuals present in the room. The system incorporates the Cognitive Services API by Microsoft (Redmond, The USA) for tasks related to text-to-speech and speech-to-text processing. Furthermore, it leverages the capabilities of GPT-3.5 by OpenAI (San Francisco, USA) to process the user's offloaded tasks. We describe all these components in detail below.

Action Triggers

The system incorporates two methods: mechanical, and ocular. The MECHANICAL TRIGGER requires the user to press a button while in the OCULAR TRIGGER, the user has to look up right (information processing ocular movement [291]). The first method uses a button on the VR controller. The second uses the integrated eye-tracking feature of the headset.

Retrieval Modalities

The retrieved information was displayed visually and auditorily. In line with previous work, [91] we hypothesized that the VISUAL RETRIEVAL offers high information density and precision

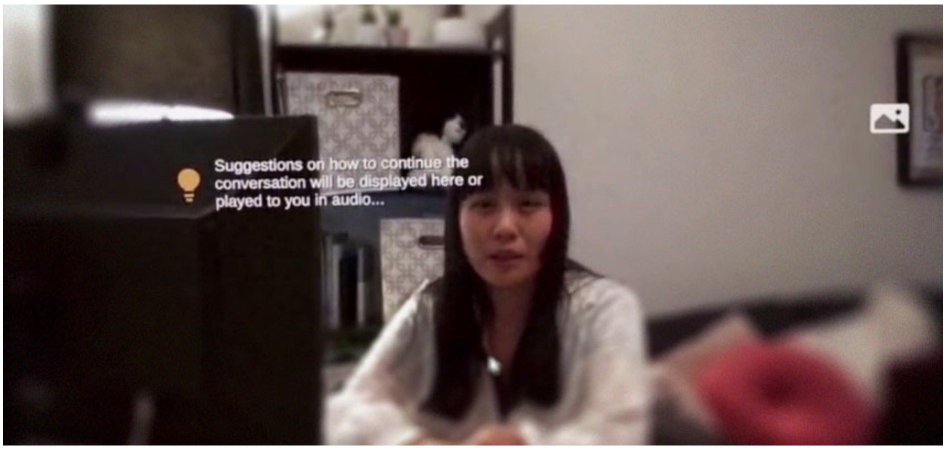


Figure 4.2: Interface for visual modality retrieval: AI suggestions are rendered the near-peripheral vision of participants, the UI element on the right is a placeholder for pictures in future implementations.

but can split attention from the conversation. In contrast, AUDITORY RETRIEVAL potentially maintains engagement but has a limited bandwidth. The former was implemented with a floating text in the near-peripheral vision of the user [292]. While the later used text-to-speech and rendered the output through headphones. These headphones allow audio rendering without blocking out the surrounding environment.

Conversation Context Processing:

We implemented the information offloading process by continuously having the system record contextual information from ongoing conversations through active listening and encoding this into text using speech-to-text. Only when the user decides to offload this contextual information to the AI using an action trigger described in section 4.1.1, it is actually processed. This approach is designed to mitigate potential bandwidth constraints. For example, in contrast with [93], the user does not have to communicate the full prompt to the system, as the system has preemptively recorded the context.

Task Solving:

In the present system, the task-solving stage is executed utilizing the OpenAI API. Subsequently, the accumulated contextual information is used as a prompt, accompanied by customized instructions, to generate alternatives for continuing the conversation. The output information is then filtered to avoid wrong-formatted responses and finally retrieved by the user.

Result Retrieval:

Once the offloaded information has been processed, the outcomes of this operation are automatically returned to the user one of the two retrieval modalities.

4.2 User Study

We conducted a within-subjects laboratory experiment to investigate the user experience and conversation dynamics when receiving AI support. We manipulated two factors: ACTION TRIGGER with the levels: MECHANICAL and OCULAR, and RETRIEVAL MODALITY with the levels VISUAL and AUDITORY. The order of the conditions was counterbalanced using Latin Square, and the order of the scenarios in the DCT was randomized. Building on prior research, we adopted an exploratory approach, as directly comparing to conventional baselines can introduce multiple confounding factors. Our primary aim was to deepen our understanding of how the modality of AI recommendations impacts conversation flow in AR-enhanced dialogue, thereby providing valuable insights for future work.

4.2.1 Discourse Completion Task (DCT)

The Discourse Completion Task (DCT) originates from pragmatics [293]. This method has been employed for research and evaluation purposes [294]. It involves using scenarios to prompt individuals to respond in writing or speech, enabling the collection of diverse and comparable cross-linguistic data [295]. We employed DCT with validated scenarios extracted from previous work that have been shown to be appropriate for discoursive analysis [296, 297, 298].

The experimenter verbally presented a randomly selected scenario to the participant, concluding with a question. The participant is then required to continue the conversation, embodying the character in the scenario. Participants had to decide whether to trigger the AI support or not. Each participant received a total of 12 DCT scenarios in a randomized way (see supplementary material).

4.2.2 Data Collection

This study aimed to investigate the influence of AI support on face-to-face conversation dynamics, user perceptions, and overall experience. We employed a mixed-method approach, including four controlled conditions and an open conversation scenario.

The initial four conditions served the purpose of familiarizing participants with the system configurations, eliciting initial insights on potential conversational integration during a normal setting, and capturing conversational behaviors under AI assistance. This prepared the participants for the open conversation phase, where they freely utilized the system in a natural dialogue. Additionally, we collected quantitative data on pre-existing perceptions of performance-enhancing technologies and per-block questionnaires to capture participants' evolving perspectives throughout the study. To complement this data, we conducted semi-structured interviews with the participants.

Conversation Dynamics Data

Using the DCT, we recorded the prompts suggested by the system and the responses provided by the participants. Afterward, we transcribed and cleaned all the responses for further processing. From these responses, we calculated response similarity, response length, and response delay.

Response Delay was defined as the time between the experimenter finishing their prompt and the participant starting their verbal response. This measure captured two key processes: (1) the system's technical processing time (recording, transcription, analysis, display) and (2) the participant's cognitive processing time (understanding, formulating, and initiating response). *Response Length* was operationalized as the total number of characters included in the participant response after cleaning transcription artifacts. *Response Similarity* we converted both participant responses and system prompts into numerical representations (embeddings). These capture the semantic meaning of each sentence. We then calculated the cosine similarity between these embeddings, resulting in a score between 0 (no similarity) and 1 (perfect similarity). This allowed us to quantify how closely aligned the participant's response was to the intended meaning of the system prompt.

Questionnaires:

Task Load: We administered the NASA-TLX task load questionnaire) [189] to assess potential variations in task load resulting from the system configurations. The NASA-TLX evaluates task load across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The NASA-TLX is a well-established and validated instrument for measuring task load across different situations [299]. *The Sense of Agency Scale (SoA)* [135] assesses an individual's overall beliefs regarding their sense of agency, which is the feeling of being in control and the initiator of their actions. The scale comprises two interrelated factors: Sense of Positive Agency (SoPA) and Sense of Negative Agency (SoNA). *Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale* is a standardized tool to measure attitudes towards performance-enhancing technologies [Het2]. It consists of two factors: Social Threat and Agency, which measure the perceived societal threat of an augmentation device, and the user's sense of ownership over their actions when using such technology. For evaluating system usability, we employed the *System Usability Scale (SUS)* developed by Brooke [300]. This questionnaire is a well-established and validated tool for assessing the subjective usability of technological systems [301].

Semi-Structured Interviews

All audio recordings from the interviews were transcribed verbatim and imported into the Atlas.ti analysis software. We applied open coding combined with pragmatic thematic analysis, as described by Blandford et al. [30]. As a first step, we familiarized ourselves with the data. Data familiarization involved multiple readings of the material to gain a comprehensive understanding. Then, two researchers coded a representative sample of 25% of the material using open coding in line with Blandford et al. [30]. Next, an initial coding tree was established through iterative discussion. The remaining transcripts were split between the two researchers and coded individually. A final discussion session was conducted to structure the coding tree after the material was coded. This was followed by a final discussion session to construct and refine themes based on our material [30].

4.2.3 Participants

We recruited participants through the university's mailing lists and our extended networks. We recruited a total of N = 21 participants, from which 8 identified as female and 13 as male. The average age of our participants was twenty-six years (M = 26.85, SD = 4.57). Participants were compensated 6 euros/30 min for participating in the study. The study was approved by an ethics committee.

Table 4.1: Participant Demographics

Participant	Age	Gender	Education	Ocupation	Participant	Age	Gender	Education	Ocupation
P1	23	Female	Some Secondary Education	Student	P12	23	Male	Some University but no degree	Student
P2	23	Male	Some University but no degree	Employed full-time	P13	29	Male	Graduate or professional degree	Employed full-time
P3	35	Male	Graduate or professional degree	Student	P14	35	Male	Graduate or professional degree	Student
P4	23	Female	University Bachelors degree	Student	P15	28	Male	University Bachelors degree	Student
P5	29	Male	University Bachelors degree	Employed part-time	P16	27	Male	University Bachelors degree	Student
P6	26	Male	University Bachelors degree	Student	P17	30	Male	Graduate or professional degree	Employed full-time
P7	22	Male	University Bachelors degree	Student	P18	36	Male	Graduate or professional degree	Employed full-time
P8	21	Female	Some University but no degree	Employed part-time	P19	32	Female	Graduate or professional degree	Employed full-time
P9	26	Female	University Bachelors degree	Student	P20	25	Male	Completed Primary	Student
P10	23	Female	University Bachelors degree	Student	P21	23	Female	Completed Secondary	Employed part-time
P11	25	Female	University Bachelors degree	Student					

4.2.4 Procedure

Informed consent was obtained from participants after providing detailed study information. A demographic questionnaire and the SHAPE scale were then administered. Following this, participants engaged with the prototype, exploring all system configurations. The initial four conditions systematically combined the two ACTION TRIGGERS and two RETRIEVAL MODALITIES. Within each block, participants provided feedback through the System Usability Scale (SUS), the Sense of Agency Scale (SoA), and the NASA Task-Load Index (NASA-TLX) after each configuration. The fifth condition involved an open conversation scenario where participants could freely choose their preferred system configuration while collaboratively planning a trip. This condition was always presented last to maximize the ecological validity of the results. Counterbalancing of the first four conditions was ensured using a Latin Square design. Subsequently, participants completed a second administration of the SHAPE scale and participated in a semi-structured interview, providing valuable post-experiment data.

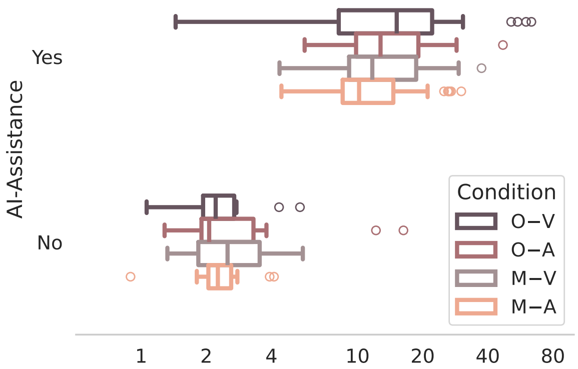


Figure 4.3: Response Delay (In seconds): All AI assistance conditions presented slower response times given the system processing requirements; however, within the AI assistance context, condition O-V presented significantly higher delays.

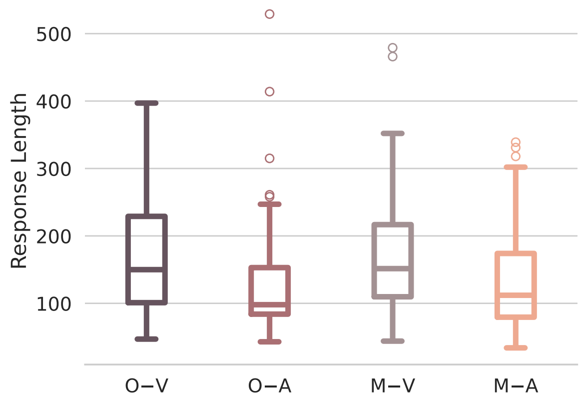


Figure 4.4: Conversation Flow plots for Response Length (In number of characters): the conditions with visual retrieval evidenced significantly longer responses

4.3 Results

In this section, we present the results of our user study. First, we share insights drawn from the Discourse Completion Task and questionnaires. Then, we detail the main themes that we identified based on the semi-structured interviews. For the sake of brevity, we abbreviate the combination of ACTION TRIGGERS and RETRIEVAL MODALITIES as M-A for mechanical trigger and auditory retrieval, M-V for mechanical trigger and visual retrieval, O-A For ocular trigger and auditory retrieval, and O-M for ocular trigger and visual retrieval.

4.3.1 Discourse Data

The data analysis for the DCT revealed non-normal distributions. To address this and ensure the robustness of our statistical tests, we employed aligned rank transformations (ART) on the data. Subsequently, we conducted two-way ANOVAs or Linear Mixed Effects Models (LME) where appropriate, followed by post-hoc tests for significant effects.

Response Delay

For response delay, the ANOVA conducted on aligned rank transformed data, utilizing Wald F tests with Kenward-Roger degrees of freedom, revealed a significant effect of Condition on Response Delay ($F(3, 160.83) = 3.25$, $p = 0.023$). Post-hoc comparisons, adjusted with the Bonferroni method for multiple comparisons, showed specific Condition differences. Notably, the O-V vs. M-A comparison revealed a significant effect on Response Delay ($est. = 24.97$, $SE = 8.75$, $df = 162$, $t.ratio = 2.85$, $p = 0.029$ adjusted). However, no other pairwise comparisons (O-V vs. O-A, O-V vs. M-V, O-A vs. M-V, O-A vs. M-A, M-V vs. M-A) reached significance after Bonferroni correction.

Response Length

In the case of response length, the ANOVA revealed a significant scenario interaction effect on response length ($F(3, 52) = 7.63$, $p < 0.001$, $\eta^2 = 0.30$). Post-hoc comparisons, adjusted for multiple comparisons with Bonferroni correction, identified significant differences in response lengths between scenarios: notably, O-V vs. M-V ($est = 32.951$, $p = 0.008$), O-A vs. M-V ($est. = -29.31$, $p = 0.021$), and M-V vs. M-A ($est. = 32.55$, p = 0.005).

Response Similarity

To understand the impact of Condition on response similarity, we used LMEs. The LME model showed an effect of Condition on Response Similarity ($F(3, 163) = 6.1$, $p < .001$). Post-hoc analyses with Bonferroni correction pinpointed specific differences: notably, scenarios M-A and O-V differed significantly ($est. = -29.09$, $p = .023$), as did M-V and O-A ($est. = 29.28$, $p = .017$) and O-A and O-V ($est. = -34.45$, $p = .004$).

4.3.2 Questionnaire Data

For the questionnaire data, we excluded 2 participants due to incomplete data. In all the questionnaires, we encountered non-normally distributed data. To address this non-normality and ensure the robustness of our statistical analyses, we applied aligned rank transformations (ART) to the data. Subsequently, we conducted two-way ANOVA tests when applicable, followed by post-hoc tests when significant effects were detected.

Statistical analyses revealed no significant differences in **Task Load** among the four conditions ($F(3, 54) = 0.68, p = 0.56$) or in the **Sense of Agency scale** ($F(3, 54) = 0.55, p = 0.64$). However, a significant effect was found for the **System Usability Scale** ($F(3, 54) = 4.1606, p =$

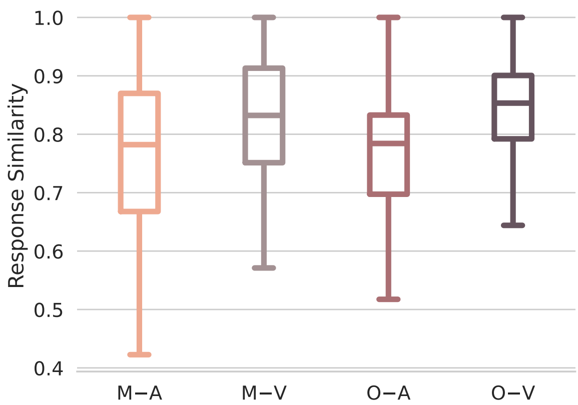


Figure 4.5: Conversation Flow plots for Response Similarity: Participants tended to follow more closely the system prompts in their responses when these prompts were presented Visually rather than Auditorily.

0.01). Post-hoc tests with Bonferroni correction indicated that mecano-visual differed significantly from O-A ($est. = 12.84, SE = 4.51, df = 54, t - ratio = 2.847, p = 0.03$) and O-V ($est. = 13.73, SE = 4.51, df = 54, t - ratio = 3.04, p = 0.02$) conditions (see Figure 4.6). No other pairwise differences reached statistical significance ($p > 0.05$).

In the pre-and post-interaction measurements of the SHAPE scale, we did not find statistically significant differences. However, regarding the Social Threat factor, 66% (N=14) of participants reported reduced or similar perceptions of augmentation technologies as a social threat. Meanwhile, 57% (N=11) indicated a decrease in the perception of individuals using augmentation technologies as being in control of their actions, whereas 21% (N=4) reported no change in this subscale (see Figure 4.7).

4.3.3 Interview Findings

Based on our qualitative inquiry, we identified three themes: (a) information moderation, integration, and balance; (b) action trigger; and (c) retrieval modality.

On a general level, our findings showcase how such a technology can support or impact the flow of a conversation, but also how individuals can adapt the use of the technology to the situation and which critical points an ideal system should address, such as correct timing and transparency to avoid disturbing the conversation. We identified mixed opinions about the interaction. While some participants felt like 'tools of the system,' others viewed the system as merely a tool they control, highlighting the variability in user experiences with cognitive augmentation. This indicates that perceived agency can vary significantly among users of a similar system. Our findings showed that the system appears to be helpful in challenging social contexts. There were two cases where the participants reported having social anxiety, the system being of extreme help given their condition, and the difficulty of finding the right words in a conversation, allowing for a continuous discussion. This illustrates that the system could potentially have assistive properties in certain scenarios. For improved readability, we slightly altered some statements (e.g. grammatical corrections), ensuring words and sentiment were maintained.

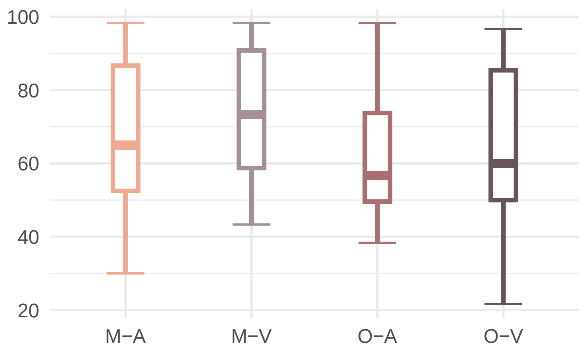
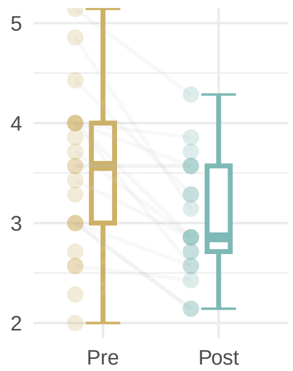
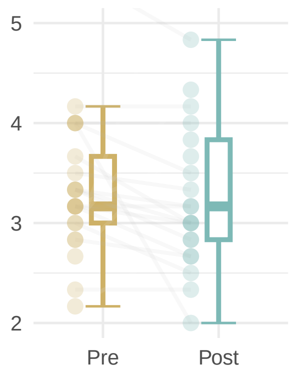


Figure 4.6: System Usability Score: the O-A and O-V presented a lower usability score.



(a) SHAPE: Social Threat



(b) SHAPE: Agency

Figure 4.7: SHAPE Scale: 66% of the participants changed their opinions about performance-enhancing technologies after interaction with the system

Information Moderation, Integration, and Balance

This theme refers to the amount of information that the system provides, the amount of information that the users integrate into their answers, and the amount of information that each agent (i.e. human and AI) contributes to the outcome, this theme has three levels:

- **Information Moderation:** Delivering an appropriate amount of information from the system to the user, avoiding both insufficient and excessive information that could lead to suboptimal user experiences.
- **Information Integration:** Describes the process of how individuals incorporate the system's prompts into their own responses.
- **Information Balance:** Balance between the amount of information and the amount of information missing (knowledge of the person/ the amount of info that the system delivers).

We observed that participants had mixed opinions regarding the amount of information that should be delivered. On the one hand, some participants highlighted the importance of having a high amount of information from which they could select and later integrate into their discourse, often correlated to the preference for the visual presentation of the information. Other participants mentioned that they would prefer a low amount of information, so there would be a lower cognitive workload on the integration process, often correlated with the preference for auditory feedback or under-use of the visual modality. The latter also use the information mostly as a trigger for themselves to elaborate on their responses.

For information integration, the participants used diverse strategies to merge the system outputs with their own thoughts. some read the prompts, some paraphrased them, and others just used specific keywords and integrated them into their responses. Some participants reported integrating information when they found it to be more natural to them or when it was relevant, on the other hand, information that was not in line with the participant's mental model was rejected.

> *So when I read some words I would never use, I would just trash it, would say, no, that's not something I would use. And if something's interesting, I would kind of think about, how I would formulate that sentence (...) and if I would use it exactly like it's suggested, I can maybe use from one word to the full sentence. Maybe I would adjust some words, depending on whether I personally would use the whole sentence.* - P15

In line with that, participants would feel in control as long as the information retrieved is coherent and plausible, but once the information breaks with these expectations, their perceived agency would be reduced.

The last level of this theme is information balance. Our findings indicate that the level of knowledge that people possess themselves determines (to some degree) if and how they offload the information. If they possessed a moderate amount of knowledge, they often integrated the additional information provided by the AI.

Other participants reported that, in the case of extensive personal knowledge, (where they had more information than the system) they would not use the system as this did not feel natural.

> *I wouldn't say I would use it for more personal conversations, it's not natural enough. It's not natural enough to say this to your friends. You wouldn't be the kind of person to say something like this blindly.* - P5

This showcases some of the strategies participants applied to integrate the system suggestions into their conversations, highlighting the relevance of the coherence of the suggestions and the contextual awareness that the system should have in order for the participants to consider the provided information. Further, the findings illustrate the need for the system to adjust to the user in terms of the amount of information delivered to keep the cognitive workload low while providing them with alternatives to select from and integrate the information.

Action Triggers

This theme refers to the experiences that participants voiced regarding the action triggers. For instance, individuals generally prefer using the button as a trigger mechanism, but for social situations, a more discreet or subtle trigger method would be preferable. The levels of this theme are the assessed triggers:

- **Mechanical**: Referring to button press trigger.
- **Ocular**: Referring to eye-tracking-based trigger.

Although we identified mixed opinions in this regard, one of the most recurrent preferred combinations was the mechanical trigger (button) in combination with the visual retrieval modality. Often linked to the concept of autonomy. For example, the mechanical trigger was described as providing more power to the user.

> “ *I have a preference for pressing the button because I think I have more control over when I want to have a suggestion.* - P1 ”

Our findings illustrate an interesting tension. Particularly, the mechanical trigger was sometimes perceived as more discreet and sometimes as more explicit. Furthermore, if the mechanical trigger was perceived as explicit, the advantage that the use of the system could be disclosed to the conversation partner was often discussed in connection with this.

> “ *So the button, I think, has the benefit of being very explicit. There you have to intentionally trigger the button. It also has the benefit of conveying to the other person that you're now using this system, which I would say is a good idea (...)* - P17 ”

A highlight of the ocular triggering method was that participants were able to perform hand gestures since there was no need to hold an additional device to trigger the system.

This elucidates the context where different types of triggering methods would be applicable, for example, using explicit methods when the person wants to inform the interlocutor about the presence of the system and using a mechanical method in stages where the user reports low agency. In contrast, implicit methods might be used when the conversation flow is prioritized over the sense of agency.

Retrieval Modality

This theme refers to the comments regarding the retrieval modality. The preference for one or the other retrieval modality was highly dependent on personal preferences. The two levels of this theme were the retrieval modalitys:

- **Auditory**: Referring to the text-to-speech method.
- **Visual**: Referring to the AR, text-based method.

While some participants felt overwhelmed with two streams of audio information (one from the system and one from their human conversation partner), others pointed out that it would be too challenging for them to read the additional information and listen to the conversation at the same time.

> *So, the audio thing obviously has the problem that I'm talking to you, which makes me want to listen to you, and that thing also is talking. It's a bit complicated to have two persons [i.e. the conversation partner and the system] talking to you at the same time, so the information coming in on the same channel makes it harder to follow that.* - P18

The other aspect participants discussed in this context was the amount of information presented. For example, some participants preferred visual information but emphasized that they struggled with the amount of text to go through. Whereas others emphasized that they could just read the text to the other person.

While other participants highlighted that audio retrieval is more effective, given that they do not have to spend time on reading and that it allows them to multitask and helps them trigger conversations.

> *I would say the audio felt nicer to me, I would say so. Because we don't really need the entire sentences. So just to start the conversation, maybe a single sentence would be enough.* - P7

The presentation method also strongly influenced the participant's sense of agency over the process. A frequent comment was that during the auditory presentation, participants had less agency given that the retrieval modality was time-locking (during the information presentation, the participant had to focus on this information exclusively), while the visual presentation allowed for more control over the process given that they were able to decide when and which parts of the retrieved information to engage with.

4.4 Discussion

We studied how the retrieval modality and action trigger impact user experience and conversational dynamics in AR-enhanced dialogue. For this, we developed a system capable of processing the contextual information of the conversation, processing the information to continue the conversation, and retrieving the information either visually or auditorily. Then, we conducted a user study using the discourse completion task (DCT), a well-known method used in pragmatics to study individuals' discourse. We then interviewed participants to gain insights into their perception of the interaction. We inferred discourse metrics from the participant responses to the DCT and conducted thematic analysis on the interview data. In this section, we synthesize the insights gained through this process and report them in two parts, one focusing on quantitative and one focusing on qualitative insights.

4.4.1 Quantitative Insights

Our study revealed that the introduction of AI-powered conversation support influences conversational flow. Notably, concerning *conversation pace,* Conversations may experience delays due to limitations in AI model processing, including limited buffer, processing power, and delays. These factors can slow down the flow of conversation. When the AI is delayed, it can cause problems for the conversation, especially if the information is relevant for continuing the conversation. The emergence of smaller, fine-tuned LLMs that run on low-power devices can potentially benefit real-time AI conversation assistance systems [302]. This trend aligns with requirements for real-world deployments, such as wearability and low processing delay [303].

Additionally, we found that participants gave significantly longer responses when using visual prompts than when using auditory ones. Our qualitative findings suggest this is because they were able to formulate their responses while processing the visual prompts at the same time. In contrast, auditory retrieval potentially requires greater cognitive effort to map the information onto an internal representation, significantly impacting both response length and adherence to the original prompt. This is further supported by the observed lower similarity between participant responses and the original prompt when presented auditorily. However, this connection has to be carefully investigated, as these measures do not account for intentionally paraphrased sentences, as this was one strategy reported by some participants. Therefore, the modality can significantly influence user engagement and response characteristics in AI-assisted conversations. In the following sections of the discussion, we touch on the qualitative insights regarding these metrics.

4.4.2 Qualitative Insights

Retrieval Modality Regarding the retrieval modality, both channels were seen as suitable to parallelize, yet auditory has the inherent disadvantage that information can overlap with the interlocutor speech. In this sense, it would be perhaps beneficial to adapt the intensity of the feedback to be subtle and not disrupt the interaction [93]. On the other hand, visual retrieval, although observed as a natural way to multi-task, was overwhelming for some participants due to the amount of information.

Autonomy, as Driving Factor of Human-AI interaction We found that when participants were put into the context of sharing cognitive tasks with the AI system, the first topic that emerged was the autonomy of humans when exposed so closely to an external cognitive agent.

We argue that this concern extends to AI systems in general, as shown in recent literature [304, 305]. Yet, this seems more crucial in the context where the AI is constantly in contact with the user while having their own thoughts and interacting in natural situations such as conversations. In line with previous work, we found that some participants felt controlled by a system, especially in situations where they blindly followed the recommendations of the system.

Effects of Action Triggers on Autonomy Participants attributed more autonomy to the more explicit mechanical trigger, as they felt a connection between their motor action of pressing a button and the information processing action of the AI system. Yet, there was a split on how and when to use each method. Participants with an inherent high agency during the interaction were comfortable using the ocular method. At the same time, some reported a loss of agency because they triggered the system unintentionally, given their natural ocular reactions. Also, the explicit trigger was seen as a way to disclose the use of the system by participants who felt it was necessary; otherwise, they would be "cheating".

Human-AI Autonomy Fluctuation Another phenomenon that we observed was that some participants started the conversation relying on the system at the beginning of the conversation, when they felt insecure about the topic, and then took over the conversation when they gained confidence. This suggests that autonomy fluctuates depending on the user's cognitive states and personality, among others, in addition to the system design.

Notably, some participants reported having less agency and having a break in the interaction whenever the recommendations were not relevant or implausible. This suggests that in the case of cognitive agency, it is necessary to maintain semantic coherence. Presumably, designing a system that is aware of the user's knowledge and that can deliver information in coherence to the user's mental model can help maintain a high sense of agency on the user. This is supported by a repetitive comment from the participants when describing the information integration in their own words: *"Is it something I would say?"*.

Information Balance Being overwhelmed by the amount of information relates to the information balance between the user's own knowledge and the system's contribution. We found that some participants relied on the system to compensate for their own lack of knowledge on specific topics and preferred to integrate more information in such cases, yet, very often, they commented that when it comes to conversations about personal matters, they will not offload any information since this information is easily accessible for them.

5

CONTROLLING FOR A PRIORI BELIEFS

This chapter is based on the following publication:

Inventory of User Expectations for Technology (iExpect) - Steeven Villa, Thomas Kosch, Agnes Mercedes Kloft, Jasper Quinn, Robin Welsch - *Computers in Human Behavior Reports (UR), Elsevier*

In this chapter, we build on previous insights that performance expectations significantly shape users' risk-taking, sense of agency, and behavior. However, assessing these expectations remains challenging due to their ambiguity. To address this, we developed and validated a questionnaire specifically designed to measure user expectations toward augmentation technologies. Our scale demonstrated strong psychometric properties, aligning with established models like the Technology Acceptance Model (TAM) while extending them by incorporating an affective dimension—anticipated enjoyment. Moreover, its predictive validity highlights its utility in capturing willingness to use technology before interaction, differentiating it from existing TAM-based measures. Beyond theoretical contributions, this instrument provides a practical tool for product development and UX research, allowing practitioners to anticipate biases in evaluation studies and refine technology adoption strategies. The following section details its development and validation process.

5.1 Constructing the Technology Expectations Questionnaire

Over the last five years, a notable trend has emerged in HCI research: the development of custom measurement instruments tailored to specific research needs, as opposed to relying solely on traditional psychometric tools from fields such as psychology and economics [Het2, 306, 307, 308]. However, as HCI continues to encounter domain-specific challenges, there has been a growing inclination to design discipline-specific questionnaires. To ensure rigor in this process, we have adhered to the best practices for questionnaire development as recommended by Boateng et al. [309], which serve as a key reference in this domain. This paper details the application of these best practices throughout our development process, as outlined in Figure 5.1.

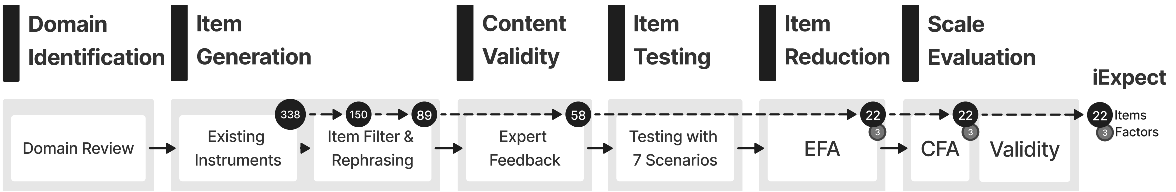


Figure 5.1: Process Overview: The creation of the iExpect was performed in accordance with the best practices suggested by Boateng et al. [309]. The numbers inside the black circles (●) describe the number of items in each stage

5.2 Item Generation

We approached the item generation in four steps: First, we identified relevant items from other instruments, second, we extracted all the items from these instruments, third, three researchers filtered the initial pool with conservative criteria based on relevance and similarity; and fourth, five researchers evaluated the remaining items based on (a) relevance, (b) formulation and rephrased and suggested items for removal.

In detail, we conducted a domain identification by collecting already existing instruments that captured a construct similar to the construct of interest: user expectations. A search in databases such as the *ACM Digital Library*, *IEEE Explore*, *ArXiv*, and *Web of Science* was conducted with the following search terms, `(expec*) AND (technolog*)`, until saturation was reached (similar items were constantly found in the search).

As a result of this process, the following established questionnaires were included: the Technology Acceptance Model (TAM) [310], the Unified Theory of Acceptance and Use of Technology (UTAUT [311, 312, 313] & UTAUT 2.0 [314]), the User Experience Questionnaire (UEQ) [315], the System Usability Scale (SUS [316] & the positive SUS [317]), the Perceived Creepiness of Technology Scale (PCTS [307]), and the modular evaluation of key Components of User Experience (meCUE [318]. Additional questionnaires from the model Motivation, Engagement and Thriving in User Experience (METUX [319]), namely the Autonomy & Competence in Technology Adoption Questionnaire (ACTA) and the Technology-based Experience Need Satisfaction (TENS) with its adaptations for the interface (TENS-Interface), the task (TENS-task) and life (TENS-Life) were also incorporated.

We collected a total of $N = 329$ items from the instruments listed above, the breakdown of the item number contributed by each instrument to the overall pool is listed in Table 5.1. During the item generation process, the researchers prioritized the relevance of the items in measuring technology expectancy prior to interaction, not posterior evaluations.

Then, a team of three researchers (1) rephrased the pool of items and (2) ranked and filtered the items based on their potential to measure technology expectancy: Items were rephrased, preserving a high similarity to the original scales while ensuring they remained technology-agnostic, thus capable of evaluating technologies before use in the broader sense. For the ranking and filtering, the criteria included relevance, potential duplication, and similarity after translation/rephrasing. Having conservative criteria, items that were literal duplicates, highly similar or out of scope were marked for removal. Then the researchers reviewed the remaining items, rating them on a scale from 1 to 5 where 1 was high relevance and 5 was low relevance. Items ranked 3 to 5 were excluded, narrowing the list to 150 items. This selection process focused on retaining only the most pertinent items for the second-to-last cut.

To refine and finalize the initial pool of items for the iExpect questionnaire, the set of 150 items was evaluated by a team of five researchers. This team included three researchers who were involved in the previous step and two researchers who were not familiar with the initial item pool. Through this evaluation process, the list was reduced to 89 items, which constituted the initial pool for the iExpect questionnaire.

Table 5.1: Summary of Questionnaires Included. Number of items by instrument with references.

Instrument	Author(s)	Items	Instrument	Author(s)	Items
ACTA	Peters, Calvo, and Ryan [319]	14	UTAUT	Madigan et al. [312]	16
Positive SUS	Lewis [317]	10	UTAUT	Maillet, Mathieu, and Sicotte [313]	40
PCTS	Woźniak et al. [307]	8	UTAUT 2.0	Venkatesh, Thong, and Xu [314]	25
TAM	Davis [310]	31	meCUE 2.0	Minge et al. [318]	35
TENS	Peters, Calvo, and Ryan [319]	37	SUS	Bangor, Kortum, and Miller [316]	10
UTAUT	Venkatesh et al. [311]	31	UEQ	Schrepp, Thomaschewski, and Hinderks [315]	72
				Total	329

5.3 Content Validity

Following the creation of the item pool, the next step involves assessing the content validity of the items to ensure they accurately measure the intended target domain. In line with Boateng et al. [309]'s recommendations, we evaluated the content validity of the 89-item pool by soliciting external expert feedback. External experts reviewed each item and rated its relevance to the construct, providing detailed feedback to refine, rephrase, add, or remove items to the pool. To measure the content validity of the pool of items, we calculated the Content Validity Index (CVI) [320, 321]. In this section, we describe the details of the process.

5.3.1 External Expert Feedback

We conducted an online survey with eleven experts to calculate the Content Validity Index (CVI). The experts were asked to rate each of the 89 items on a 4-point scale based on its relevance to the topic.

Procedure Before the item rating process, participants were informed of the study's purpose, with particular emphasis on the items being potential candidates for a scale aimed at measuring *A Priori Expectations of Novel Technology* from the user's perspective. To ensure consistent interpretation of the scale's purpose, the construct definition was displayed on every page of the survey (see Definition #1). Participants were instructed to rate the relevance of each of the 89 initial items for inclusion in the questionnaire. The items were presented in a randomized order, each accompanied by a question regarding its relevance to the intended construct. Experts rated each item on a 4-point scale: Not Relevant, Somewhat Relevant, Quite Relevant, or Very Relevant. This rating procedure was based on previous work in HCI scale development [306] and content validation research [320, 321].

Additionally, experts were given the option to provide comments on each item in a free text field. Finally, participants were asked to self-assess their expertise in related areas using 5-point scales and to provide demographic information.

> **Definition – A Priori Expectations of Novel Technology:** These are the anticipated functionalities or perceived usefulness of a technology or system that a user has not yet experienced. Since the user lacks firsthand experience or objective data about the technology, these expectations are not based on informed assessments. Instead, these assessments are constructed on assumptions, public opinions, media, marketing, or hype surrounding the technology or system. In this sense, it differs from the Expectation-Experience models as this concept is intended to capture the prior expectations of a technology and not the mismatch between a priori and a posteriori assessment of a technology.

Participants We conducted a study involving 11 experts in Human-Computer Interaction (HCI), Psychology, and Questionnaire Development. These experts were selected based on their publication records within these domains and their contributions to the ACM CHI Conference. Out of the 20 experts who were invited to participate, 11 accepted the invitation. Participation in this study was entirely voluntary, with no compensation provided. The average time required for each expert to review the complete item list was approximately 33 minutes ($M = 33.17, SD = 12.43$). Of the 11 experts, 5 identified as female, and 6 as male. Participants had an average age of 33 years old ($M = 33.09, SD = 3.01$). A detailed summary of the experts' backgrounds is provided in Table 5.2.

Data Analysis We calculated the Content Validity Index (CVI) for each item based on the ratings provided by the experts. According to the guidelines proposed by Polit and Beck [321], a threshold of 0.57 is deemed acceptable. For items that fell within a borderline range of 0.55 to 0.65, we conducted further inspections to ensure their validity. Following a consensus among the research team, considering both expert feedback and the calculated CVI values, a total of 58 items were retained for testing in the next step.

Table 5.2: Overview of the backgrounds and expertise of the 11 experts who participated in the item validation. Including scientific degree and domain as well as self-assessed expertise in psychology (Psy.), HCI, System Development (Sys), and development (Dev.) of quantitative instruments

	Gender	Age	Degree	Domain	Occupation	Psy. Exp.	HCI Exp.	Sys. Exp.	Dev. Exp.
1	Female	33	Doctoral	Res.	Researcher	■■■□□	■■■■■	■■■■□	■■□□□
2	Female	36	Doctoral	HCI	Professor	■■■■□	■■■■■	■■■■□	■■■■■
3	Male	35	Doctoral	HCI, AI	Professor	■■■□□	■■■■■	■■■■■	■■■■■
4	Male	27	Master's	HCI Res.	Researcher	■■■□□	■■■■■	■■■■□	■■■■■
5	Male	31	Master's	CS	Researcher	■■■□□	■■■■□	■■■□□	■■■□□
6	Male	37	Doctoral	HCI	Researcher	■■□□□	■■■■■	■■■■□	■■□□□
7	Female	33	Doctoral	Acad. Res.	Researcher	■■■□□	■■■■■	■■■■□	■■□□□
8	Male	29	Master's	HRI	Researcher	■■■□□	■■■■□	■■■□□	■■■■□
9	Female	34	Doctoral	HCI	Researcher	■■■■□	■■■■■	■■■■□	■■■■■
10	Female	34	Doctoral	HCI	Researcher	■■■■□	■■■■■	■■■■■	■■■■■
11	Male	35	Doctoral	CS, AI	Researcher	■□□□□	■■■■■	■■■■■	■■■■□

5.4 Item Testing

In the next phase of the scale development process, we tested the validated item set and collected data for further item reduction. Based on the recommendations of Boateng et al. [309], which suggest a minimum sample size of 200 participants to ensure robust analysis, we conducted a study with 259 participants. The participants assessed seven scenarios involving novel technologies using a set of 58 items. These scenarios were selected and validated through a focus group to ensure their suitability as stimuli for the questionnaire. In this section, we provide details about the data collection for the item testing, as well as the procedure to generate and discuss the stimuli scenarios.

5.4.1 Scenarios

We performed a targeted search for technology articles (for visual representations, see Figure 8.8 that report on novel technologies unlikely to have been tested by average users. Articles from *The Verge*[1] were chosen as references and were subsequently curated and edited to remove elements that could influence user expectations based on individual status, such as explicit pricing information. We selected scenarios specifically designed to evoke both high and low expectations concerning the presented technology and the article's writing style. In detail, we selected the following scenarios:

S1-TinyPod: A case that transforms an Apple Watch into a vintage MP3 player, providing essential functions like messaging and music playback. The article's description of the TinyPod as both a functional and aesthetically pleasing accessory may elevate expectations,

[1]https://www.theverge.com

especially among users who have tried the Apple iPod before. Yet, it can also lower user expectations as the functionalities offered are known.

S2-Fiio DM13: A portable CD player with Bluetooth support, various audio connections and compatibility with multiple digital formats. Similar to the previous scenario, this technology presents incremental innovation while simultaneously setting expectations for a modern classical device. The article's discussion sells the product as a "premium device that blends old and new technology".

S3-Sim-Lab Steering Wheel: A high-fidelity replica for sim racing, featuring detailed controls, carbon fiber construction, and real-time telemetry data display. This technology sets high expectations for an immersive and realistic racing simulation experience. The article's emphasis on the precision and quality of the wheel may heighten user expectations, selling it as a top-tier product for serious sim racers.

S4-Unistellar Envision Binoculars: Binoculars that use augmented reality to identify and label landmarks and celestial objects, connecting to a smartphone for real-time contextual information. This technology heightens user expectations by offering an advanced, interactive experience that combines outdoor exploration with digital convenience. The article's description presents the product as a cutting-edge tool for both casual and serious users.

S5-Rabbit R1d Handheld: An AI-powered handheld device with a touchscreen, camera, and scroll wheel, running on Rabbit OS to control apps and services through a single interface. This technology raises expectations by promising a simplified, unified interface for managing various digital tasks, appealing to users seeking innovation. The article's focus on the novel design and functionality of the device sets expectations for a *highly capable, all-in-one tool.*

S6-Nothing AI-Earbuds: A ChatGPT integration in the manufacturer *Nothing* earbuds, enabling direct AI interaction via a pinch-to-speak feature in the latest Nothing OS. This technology raises expectations by promising seamless access to AI-powered services through familiar devices, suggesting a future where voice interaction becomes more intuitive and integrated. The article's presentation of the integration plans sets a neutral expectation, leaving users curious about how well this feature will be implemented and function in practice.

S7-AI Pin: The Humane AI Pin is a wearable computer designed to replace smartphones by using an AI assistant and a custom operating system to eliminate the need for a screen. This technology raises user expectations by promising a more streamlined, hands-free interaction with digital tools, but current limitations may lead to user disappointment if the performance does not meet these expectations. The article's emphasis on the unfinished and problematic

Figure 5.2: Pictures of products used for the scenarios seven scenarios from their original article.

state of the AI Pin may lower user expectations, making them cautious about the product's current usability.

5.4.2 Focus Group: Scenario Validation

As there is no pre-defined way to induce people's expectations, we conducted a focus group with five participants to discuss the validity of the proposed seven scenarios and potential strategies to improve them. Participants could propose alterations to scenarios during the discussion. The goal of this step was to ensure that the selected set of scenarios effectively is suitable as stimulus for eliciting user expectations.

Procedure Initially, participants received a brief introduction to the study and completed a demographic survey. Following this, we distributed printed copies of the seven scenarios to each participant. The group discussion was structured into two phases. In the first phase, participants individually reviewed each scenario, noting their thoughts and ranking the scenarios from most to least expectation eliciting. In the second phase, we facilitated an open discussion where participants were encouraged to explain the reasoning behind their rankings. The facilitator ensured that all scenarios were thoroughly discussed. In the final stage, the group collaboratively agreed on a consensus ranking of the scenarios. Subsequently, participants were asked to propose modifications to the scenarios to enhance the range of their impact, aiming to increase the highest eliciting (positive expectations) scenarios while decreasing the impact of the lowest ones (negative expectations).

Participants Five participants (1 female, 4 male) were recruited from the facilities of Aalto University through convenience sampling to participate in an in-person group discussion. Participation was voluntary, and participants were informed that they could withdraw at any time. The mean age of the participants was 23 years old ($M = 23.60$, $SD = 2.33$). The session lasted approximately 45 minutes, and participants did not receive any compensation for their involvement.

Results We collected individual participant rankings, the overall group ranking, and participants' comments from the printed articles. These results are summarized in Table 5.3. From the independent rankings, we observed that each scenario, except for the Nothing Earbuds, was ranked among the top three in terms of positive expectations by at least one participant. The Earbuds scenario also had the lowest standard deviation, indicating consensus among participants regarding its low expectation level. In contrast, the other scenarios exhibited a higher spread, reflecting a broader range of opinions. The Binoculars and Rabbit R1 devices elicited the highest expectations according to the rankings. This evaluation confirms that the selected scenarios effectively elicit a diverse range of user expectations and are suitable as stimuli for the current questionnaire.

Table 5.3: Categorization of Scenarios, Rater scores with Average, Standard Deviation (SD), and Overall Ranking

Scenario	Code	Innovation	Writing Style	R1	R2	R3	R4	R5	Mean	SD	Rank
TinyPod	S1	Incremental	Neutral	● 3	● 1	● 2	❍ 6	❍ 7	3.8	2.32	3
Fiio DM13	S2	Incremental	Neutral	❍ 7	● 3	❍ 6	● 2	● 3	4.2	1.94	4
Sim-Lab Steering Wheel	S3	High	Heightening	● 1	❍ 5	❍ 7	❍ 7	❍ 4	4.8	2.23	5
Unistellar Binoculars	S4	High	Heightening	● 2	❍ 4	● 3	● 1	● 1	2.2	1.17	1
Rabbit R1	S5	Moderate	Heightening	❍ 6	● 2	● 1	❍ 4	● 2	3.0	1.79	2
Nothing's ChatGPT Earbuds	S6	Moderate	Neutral	❍ 4	❍ 6	❍ 4	❍ 5	❍ 6	5.0	0.89	7
Humane AI Pin	S7	High	Lowering	❍ 5	❍ 7	❍ 5	● 3	❍ 5	5.0	1.26	6

5.4.3 Procedure

Participants were first requested to provide informed consent. Following consent, we collected demographic information before presenting them with one of seven scenarios described in the survey. To ensure that the respondents had no previous experience with the devices mentioned, we included a button labeled, "I have used, tested, or experienced this technology personally." If participants clicked this button, they were shown a different scenario. If a participant reported having experienced all the technologies in the scenarios, they were excluded from the sample.

After reading a scenario, participants responded to three comprehension check questions and were then asked to explain the technology in one sentence. Subsequently, they rated 58 items on a Likert scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree).

5.4.4 Participants

To collect user data for the factor analysis, we distributed a survey to participants from the USA and UK using Prolific[2] as the recruitment platform. We collected data from native English speakers, as the current scale is developed in English. We obtained responses from 325 participants, from which, after filtering for comprehension of the text, attention checks, and click latency analysis, we removed 67 participants. This led to a total of 259 valid responses from 126 females, 129 males, and four non-binary individuals. Participants had an average

[2] https://prolific.com/

age of 37 years old ($M = 37.28, SD = 12.72$). The survey took approximately 10 minutes to complete ($M = 10.02, SD = 5.92$).

5.5 Item Reduction

In the next stage of questionnaire development, we analyzed the data obtained from the previous step to determine the optimal number of factors and reduce the number of items. This was done using Exploratory Factor Analysis (EFA), a method that reveals the underlying structure of the items by modeling the observed variables in relation to latent factors. The detailed procedure is described in the following section.

5.5.1 Item Pre-processing and Adequacy Testing

For the item analysis, we inverted the negatively worded items, examined the densities of all items, and eliminated those with high skew and kurtosis or insufficient item discrimination (n = 10). Next, we conducted a Kaiser-Meyer-Olkin (KMO) factor adequacy test to evaluate the data's suitability for factor analysis. In a KMO test, values close to 1.0 are desirable, and our dataset produced a KMO Measure of Sampling Adequacy (MSA) of 0.95. Subsequently, we performed Bartlett's Test to evaluate the null hypothesis that the inter-correlations among the variables in the dataset are equal to zero, thereby eliminating the possibility of an identity matrix and ensuring that the variables are suitable for factor analysis ($\tilde{\chi}^2(1081) = 10933.87$).

5.5.2 Exploratory Factor Analysis

We conducted an exploratory factor analysis to identify the underlying factors that explain the patterns of correlation among the items [322]. To determine the optimal number of factors, we employed both parallel analyses [323] and scree plot analysis [324]. The scree plot inspection suggested that a three-factor solution was appropriate, corresponding to factors with eigenvalues greater than 3.32.

We then used an oblimin rotation[3] to account for the possibility of correlation among factors [325]. An oblimin rotation produces correlated factors and is an oblique rotation method used in factor analysis to maximize the variance of the variable factor loadings while minimizing the number of variables with high factor loadings [325]. From this model, we eliminated all items with loadings below 0.4 and those loaded on multiple factors (above 0.2 on more than one item; n = 3). We deleted another three items with loadings that were too high or too low on the factors. We merged two items with high similarity as a final step. The scale now encompassed 22 items distributed across three factors, with nine, seven, and

[3] Note that we have tried an orthogonal rotation that had an inferior fit as compared to the oblique rotation, which did not converge in the CFA.

Table 5.4: The revised version of the iExpect scale consists of 22 items grouped in three factors: Anticipated Ease of Use (EU), Anticipated Usefulness (US), and Anticipated Enjoyment (AE), with item loadings reported. Loadings under 0.2 are omitted. Primary loadings used for scoring are in **Bold**.

Item	Code	Origin	Factor Loadings		
			EU	**US**	**AE**
Learning how to use the technology will be easy for me	EU_1	UTAUT	**0.884**		
I am knowledgeable about how to use the technology	EU_2	TAM	**0.819**		-0.131
The operating procedures will be simple to understand	EU_3	meCUE 2.0	**0.779**		
It will be quickly apparent how to use the technology	EU_4	meCUE 2.0	**0.778**		
The technology will be easy to use	EU_5	UTAUT	**0.750**	0.150	
The technology will be confusing **(R)**	EU_6	UEQ	**0.742**		
I feel confident in my abilities to use the technology efficiently	EU_7	ACTA	**0.740**		0.108
I will be able to use the technology without the support of a technical person	EU_8	Positive SUS	**0.731**	-0.186	0.138
It will be easy to get the technology to do what I want it to	EU_9	ACTA	**0.637**	0.108	0.133
The technology will make it easier to do my daily activities	US_1	UTAUT		**0.903**	
The technology will increase my productivity	US_2	UTAUT		**0.885**	
The technology will increase the quality of my work output	US_3	UTAUT		**0.842**	
The technology will be useful in my daily life	US_4	UTAUT 2.0	0.109	**0.832**	
The technology will enable me to accomplish tasks more quickly	US_5	UTAUT		**0.826**	
The technology will help me reach my goals	US_6	meCUE 2.0		**0.821**	
The technology will increase the effectiveness of performing tasks	US_7	UTAUT		**0.688**	0.150
The technology will be enjoyable	AE_1	UEQ			**0.858**
The technology will be entertaining	AE_2	UTAUT 2.0			**0.860**
The technology will be fun to use	AE_3	ACTA			**0.860**
The technology will be boring **(R)**	AE_4	UEQ	-0.117		**0.689**
The technology will be stylish	AE_5	meCUE 2.0		0.131	**0.664**
The technology will make me happy	AE_6	meCUE 2.0		0.200	**0.621**

six items, respectively. The model had a good fit, with KMO $MSA = 0.94$, Tucker Lewis Index of factoring reliability $TLI = 0.956$, and the Root Mean Square Error of Approximation $RMSEA = 0.056$. Table 7.2 presents the results of the exploratory factor analysis, namely the factors, items, and their corresponding loadings.

Based on the content of each factor, we named the first factor Anticipated Ease of Use (EU), as it includes elements related to learning how to use the technology and the perceived ease of use. The second factor, named Anticipated Usefulness (US), encompasses items that focus on how well the technology supports the user's tasks. Finally, we labeled the third factor Anticipated Enjoyment (AE), as it includes items assessing the potential enjoyment derived from using the technology. The internal consistency of the scales, as measured by Cronbach's alpha, was high: $\alpha = .932$ for EU, $\alpha = .945$ for US, and $\alpha = .908$ for AE, suggesting good internal consistency[326].

Dimension Visualization: In Figure 5.3, we show the three-factor solution scores for each scenario in a radar plot. Notably, scenarios 3 and 4 scored higher in Anticipated Usefulness, while scenarios 1 and 2 scored higher in Anticipated Ease of Use, scenarios 5, 6, and 7 had comparatively lower scores, with scenario 7 having the overall lower score.

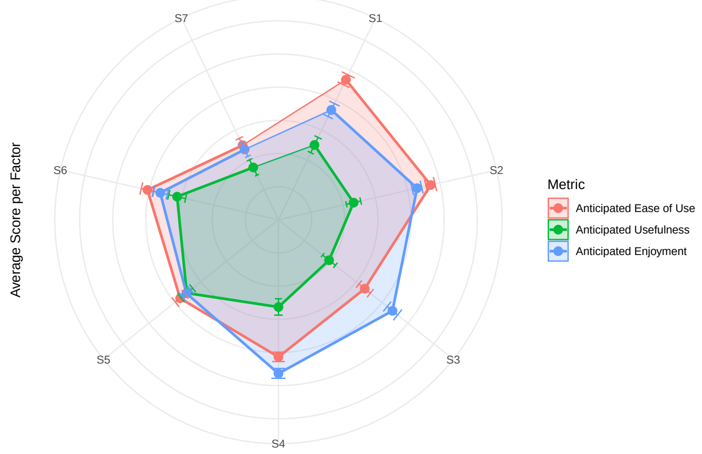


Figure 5.3: Scenario scores for the resulting dimensions of the factor analysis

5.6 Scale Evaluation

In this step, we evaluated the obtained three-factor and 22-item solution found in the previous step across three assessments: (1) We first conducted a CFA to verify the factor structure identified in the previous step; (2) To assess the temporal stability of the scale, we performed a test-retest reliability analysis; (3) We then assessed discriminant, convergent and concurrent validity to ensure that the scale measures a construct distinct from related constructs, and can predict outcomes of connected concepts. A new sample (N = 278) was gathered to test the dimensionality, and a subset (N = 93) was recruited to participate again two weeks later for the temporal stability assessment.

5.6.1 Dimensionality test

To confirm the validity of the identified factor structure, it is relevant to verify whether the dimensions of the solution remain consistent (Anticipated Ease of Use, Anticipated Usefulness, Anticipated Enjoyment). Previously, we identified a three-factor solution with 22 items. Following Boateng et al. [309], we conducted a CFA with a new participant sample ($N = 278$) to test if this structure holds.

5.6.2 Participants

To collect user data for the factor analysis, we distributed a survey to participants from the USA and UK using Prolific as the recruitment platform. We collected data from native English speakers, as the current scale is developed in English. We obtained responses from 332 participants, from which, after filtering for comprehension of the text, attention checks, and click latency analysis, we removed 54 participants. Our sample thus consists of 278 participants, including 140 females, 136 males, and 2 non-binary individuals. Participants had an average age of 37 years old ($M = 37.12, SD = 12.96$). The survey took approximately 4 minutes to complete ($M = 7.23, SD = 3.87$). The participants were compensated with 9 GBP per hour. All participants were informed of the voluntary nature of their participation and provided with the option to withdraw at any time if they felt uneasy. Participants were also informed that the collected data would be anonymized prior to processing.

Procedure Similar to the previous data collection. participants were first requested to provide informed consent. Following consent, we collected demographic information before presenting them with one of the seven scenarios described above. To ensure that the respondents had no previous experience with the devices mentioned, we included a button labeled, "I have used, tested, or experienced this technology personally." If participants clicked this button, they were shown a different scenario. If a participant reported having experienced all the technologies in the scenarios, they were excluded from the sample. After reading a scenario, participants responded to three comprehension check questions and were then asked to explain the technology in one sentence. Subsequently, they rated the 22 items on a Likert scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree).

Confirmatory Factor Analysis In order to assess the validity of the iExpect scale's structure, we conducted a CFA. This statistical procedure allowed us to confirm the dimensionality of our proposed factor model. The solution had three intercorrelated factors; see again Table 7.2. We fitted the model with lavaan Rosseel [327] using a maximum likelihood estimator with the NLMINB optimizer (56 iterations).

The confirmatory factor analysis revealed an RMSEA of $.087$, which falls within the acceptable range. Additionally, the CFI was calculated to be $.937$, and the Standardized Root Mean Square Residual (SRMR) was determined to be $.077$. Both values are within the desirable bounds, with CFI values above $.95$ and SRMR values below $.08$ being considered indicative of a good fit of the data to the model. Again, Cronbach's Alpha for EU was $\alpha = .964$, for US $\alpha = .953$, and for AE was $\alpha = .926$ can be deemed a good level of internal consistency. Providing evidence for the validity of the three-factor 22-item solution.

5.6.3 Temporal Stability

Temporal stability is an essential aspect of scale assessment, as it ensures validity for longitudinal studies and repeated measures, indicating that the instrument is not significantly influenced by external factors. Temporal stability refers to the consistency of a scale in producing similar results when administered to the same participants at different time points [309]. To evaluate the temporal stability of the iExpect solution, we conducted a test-retest reliability assessment. This psychometric evaluation, commonly employed in scale development (e.g., [307, 308]), estimates reliability based on temporal stability. We recruited 93 participants from the initial CFA data collection and analyzed their responses two weeks later to compare results across the two time points.

Procedure The procedure was identical to the one described in paragraph 5.6.2. The main difference from this procedure is that the participants were automatically assigned to the same scenario assigned in the first data collection. Therefore, participants were not allowed to change the scenario this time. Additionally, after participants completed the iExpect items, we distributed the Basic Psychological Needs Scale for Technology Use (BPN-TU) [328], the Big Five Inventory (BFI-10) [329] and the Abbreviated Technology Anxiety Scale (ATAS) [330] for discriminant and convergent validity.

Participants Approximately 33% of participants ($n = 93$) from the previous study were recruited again via Prolific for this follow-up data collection. The sample included 49 females and 44 males; none identified as non-binary or other, with a mean age of 39 years old ($M = 39.20, SD = 13.48$). The compensation and consent procedures remained consistent with those in the prior study. The survey was administered online and took respondents an average of 8 minutes to complete ($M = 8.53, SD = 4.99$).

Results To assess the reliability of the overall model, we computed the Intraclass Correlation Coefficient (ICC) using a two-way random effects model. We considered both consistency and agreement types of ICCs. Additionally, we calculated Spearman correlations for all the subscales and total scale. The consistency (ICC = 0.867, 95% CI = 0.806 to 0.91, $p < 0.001$) and the agreement (ICC = 0.867, 95% CI = 0.806 to 0.91, $p < 0.001$), and Spearman correlation ($\rho = .821$, $p < 0.001$) were found to be high for the overall scale suggesting a high degree of reliability in the overall scale items. Similarly, each of the three subscales yielded satisfactory consistency and agreement values; EU Consistency (ICC = .851, 95% CI = .783 to .899, $p < 0.001$), Agreement (ICC = .843, 95% CI = .767 to .895) and $\rho = .850$, $p < 0.001$, US Consistency (ICC = .747, 95% CI = .642 to .825, $p < 0.001$), Agreement (ICC = .749, 95% CI = .644 to .827, $p < 0.001$) and $\rho = .722$, $p < 0.001$, AE Consistency (ICC = .863, 95% CI = .800 to .907, $p < 0.001$), Agreement (ICC = .862, 95% CI = .799 to .906, $p < 0.001$) and $\rho = .837$, $p < 0.001$. Table 7.5 shows a summary of the results per factor.

Table 5.5: The two-way single-measurement intraclass correlation coefficients (ICC) calculated for Anticipated Ease of Use (EU), Anticipated Usefulness (US), and Anticipated Enjoyment (AE) factors of the Inventory of User Expectations for Technology (iExpect) scale.

Factor	iExpect-EU			iExpect-US			iExpect-AE		
	κ	F	p	κ	F	p	κ	F	p
Consistency	.851	F(92,92) = 12.4	$< .005$	.747	F(92,92) = 6.91	$< .005$	.863	F(92,92) = 13.6	$< .005$
Agreement	.843	F(92,74.1) = 12.4	$< .005$	.749	F(92,92.1) = 6.91	$< .005$	.862	F(92,92.7) = 13.6	$< .005$

To visualize the reliability of the iExpect scale, we employed the Bland-Altman method [331]. We plotted each participant's mean difference between the initial test and retest against the average of both test sessions in Bland-Altman plots. The dashed horizontal lines in these plots represent the limits of agreement, corresponding to the 95% confidence interval around the mean difference between sessions. These limits delineate the range within which 95% of the differences between test sessions are expected to fall [331, 332].

In the plot (see Figure 5.4), the mean difference near zero (dotted line) indicates that the iExpect scale demonstrates absolute temporal stability on average. The distribution around zero further suggests that reliability is not influenced by the mean score. This supports the scale's suitability for administration at different time points, making it appropriate for use in both between-groups and repeated-measures designs.

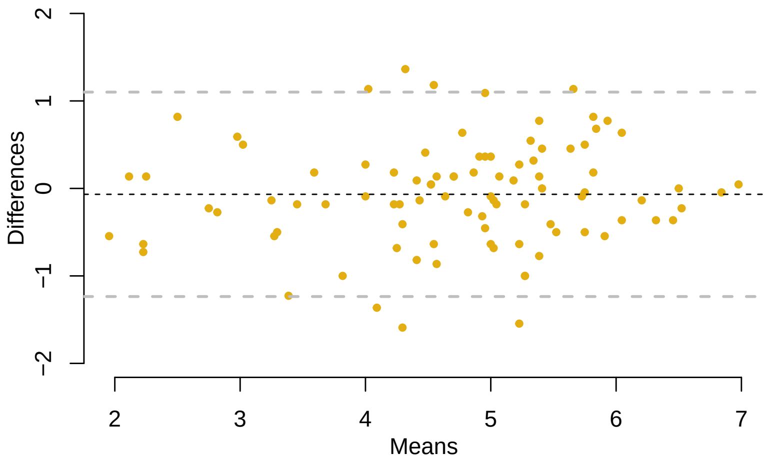


Figure 5.4: Bland Altman plot showing the iExpect Mean score for both test sessions in each participant as a function of the difference between both testings. There is no indication of skew in the data. Therefore, the small error found seems to vary independently of the Mean score.

5.6.4 Construct Validity

Following the validation of the scale's factor structure and temporal stability, one of the last steps involves assessing the construct validity, as recommended by Boateng et al. [309]. In this section, we evaluate the construct validity of iExpect across three dimensions: (1) Discriminant validity, to ensure that the construct is distinct and independent from related constructs; (2) Convergent validity, to confirm that the construct correlates with similar concepts; and (3) Concurrent validity, to demonstrate that the construct can predict the outcomes of associated concepts.

Measures: To test discriminant and convergent validity of the iExpect scale, we selected three instruments available in the literature, including all their subscales; these instruments are listed below:

BPN Basic Psychological Needs Scale for Technology Use (BPN-TU) [328]: The BPN-TU scale is a psychometric tool to assess how well the use of technology satisfies three core psychological needs: **autonomy** (BPN_{Aut}), **competence** (BPN_{Com}), and **relatedness** (to others-BPN_{Rel_O}-, and to the technology -BPN_{Rel_T}-). It is grounded in Self-Determination Theory and evaluates users' experiences with technology by measuring how these interactions fulfill their intrinsic motivations.

BFI Big Five Inventory (BFI-10) [329]: is an assessment tool used to measure five major dimensions of personality: **Extraversion** (BFI_{Extra}), **Conscientiousness** (BFI_{Consc}), **Neuroticism** (BFI_{Neuro}), **Openness** (BFI_{Open}), and **Agreeableness** (BFI_{Agre}). It provides insights into individual personality traits and is widely utilized in psychological research.

ATAS Abbreviated Technology Anxiety Scale (ATAS) [330]: is a tool that is **unidimensional** to measure an individual's anxiety or discomfort related to the use of technology. ATAS is particularly useful in contexts where understanding the psychological barriers to technology adoption and use is critical.

Given the nature of these instruments, we hypothesize that the BPN scale, particularly the BPN_{Aut} subscale, will positively correlate with the iExpect-EU. This is because the BPN_{Aut} subscale reflects the user's sense of control over the technology and their ability to use it predictably. Similarly, we expect a positive correlation between the BPN_{Com} subscale and the iExpect-US, due to the alignment with the goal-oriented anticipation measured by US.

While personality traits, as measured by the BFI-10 inventory, may influence attitudes towards novel technologies, we anticipate that the iExpect construct will be largely uncorrelated with these traits. The iExpect is specifically focused on the technology itself and the anticipation of its use, which we believe remains independent of personality traits, particularly in the pre-use stage when users lack experience with the technology. However, it is plausible that certain factors, such as BFI_{Agre}, which may involve elements of traits like trust, could show a slight correlation. This is because trust in the potential benefits of the technology may play a role in shaping expectations.

Finally, we hypothesize that the $ATAS$ scale will show a negative correlation with our construct. This expectation is based on the idea that the anticipation of benefits from using technology is inversely related to the anxiety that using technology might provoke.

Participants: The construct validity assessments were made in the same subset of participants from the temporal stability assessment presented in paragraph 5.6.3.

Internal Consistency The split-half reliability of the iExpect was evaluated using the split-half function with 1000 samples, demonstrating strong internal consistency across multiple metrics. The maximum split-half reliability (λ_4) was 0.99. Both the average split-half reliability ($\bar{\lambda}$) and Cronbach's (α) were 0.96. The minimum split-half reliability (β) was observed at 0.74.

Table 5.6: Discriminant Validity calculations for the ATAS questionnaire, the BFI-10 subscales and the iExpect factors

Scale	iExpect-EU				iExpect-US				iExpect-AE			
	$Corr$	*df*	χ^2	$RMSEA$	$Corr$	*df*	χ^2	$RMSEA$	$Corr$	*df*	χ^2	$RMSEA$
$ATAS$	-0.49	459	820.56	0.39	-0.01	459	876.33	0.87	-0.33	459	837.47	0.58
BFI_{Extr}	-0.08	302	661.62	0.82	-0.10	302	660.04	0.81	0.01	302	672.38	0.89
BFI_{Cons}	0.10	302	660.04	0.81	0.04	302	667.79	0.86	0.10	302	659.92	0.81
BFI_{Neur}	-0.24	302	641.63	0.68	-0.03	302	668.33	0.87	-0.22	302	644.39	0.70
BFI_{Open}	0.01	302	671.93	0.89	-0.11	302	658.00	0.80	-0.15	302	653.10	0.70
BFI_{Agre}	0.32	302	632.71	0.61	0.26	302	639.82	0.67	0.35	302	629.07	0.57

We also observed similar values for iExpect-EU ($\lambda_4 = .97$, $\alpha = .96$, $\overline{\lambda} = .96$, $\beta = .93$), iExpect-US ($\lambda_4 = .97$, $\alpha = .96$, $\overline{\lambda} = .96$, $\beta = .95$), and iExpect-AE ($\lambda_4 = .95$, $\alpha = .93$, $\overline{\lambda} = .93$, $\beta = .91$) indicating high internal consistency and suggesting that the items within the iExpect are reliably measuring the intended construct.

Discriminant Validity To ensure the independence of our construct from related scales, we followed the method proposed by Rönkkö and Cho [333]. We first computed models for the iExpect scale along with the previously mentioned instruments. Subsequently, for the $ATAS$ scale, we observed a negative correlation with both the iExpect-EU and iExpect-AE subscales, aligning with the anticipated relationships based on the underlying concepts of each construct. As for the BFI-10 subscales — BFI_{Extra}, BFI_{Consc}, BFI_{Neuro}, BFI_{Open}, and BFI_{Agre} — each showed small, negligible correlations with the proposed constructs, in fact, all the upper confidence intervals (CI_{Upper}) of the correlations where lower than 0.68, Rönkkö and Cho [333] suggest discriminant validity when $CI_{\text{Upper}} < 0.8$. Table 5.6 shows the computations for the discriminant validity instruments. Given these results, it is possible to assert that the iExpect scale factors are independent and distinct from connected instruments and concepts.

Table 5.7: Convergent Validity Calculations for the BPN subscales and the iExpect factors, SFL represents the Standardized Factor Loading, (which resembles the correlation between the factors)

Scale			iExpect-EU			iExpect-US			iExpect-AE		
	ω	AVE	SFL	*df*	χ^2	SFL	*df*	χ^2	SFL	*df*	χ^2
BPN_{Aut}	0.90	0.77	0.70	507	983.77	0.57	507	990.75	0.79	507	982.83
BPN_{Com}	0.87	0.70	0.59	507	988.79	0.67	507	984.21	0.78	507	982.83
BPN_{Rel_O}	0.88	0.68	0.24	507	1020.96	0.51	507	994.73	0.43	507	1001.48
BPN_{Rel_T}	0.83	0.62	0.33	507	1009.56	0.39	507	1004.71	0.48	507	996.71

Convergent Validity For the convergent validity assessment, we applied the same procedure as we did for the discriminant validity. We observed correlations across all subscales for the instruments used to assess convergent validity. Specifically, BPN_{Aut} and BPN_{Com} showed positive correlations with the three factors of the iExpect scale. BPN_{Rel_O} and BPN_{Rel_T} demonstrated weak correlations with iExpect-EU but had stronger correlations with iExpect-US and iExpect-AE, particularly with the latter. Table 5.7 provides a summary of the convergent validity results. Additionally, following Cheung et al. [334]'s recommendations, we analyzed the following criteria: (1) The ω-values > 0.7 , (2) the standardized factor loadings > 0.4, and (3) the AVE values are > 0.5. All the scales presented an ω higher than 0.7, similarly the AVE was also higher than 0.5. Regarding the standardized factor loadings, we found that BPN_{Aut} and BPN_{Com} presented a high value for all three factors, yet BPN_{Rel_O} presented a high value for US and AE mostly, while BPN_{Rel_T} presented a high value only for AE. In all, considering the redundant number of tests, the result evidences a convergent validity, highlighting that the proposed scale aligns with conceptually related instruments.

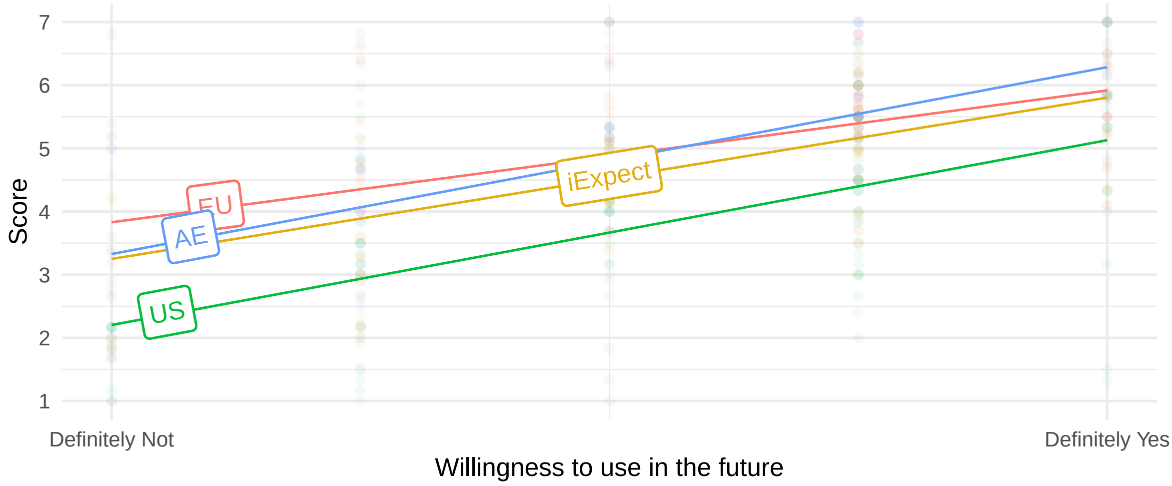


Figure 5.5: Concurrent Validity of the iExpect scale; Willingness to use a technology in the future as predicted by the iExpect scale and each of its factors (Anticipated Ease of Use (EU), Anticipated Usefulness (US), and, Anticipated Enjoyment (AE)). The results show that at a higher score of the iExpect, there is a higher willingness to use a given technology in the future as measured by the responses of the study participants

Concurrent Validity To evaluate the concurrent validity of the iExpect, we examined how effectively the scale and its factors predict the willingness to use technology in the future, Figure 5.5. A multiple linear regression analysis was conducted to assess the impact of three independent variables (EU, US, and AE) on participants' responses to the question, *"Would you be willing to use this product in the future?"*. The regression model was statistically significant, $F(3, 89) = 37.887, p < 0.001$, indicating that the independent variables collectively accounted for a significant portion of the variance in the dependent variable.

5.7 Discussion

In this paper, we introduce the development and validation process of a 22-item measure, the iExpect scale, designed to measure user expectations of technologies before interaction. We identified a three-factorial structure that encompassed the EU factor, which measures the anticipated ease of use, US, which describes the anticipated usefulness of the system, and AE that captures the anticipated enjoyment from using the interactive system.

5.7.1 Scale Scoring and Independent use of the Subscales

The iExpect scale is scored on a seven-point Likert scale from Strongly Disagree (1) to Strongly Agree (7). Items EU-6 and AE-4 are reverse-scored. Higher scores indicate higher user expectations about a given technology:

In the full scoring system of the iExpect scale, it is advisable to calculate the arithmetic mean of all the items to obtain the overall iExpect score. This scoring provides an overview of user expectations about a technology.

$$iExpect = \frac{EU + US + AE}{3}$$

Yet, we also recommend analyzing the individual factors of the scale given their individual relevance and the length of the overall questionnaire, so if a researcher would need to look at the specific insight of AE for example, distributing and computing only the elements of this subscale would be possible. In order to do this, we suggest first reversing the relevant items for EU and AE and then calculating the arithmetic mean of the items of the factor. The interpretation of the independent subscales is similar to the overall scale as they have the same valence: higher scores indicate higher expectations about a given technology.

$$EU = (EU_1 + EU_2 + EU_3 + EU_4 + EU_5 + EU_{6_R} + EU_7 + EU_8 + EU_9)/9$$
$$US = (US_1 + US_2 + US_3 + US_4 + US_5 + US_6 + US_7)/7$$
$$US = (AE_1 + AE_2 + AE_3 + AE_{4_R} + AE_5 + AE_6)/6$$

5.7.2 Scale Structure and Content

The iExpect sub-scales were validated by comparing different scenarios and in confirmatory factor analysis, showing a good fit based on several fit indices, excellent internal consistency, and very good test-retest reliability. We have evaluated the validity of iExpect across studies. In subsection 5.3.1, we could show experts' agreement regarding our tested items for the iExpect and, therefore, the validity of the content of our item pool is substantiated. Furthermore, our three-factor structure aligns with the TAM model of technology acceptance. The TAM model, originally developed by Davis [310], posits that user acceptance of technology is primarily determined by two key factors: Perceived Usefulness and Perceived Ease of Use. Our three-factor structure, encompassing Anticipated Ease of Use, Anticipated Usefulness and Anticipated Enjoyment, not only aligns with these foundational constructs but also extends the model by incorporating an affective dimension — anticipated enjoyment — which has been increasingly recognized as an important factor in user experience and technology acceptance in subsequent extensions of TAM [310, 335, 336, 337, 338, 339], such as UTAUT2 (Hedonic Motivation)[311, 312, 313, 314].

Aside from content validity, we could also establish concurrent validity. The factors of the scale could predict willingness to use technology in the future, demonstrating the criterion validity of our scales. Therefore, we deem our scale to be practically useful for capturing aspects of the intention to use the technology before any interaction, which sets it apart from the TAM-based questionnaires and scales that focus on acceptance and intention to use technology once participants have interacted with it [310, 311, 312, 313, 314]. Interestingly, in the final set of items of iExpect, there is no reference to social aspects of technology use, as covered in the meCue scale [318]. and our adaption, hinting at the possibility that before interaction, the social acceptability aspects of technology expectations are not as clear. While this view aligns with technological determinism (see for a socio-technical system discussion, see Wyatt [340]), social change coming from technological adoption, social acceptability might still be a key part of not yet-used technologies if the product vision is defined and in public discourse such as in social robotics [341] or human augmentation [Het2].

We found a distinct pattern regarding the correlation to related and unrelated measures showing convergent and discriminant validity. While we found no correlations for most BFI-10 measures and a negative correlation to the ATAS, we found that our subscales correlated with the BPN scale. Given that both the ATAS and the BPN scale relate to emotional and motivational aspects regarding the anticipation of technology use, we can speculate that expectations measured with the iExpectmight link the distal emotional and motivational aspects to expectations of interaction which are relatively proximal, i.e., map well onto the willingness to use. Moreover, while our measure was highly reliable, with very few relative and absolute changes across time, given this link to emotional and motivational aspects of technology use, the measure should be sensitive to dynamic changes in expectations across time if new information about technology is given to potential users (e.g., from sneak-peak of a product to the presentation keynote).

5.7.3 Implications

Our research has three sets of implications for practice, theory, and methods. First, with our scale at hand, practitioners can now evaluate how descriptions of different technologies, even in the prototyping stage, can elicit expectations. In product development, our scale could be used to understand aspects of purchase intentions for interactive technologies and map out which aspects are relevant to different user groups and demographics. However, it is important to note that in this regard, further research is needed to understand individual differences in expectations towards interactive technologies.

Second. with the iExpect, we have a new method. We now have a tool for measuring expectations regarding technologies before interaction. Given that scales like the ones used Brown, Venkatesh, and Goyal, Brown et al. [342, 343] have been used to anticipate technology use before interaction in practice but were constructed for acceptance after interaction and use of technology, we can resolve a long-standing mismatch in the TAM literature and provide now a new measure that is designed to measure expectations of technology.

Third, on a larger scale, the iExpect could be a new fundamental tool for evaluation efforts in HCI research as expectation-confirmation/disconfirmation Brown, Venkatesh, and Goyal [342], as well as placebo effects [Het2, 344, 345, 346], can undermine evaluation efforts. By capturing expectations upfront, the iExpect allows researchers and UX practitioners to anticipate biases (due to very high or very low expectations) that can skew evaluation results, ensuring that assessments reflect the technology's true performance, usability and user experience rather than being influenced by preconceived notions or expectations. Nevertheless, it is crucial to investigate whether these expectations stem from unrealistic perceptions, misunderstandings, exaggerated marketing, or genuine beliefs. To address this, researchers and practitioners should interview users to understand the reasons behind their expectations, ensuring that subsequent guidance or product adjustments are based on whether expectations are well-founded or if there is a misalignment between user perception and the technology's actual capabilities.

Reflections on Part II

This part explored the role of cognitive augmentation technologies in shaping user expectations, agency, and decision-making. We first examined how performance expectations influence risk-taking and the placebo effects of augmentation, highlighting the importance of user beliefs in shaping behavior. Our findings indicate that expectations alone can alter decision-making, reinforcing the need to design augmentation technologies with appropriate mental models to mitigate potential risks.

Next, we investigated the neurophysiological underpinnings of agency in cognitive augmentation. Through EEG-based spectral analysis, we explored how perceived AI assistance reduces the sense of agency at a neural level. Building on these insights, we explored augmented reality-based AI assistance in conversations. Our study revealed how different retrieval modalities and action triggers impact user experience and autonomy. We found that information balance and control mechanisms play a critical role in ensuring seamless interaction without compromising the user's sense of control.

Finally, we addressed the challenge of measuring and assessing user expectations toward augmentation technologies. We developed and validated a new questionnaire, offering a tool to quantify expectations before interaction. This scale extends traditional technology acceptance models by incorporating an affective dimension, allowing researchers and practitioners to better anticipate biases and user perceptions.

III

Augmented Human Heteronomy

6 MAPPING ATTITUDES

"These days, the problem isn't how to innovate; it's how to get society to adopt the good ideas that already exist."

– Douglas Engelbart

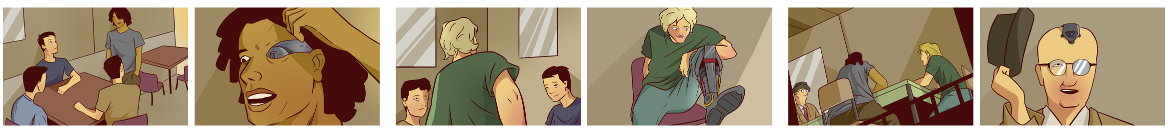
This chapter is based on the following publications:

Understanding Perception of Human Augmentation: A Mixed-Method Study - Steeven Villa, Jasmin Niess, Takuro Nakao, Jonathan Lazar, Albrecht Schmidt, Tonka-Katrin Machulla - *In 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery.*

In the previous part, we explored factors influencing autonomy. Here, we shift focus to heteronomy, or the external factors influencing augmented human autonomy, beginning with an examination of societal attitudes toward human augmentation. Human augmentation has been explored extensively in the context of technological possibilities but less so from a societal perspective. This chapter examines societal attitudes toward augmented humans and the factors influencing these perceptions across cultural contexts. To address this, we conducted a comparative study in Japan, the United States, Colombia, and Germany. The findings highlight how cultural differences shape public perceptions of human augmentation.

6.1 Method

We followed a two-stage approach to gain knowledge on the factors that influence attitudes toward augmented humans (Figure 7.1). First, we examined the dimensions influencing the



(a) Sensory augmentation vignette (b) Motor augmentation vignette (c) Cognitive augmentation vignette

Figure 6.1: Visual representation of text-based scenario applied in this inquiry. A visual representation of vignettes is only used in this manuscript for reference and to improve clarity. In the two studies reported in this work, vignettes were text-based only to avoid potential biases.

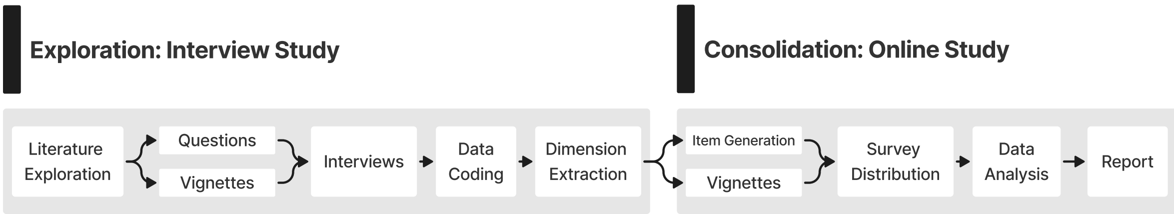


Figure 6.2: Study diagram: In the Exploration phase, we addressed **RQ1**: Which factors influence the perception of augmented humans? and **RQ2**: How do the different augmentation types affect the perception of augmented humans?

population's judgments of augmented humans through 16 interviews. Second, we examined the perception and acceptance of the three different types of augmentations in a between-subject vignette study as described in Section 6.4. By running semi-structured interviews with a diverse sample (countries from the Americas, Europe, and Asia), we followed a call to diversify HCI participant samples by Linxen et al. [347]. We aimed to gather different perspectives and concerns regarding the judgment of augmented humans. Although it does not cover all the possible opinions that can emerge from other cultures (i.e., Slavic, African, and Middle Eastern among others), we consider this an initial approximation toward comprehensive mapping of human augmentation technologies and a first step toward building an understanding about how the world perceives augmented humans. Furthermore, for the between-subject study, we included participants from the same set of countries. The study was conducted during the months of June to September of 2021 in different academic institutions across the four included countries (USA, Germany, Japan, and Colombia).

This study has been approved by the Institutional Review Board of the University of Maryland (approval number 1645727-2).

6.2 Which factors impact the perception of augmented humans? An interview study

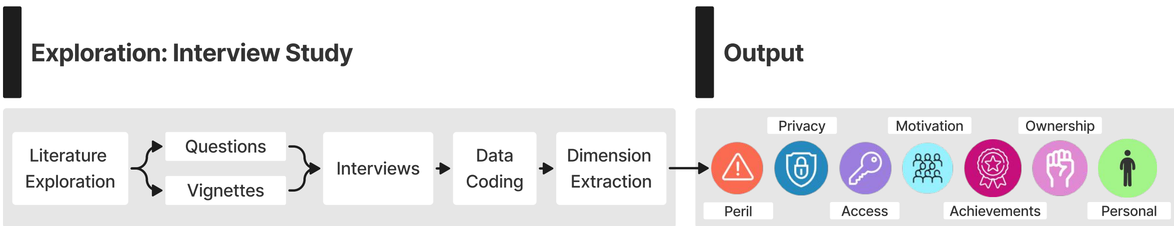


Figure 6.3: Exploration phase: we addressed **RQ1**: Which factors influence the perception of augmented humans?

We conducted 16 semi-structured interviews with participants from four countries to gain knowledge about attitudes and perceptions toward human augmentations. This constituted the first step of our inquiry and aimed to collect different perspectives of human augmentation technology users. All interviews were conducted in a one-on-one session with a single researcher and in the native language of the participant. All interviews were conducted using online video-conferencing software with audio-only recording upon receiving consent from participants (see Figure 6.3 for an overview).

6.2.1 Vignette Design A

We decided to use vignettes in both studies (in a vignette study participants are asked to see the world through the eyes of a hypothetical person in a specific scenario). Our decision is motivated by past work showing that vignette studies offer the means to balance the benefits of experimental research with high internal validity and the advantages of applied research with high external validity [348].

The choice of scenarios informed by Findler et al. [349] and Riasmo et al. [35] follows the rationale of covering human augmentation from three different standpoints: cognitive, motor, and sensory. Therefore, we developed three vignettes exhibiting the three types of augmentations [35]: *Sensory* augmentations represented by an eye augmentation, *motor* augmentations represented by augmented legs, and *cognitive* augmentations represented by a brain augmentation. We consciously decided not to describe the level of integration (implant, embodied, or environment) to learn more about participants' initial assumptions about this aspect. The augmentations used in the vignettes were brand-agnostic; this is, we did not prime the participants by mentioning any companies or form factors that can bias their answers. In the accessibility literature, it has been demonstrated that individuals have different judgments depending on the type of condition of the assessed individual (Fidler, 2007). Therefore, we wanted to individualize these three points of view to gain more detailed knowledge.

In each scenario, a man, named Michael, has a first-time encounter with another person during a social gathering. This person tells Michael that they have an augmentation and what the consequences of this augmentation are (see Figure 10.20 for visual reference). The vignettes differed regarding the type of augmentation—it either improved perception (artificial eye), cognitive abilities (brain implant), or motor skills (artificial legs). The following is an example vignette.

> “Michael went out for lunch with some friends to a coffee shop. A man with improved sight, with whom Michael is not acquainted, enters the coffee shop and joins the group. Michael is introduced to this person. Shortly after that, everyone else leaves, with only Michael and the man with the improved senses remaining alone together at the table. Michael has 15 minutes to wait for his ride. Michael has heard that it is possible to see small details at long distances and even infrared with this augmentation. Try to put yourself in the described situation and see the world through Michael's eyes.”

Table 6.1: Overview of the interview participants

ID	Country	Gender	Age	ID	Country	Gender	Age	ID	Country	Gender	Age	ID	Country	Gender	Age
C1	Colombia	Male	27	**J1**	Japan	Male	26	**G1**	Germany	Male	36	**U1**	USA	Female	50
C2	Colombia	Female	26	**J2**	Japan	Male	30	**G2**	Germany	Male	25	**U2**	USA	Female	23
C3	Colombia	Female	25	**J3**	Japan	Male	65	**G3**	Germany	Female	62	**U3**	USA	Female	23
C4	Colombia	Male	40	**J4**	Japan	Female	27	**G4**	Germany	Female	33	**U4**	USA	Male	31

6.2.2 Interview Protocol

During the interview, we first obtained demographic data. This was followed by defining the concepts "augmentation" and "bionic person" to familiarize participants with the idea and to avoid confusing similar words in the different languages. We then presented one of the three vignettes to the participants and asked them to voice their thoughts. Based on what the participants voiced, we then inquired in more detail about aspects such as interest in, avoidance of, and other thoughts on human augmentations. In the final part of the interview, we gave the participants the opportunity to ask follow-up questions and thanked them for their participation in our study.

6.2.3 Translation and Transcription:

The interview script was developed in English and subsequently translated to Japanese, German, and Spanish. The translations were executed by individuals knowledgeable in human augmentation. After the interview, we used a two-translator approach to translate and transcribe the interviews back into the English language: one translator transcribed the audio file while the second validated the accuracy of the transcription and translated it to English. After this, the first translator checked the document in English again. We repeated this process across the three non-English languages.

6.2.4 Participants

First, we had a local researcher in every sampled country, thus ensuring that every sampled country had a representative in the research team. Then, we recruited participants for the interview using the snowball strategy. Although some of the participants could also have had the interview in the English language, we opted to have all the interviews in the participant's native language and led by a researcher from the same nationality. We invited 16 participants using snowball sampling. All participants were compensated for their participation according to the average income of the respective countries. We searched for participants with heterogeneous age ranges and individuals from industry and academia from diverse subject areas. Table 6.1 presents demographics of the participants.

6.2.5 Analysis

All translated and transcribed interviews were imported into the qualitative data analysis software Atlas.ti. We conducted a thematic analysis as follows: In the first step, two researchers coded a representative sample of 25% of the material using open coding in line with Blandford et al. [30]. Then, we conducted an iterative discussion to establish an initial coding tree. The remaining transcripts were split between the two researchers and coded individually. Finally, we conducted a concluding discussion session to finalize the coding tree. This was followed by a thematic analysis to identify emerging dimensions from the material as described by Blandford et al. [30].

6.3 Interviews: Results

Here, we present the findings of our qualitative inquiry. Based on our analysis, we conceptualized six dimensions: Peril, Privacy, Access, Motivation, Ownership, and Achievement, plus an overarching topic consistently found in the interviews, which is the Personal preference to have an augmentation. Our findings are described below and illustrated with excerpts from the interviews. Each excerpt is marked with the respective participant ID.

6.3.1 Peril & Privacy

The first theme focused on the potential danger emanating from human augmentations, encompassing aspects such as human augmentations as weapons or invisible threats, privacy issues, and defense strategies. Interestingly, participants considered the potential risks of human augmentation from different perspectives. Risks caused by augmented humans as well as risks caused by the actual augmentation were discussed. The fact that it was unclear who or what was the source of risk in human augmentations (e.g., the human or the technology) and who could possibly be harmed by the augmentation (e.g., the augmented human or a non-augmented human) illustrates the complexity of the issue. The need to assess the risks posed to the augmented humans by the augmentation was emphasized by many participants. The following statement highlights this consideration:

> “For public use, the legal side of it should be checked (...) that person that is going to be using it, is this [i.e., the augmentation] going to pose any kind of threat to their body or physiology. I would like to check all this first, then I'll go for that. (U4)”

All participants discussed the potential threat to individuals or society caused by different augmentations. For example, they often either explicitly or implicitly compared them to weapons. This is highlighted by a statement of one participant discussing the potential danger of bionic vision compared with a motor augmentation:

> “When it comes to bionic vision, yes, it could interfere, but in this case (motor augmentation), it shouldn’t unless this person does, I don’t know, some sort of martial art. Their legs would allow this person to give a faster kick or a better punch. They would be stronger and hurt the person they’re fighting against. It could also be dangerous because having such strong legs could lead you to kill someone by simply kicking them. (C1)”

Another aspect that emerged from the interview data was that participants often expressed their worry about potential privacy issues related to human augmentation. Controlling other people or being controlled by other people with the means of the augmentation was a recurring topic in our interviews. On the other hand, participants who discussed the potential privacy issues of augmentations also considered potential positive aspects of human augmentations such as understanding emotions.

The need to develop a strategy to react to potential threats caused by augmentations or augmented humans was a recurring topic. One participant reflected on strategies to mitigate the potential threat of augmented humans. The scenario he described could almost be compared to an augmented human arms race:

> “Often with things like this, if there is an attack vector, a defensive strategy is developed in return. I don’t know. Other people would start bionically changing themselves too, to ensure you can’t see anything, that they hide their sweat or something or cool themselves in some other way or somehow stop the stress reaction or something. (G2)”

6.3.2 Access

Many participants reflected on the prerequisites of accessing human augmentations. Their opinions ranged from the need for all people to *have* access to human augmentations, to all people *must have* human augmentations, to people with special needs should have priority access to human augmentations:

> “No, I don’t think it’s ethical if only some people have it [i.e., the augmentation]. I believe everyone should have it. One could say: "Okay, let’s give it to the engineers and scientists since they are the ones who are technically in charge of the world’s progress and are working with all those things." But what about psychologists, philosophers, and teachers? Why shouldn’t they have it as well? Or why can’t I have it? If I’m an employee at a company, why can’t

> I have it if that would increase everyone's overall performance? Ethically, I don't think it's okay to limit this knowledge to a select group of people; we should all have it. It should even be mandatory. It is something everyone should have from day one. (C1)

Many participants discussed the need to provide access to human augmentations for everyone. Concurrently, almost all participants agreed that it would be completely acceptable if only people with special needs would have access to a specific augmentation to improve their quality of life (illustrated by the next quote). This is an interesting contradiction as both principles cannot be implemented at the same time. Consequently, based on the statements of the participants, some kind of eligibility analysis may be necessary to award human augmentations.

> If there is a certain kind of regulation that requires people to be selected, those who need it the most should come first. Naturally, if it can be used to bring people with low abilities to the level of ordinary people, I suppose those people should be given priority. (J3)

Some participants commented on the price of the augmentation. Most participants critically reflected on the potential issue that only wealthy individuals would have access to augmentation, which in turn could potentially lead to a larger socio-economic divide. However, some participants considered potential solutions, such as regulating the price of Human Augmentations:

> [Similar to] anything else someone wants really badly. The price must be within a feasible, affordable range. For example, every fool can afford a car. So that's exactly the price range that this [i.e. Human Augmentation] should be in. But I think that tough legal regulations are needed [to ensure that augmentations stay within that affordable price range], and that worldwide. (G3)

6.3.3 Motivation

The motivation to have an augmentation was discussed by the participants from two different dimensions. First, they reflected on the users' core values (e.g., socially altruistic or egoistically motivated). Second, the participants reflected on whether they would like to have a human augmentation themselves. It was particularly valuable to learn that despite general caution, some participants even expressed certain jealousy toward augmented humans. Better understanding of the perceptions and assumptions regarding users' motivations in this context is essential. These insights can then be considered in the design of future human augmentations. The augmentations that most participants assessed most positively were cognitive and motor augmentations.

> I would love such a technology. If that's something that can be done, I would be jealous of that person to be honest. [chuckles] That's the thing with me, probably I would want it too. (U3)

In contrast, some participants questioned the motives of people who would be interested in augmentations. Hence, interviewees contemplated potential reasons why someone would get an augmentation. In this context, egoistic motives and criminal intentions were discussed:

> I think there needs to be a social reason. (...) I think it's a very small minority of people who are interested in the latest technology, like 1% of all people, and most of them are people who came up with something bad. Like criminals. (J2)

There was no agreement among participants about the value of egoistic versus social motivations to get an augmentation, meaning that some participants emphasized the importance of doing something positive for society (as highlighted above). In contrast, others focused on the potential benefits for individuals. Interestingly, some participants critically questioned having an augmentation for egoistic motives. However, for most participants, the main justification for having an augmentation would be a specific need to address. For instance, many participants expressed that it would be self-evident to get a motor augmentation to mitigate the after-effects of an accident or an injury:

> That's fine. A prosthetic leg means you've lost a leg. If you can supplement what you've lost and get a higher ability [than other people], that's fine. (J4)

6.3.4 Ownership & Achievement

Another theme conceptualized based on our analysis describes augmented humans' perception on a spectrum from humans to artificial beings. Aspects such as responsibility for and ownership of augmentations (e.g., (dis)advantages in competitions), stigmatization due to augmentations, and the need to hide or disclose augmentations based on the social context were discussed. Many participants reflected on the essence of human beings and whether one or many augmentations could change this essence. Further, fear related to losing agency was mentioned by some participants. It is essential to emphasize that this dimension only takes into consideration the apparent ownership that an observed augmented human possesses, and not necessarily the actual ownership or agency that an augmented human possesses over their augmentation.

Many participants of the Colombian user group and some participants of the Japanese user group discussed social consequences such as stigma due to augmentations. Interestingly, this aspect was not mentioned by any German participant. One participant envisioned a situation where stigma would lead to a situation where a doctor would mistreat a patient that has an augmentation:

> “ If that stigma is extended to, for example, healthcare, it will be necessary to educate doctors who are going to see patients who have augmentations because there might be opinions that– For example, if they are people who are not included in a healthcare system because doctors don't have good opinions on that. "Doctor, listen, I don't know. It seems like this part of the joint in my augmentation is causing me a rash on my hip". "Oh, and who told you getting an augmentation was a good idea?". The stigma can be generalized through all different areas and I think it will be important to prepare the person who is going to have something that is different. (C3) ”

On a similar note, participants discussed making augmentations invisible to avoid stigmatization. In contrast, many participants emphasized the need to make augmentations visible for disclosure reasons and that it would be an ethical issue if the augmentation would be hidden.

One aspect that can be associated with both specific stigmatization of augmented humans and wondering about the agency is the attribution of success or achievements:

> “ It's strange to say congratulations. If a person transforms like that, I can't say he is great because it's just an ability of the prosthesis, right? (J2) ”

In the first step of our study, four dimensions were identified. These dimensions will now be explored quantitatively in the next step. For a detailed quantitative analysis, two of the identified dimensions are subdivided again so that we now have seven dimensions for the second study: PERIL: How dangerous or safe is an augmented human perceived to be? PRIVACY: Does human augmentation hinder privacy? ACCESS: Should everyone have access to augmentations or should access to augmentations be regulated? MOTIVATION: What is the motivation of the augmented human to acquire the augmentation: personal benefit or social benefit? OWNERSHIP: Are augmented humans the owners of their augmentations or do the augmentations control the augmented human? ACHIEVEMENTS: Are the achievements of augmented humans legit? PERSONAL: Do I want to acquire an augmentation? In the following section, we analyze these dimensions in more detail.

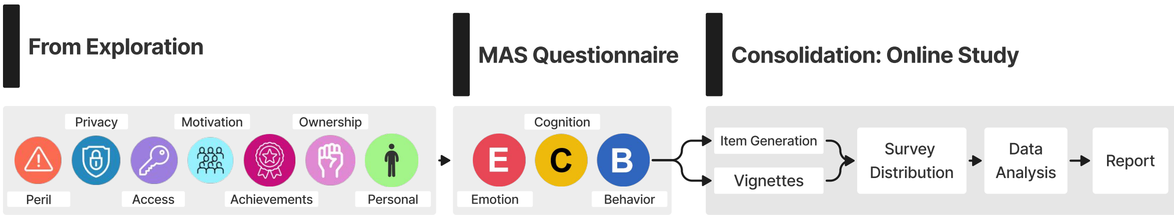


Figure 6.4: Consolidation phase: we addressed **RQ2**: How do the different augmentation types affect the perception of augmented humans? and **RQ3**: How does it affect the perception of augmented humans whether the augmented human has a disability or not?

6.4 Impact of type of augmentation and disability condition on people's perception of augmented humans: an online survey

The online study builds on the findings of the interview study and examines how the Type of augmentation and Condition of the augmented human impacts people's perceptions (Between-factors Type of augmentation and Condition). We conducted a between-subject online study using six vignettes that included one out of three Types of augmentation (Levels: Cognitive, Motor, and Sensory) and the Condition of the augmented human (Levels: Disability, No Disability). The vignettes were informed by the work of Findler et al. [349], Riasmo et al. [35], and the findings of the interview study. See Figure 6.4 for an overview.

6.4.1 Participants

We collected data from 751 participants; after filtering, we ended up with 506 respondents (50.5% female, 48.6% male, and 0.9 % non-binary) from Colombia (n = 149), Germany (n = 205), Japan (n = 65), and the USA (n = 87) (see Table 6.2). Participants' average age was 36.87 years (SD = 5.36).

We refrained from using survey platforms such as Mturk and Prolific to facilitate consistency of the sampling given that such platforms have different payment systems (incentives) and are mostly Western-oriented, the pool of participants does not cover South American or Asian countries [1] [2] and using a different platform per country could induce confounding factors into the data. Therefore, we applied the snowball strategy. We did this by contacting multiple university faculties, explaining the study's purpose, and distributing the surveys to students after contact was established. Although we know this can potentially end up sampling a specific population inside the country of origin, this is consistent across countries and is, to the best of the research team's knowledge, the most ecological way to guarantee data integrity

[1]https://www.mturk.com/help

[2]https://participant-help.prolific.co/hc/en-gb/articles/360021985613-Who-can-participate-in-studies-on-Prolific-

while sampling from diverse sources. We also filtered respondents who reported a country of origin diverging from the target samples.

We filtered the data based on the following criteria to remove random answers and bot responses:

1. Exclude responses with unrealistically short completion time.
2. Exclude responses from countries not belonging to the selected countries.
3. Exclude responses with poor open-ended questions coherence.
4. Exclude incomplete responses.

Table 6.2: Demographic distribution of the survey participants (506 responses collected via snowball sampling and survey platforms)

	Age		Gender			
	M	**SD**	**Female**	**Male**	**Non-binary**	**Total**
USA	43.65	17.12	45	42	0	87
Japan	37.2	11.2	30	35	0	65
Germany	38.0	15.7	112	92	1	205
Colombia	28.6	9.3	69	77	3	149
Total	36.8	5.36	256	246	4	506

6.4.2 Vignette Design B

Each participant was presented randomly with one of the six vignettes presenting a fictitious scenario based on the interview results and *The Multidimensional Attitudes Scale Toward Persons With Disabilities (MAS)* [349]. The story behind the fictitious scenario is similar to the one presented in section Vignette Design A. In each scenario, a man, named Michael, has a first-time encounter with another person during a social gathering. This person tells Michael that they have an augmentation and what the consequences of this augmentation are (see Figure 10.20 for visual reference). The vignettes differed in two aspects: i) the TYPE OF AUGMENTATION—it either improved perception (artificial eye), cognitive abilities (brain implant), or motor skills (artificial legs), and ii) the CONDITION of the augmented human—the person either wanted to improve their abilities because he had a disability condition or he wanted to extend their abilities beyond the normal human range. The following is an example vignette for the combination of improved perception $\times$ non-disability condition:

> Michael went out for lunch with some friends to a coffee shop. A man with an artificial eye, with whom Michael is not acquainted, enters the coffee shop and joins the group. Michael is introduced to this person. During the chat, the man

> tells them that he replaced his healthy eye with an artificial eye to augment his vision beyond the normal range. Shortly after that, everyone else leaves, with only Michael and the man with the artificial eye remaining alone together at the table. Michael has 15 minutes to wait for his ride. Michael has heard that it is possible to see small details at long distances and even infrared with this augmentation. Try to put yourself in the described situation and see the world through Michael's eyes.

In addition, we include participants' Country of Origin (country where respondent was raised) as a third factor of interest in our analysis. Thus, the survey was designed as a quasi-experiment with three independent variables (type of augmentation, Condition, and Country of origin).

6.4.3 Measures

As dependent variables, we constructed 46 Likert-type items inspired by the seven dimensions that we derived from the interview study: peril, access motivation ownership achievements, privacy and personal. Additionally, we included the *The Multidimensional Attitudes Scale Toward Persons With Disabilities (MAS)* [349] scales (emotion, cognition, behavior).

We also included a section on *Value conflicts,* contrasting three aspects: 1) Purpose of the augmentation (Individual vs. Social), Disclosure of the augmentation (Aesthetics vs. Disclosure), and Access (Augmentations should be Regulated vs. Open Access). We asked participants to indicate on binary scales which of the two values regarding augmentation they considered more important. The survey ended with the question, "What would you ask this person?" and an open text field to provide any additional comments or considerations. The survey was developed in English and subsequently translated by professional translators to the remaining languages. Afterward, authors who were native speakers of the target language double-checked the translation's consistency with the English original.

Participants filled in their demographic data and were then debriefed about their rights and the purpose of the study right after opening the survey. Afterward, they read one of the vignettes. They were then asked to take the perspective of the protagonist of the vignette to rate their agreement to statements on Likert-type items. Participants took part of the experiment voluntarily without receiving any compensation. The average survey completion time was 32 minutes.

6.5 Online Survey: Results

We applied the Aligned-Rank Transform procedure [350] to analyze whether average ratings differed across conditions. We applied this to the data before performing analyses of variance

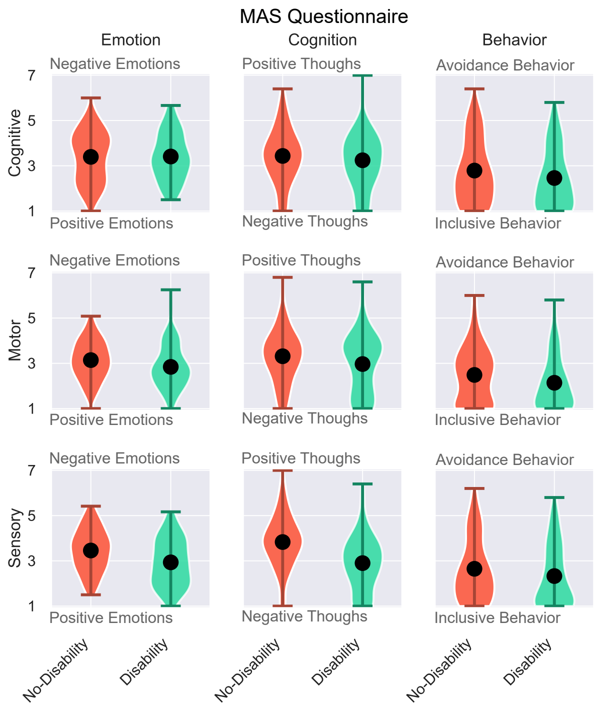


Figure 6.5: Perception toward augmented humans (MAS questionnaire dimensions): Emotion (1. Positive emotions toward augmented humans, 7. Negative emotions toward augmented humans), Cognition (1. Negative thoughts toward augmented humans, 7. Positive thoughts toward augmented humans), Behavior (1. Inclusive behavior toward augmented humans, 7. Avoidance behavior toward augmented humans).

with the between-subject factors TYPE OF AUGMENTATION (Levels: SENSORY, COGNITIVE, MOTOR). The summary of ratings for the latter can be found in Table 6.3 and Table 6.5 and the distribution can be observed in Figure 6.6 and Figure 6.5. On the other hand, we also analyzed respondents' COUNTRY OF ORIGIN (Levels: USA, JAPAN, GERMANY, COLOMBIA) influence in the perception of augmented humans . A summary of ratings for the latter is shown in Table 6.4 and Table 6.6. We further explored significant main effects using within-factor post-hoc pair-wise comparisons.

All the p-values were adjusted for the number of comparisons. Due to the high number of possible comparisons and the limited space of this paper format, we focus on the main effects and leave out interaction effects. This section also provides descriptive statistics such as the ratings' arithmetic mean (M) and the associated standard error (SE).

Table 6.3: Summary of the main effects for the the MAS questionnaire in terms of Condition and Type of Augmentation, **C** = Cognitive, **M** = Motor, **S** = Sensory

	Disability						**Type of Augmentation**											
	Yes		**No**		**Yes** vs **No**		**Cognitive**		**Motor**		**Sensory**		**C** vs **M**		**C** vs **S**		**M** vs **S**	
	M	***SE***	***M***	***SE***	***t(482)***	***p***	***M***	***SE***	***M***	***SE***	***M***	***SE***	***t(482)***	***p***	***t(482)***	***p***	***t(482)***	***p***
Emotion	3.072	0.067	3.317	0.062	2.582	**0.010**	3.397	0.082	2.997	0.073	3.181	0.079	3.228	**0.004**	1.688	0.211	-1.561	0.264
Cognition	3.041	0.082	3.535	0.078	4.115	**<0.001**	3.352	0.097	3.174	0.102	3.342	0.101	0.967	0.598	-0.060	0.998	-1.031	0.558
Behavior	2.316	0.084	2.641	0.086	2.624	**0.009**	2.633	0.107	2.322	0.098	2.478	0.108	1.446	0.318	1.092	0.520	-0.364	0.929

Table 6.4: Summary of the main effects for the MAS questionnaire in terms of Respondent's Country of Origin, **Emo** = Emotion, **Cog** = Cognition, **Beh**= Behavior

	CO		DE		JP		US		CO vs DE		CO vs JP		CO vs US		DE vs JP		DE vs US		JP vs US	
	M	SE	M	SE	M	SE	M	SE	t(482)	p	t(482)	p	t(482)	p	t(482)	p	t(482)	p	t(482)	p
Emo	2.936	0.087	3.309	0.069	3.418	0.126	3.211	0.107	-3.516	**0.003**	-3.608	**0.002**	-1.911	0.225	-1.112	0.683	0.936	0.786	1.694	0.328
Cog	3.183	0.120	3.358	0.085	3.452	0.149	3.202	0.130	-0.826	0.842	-1.302	0.562	0.113	0.999	-0.735	0.883	0.813	0.848	1.272	0.581
Beh	2.274	0.097	2.444	0.099	2.763	0.160	2.715	0.164	-0.886	0.812	-2.657	**0.040**	-2.339	0.091	-2.106	0.153	-1.724	0.312	0.481	0.963

6.5.1 General Assessment

Value conflicts and the subscales of the MAS scale describe in a broad sense the interaction with an augmented human but also enforce the respondent with a set of fundamental questions: Should the use of augmentation devices be regulated or free? What emotions, thoughts, and behaviors would someone experience or execute when interacting with a human augmentation user? We address these questions in this subsection.

Value conflicts We presented participants with three value conflicts. Each time, they had to choose which of two values they considered more important. Regarding the first conflict—using augmentations for personal improvement vs. augmentation for improving society—more participants chose the latter option (61.8%). Colombia and Germany, in contrast with Japan and the USA, gave priority to the social role of augmentation devices ($\tilde{\chi}^2(3) = 17.69$, $p < 0.001$). Regarding the second conflict—aesthetic appearance (Augmentation not explicitly visible) of the user vs. disclosure (augmentations should be visible)—more participants indicated to prefer an aesthetic appearance (63%). No significant difference was found between countries. Lastly, participants were asked to choose between regulating access to augmentations vs. providing access to all. There was a slight majority in favor of regulation (52.7%).

Emotion The EMOTION subscale of the MAS questionnaire analyzes the tendency of an individual to elicit positive or negative emotions on the respondent. A lower value represents a tendency toward positive emotion. The condition of the augmented human significantly impacted the emotions towards the augmented humans.

In this sense, the pair COGNITIVE-MOTOR augmentation was the only pair with significant contrast, in which COGNITIVE augmentation (M = 3.397, SE = 0.082) tended to elicit more negative emotions than the MOTOR augmentation (M = 2.997, SE = 0.073; $t(482) = 3.228$ $p < 0.01$). Seen from the COUNTRY OF ORIGIN lens, COLOMBIA (M = 2.936, SE = 0.087) had the most positive emotions. We found significant differences in the pairs COLOMBIA-GERMANY ($t(482) = 3.516$ $p < 0.001$) and COLOMBIA-JAPAN ($t(482) = 3.608$ $p < 0.001$).

Cognition The COGNITION subscale of the MAS QUESTIONNAIRE analyzes the tendency of the observer to have negative or positive thoughts regarding an individual. The tests applied on the data failed to find a main effect in terms of type of augmentation or country of origin of the respondent.

Behavior The BEHAVIOR subscale of the MAS QUESTIONNAIRE analyzes the inclusive or avoidance behavior of the observer regarding an individual. No effect was observed on the TYPE OF AUGMENTATION regarding the respondent COUNTRY OF ORIGIN; only the pair COLOMBIA-JAPAN was significant ($t(482) = 2.657$, $p < 0.05$), with COLOMBIA being the country with the most inclusive behavior ($M = 2.274$, $SE = 0.097$). However, all countries leaned toward inclusive behavior.

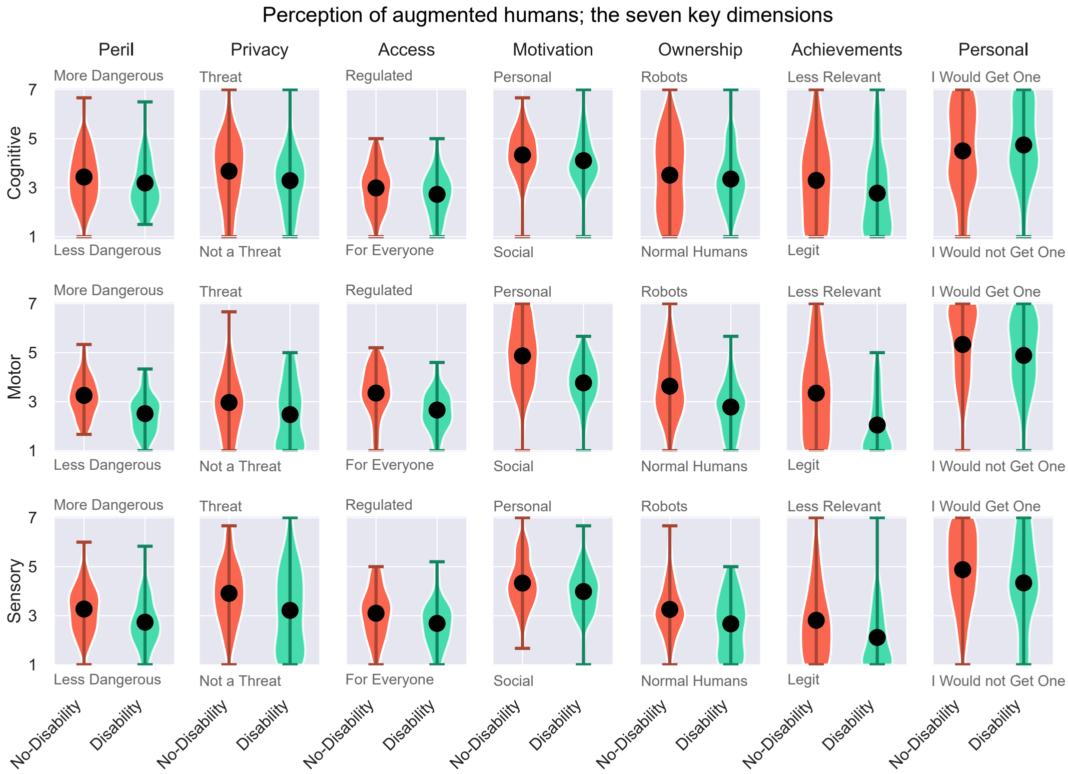


Figure 6.6: Perception toward augmented humans: Perilous perception (1. Less dangerous, 7. More Dangerous); Privacy (1. Augmented humans are not a threat for privacy, 7. Augmented humans are a threat for privacy); Access (1. Augmentations should be available for everyone, 7. Augmentations should be regulated); Motivation (1. Individuals use augmentation for social benefit, 7. Individuals use augmentations for personal benefit); Ownership (1. Augmented humans are normal humans, 7. Augmented humans are robots); Achievements (1. Augmented humans' achievements are legit, 7. Augmented humans' achievements are less relevant); Personal (1. I would not like to have an augmentation, 7. I would like to have an augmentation). The black circle in the graphs indicates the distribution's mean value.

6.5.2 Dimensions

In this section, we report the results for the set of dimensions extracted from section 6.2 (see Figure 6.3 for reference). These dimensions answer questions such as: How dangerous is an augmented human perceived to be? Should everyone have access to augmentations? Should augmentations have a social or individual purpose? Does human augmentation hinder privacy? Are augmented humans the owners of their augmentations or do the augmentations turn them into robots? Are the achievements of augmented humans legit? Would the observer want to have an augmentation?. **The collected data yielded that the previous CONDITION (Disability, No-disability) of the augmented human (observed human) impacted all the dimensions presented below with the exception of the personal preference** for acquiring an augmentation (see Table 6.5 for a summary). This behavior is coherent with the formulation of the dimension given that it does not reference the observed human but the observer. In the following, we report the results for every dimensions in terms of type of augmentation and country of origin of the observer. In our analysis, we also accounted for technological preference, but no significant changes occurred.

Peril The PERIL dimension analyzes how much an observed augmented human is perceived to be dangerous, with higher values representing a higher perception of threat. Our sample yielded that the perceived threat posed by an augmented human is modulated by the TYPE OF AUGMENTATION and COUNTRY OF ORIGIN. COGNITIVE augmented individuals were considered more dangerous ($M = 3.310$, $SE = 0.088$) than MOTOR ($M = 2.910$, $SE = 0.071$; $t(482) = 3.161$, $p < 0.01$) and SENSORY augmented humans ($M = 2.984$, $SE = 0.079$; $t(482) = 2.462$, $p < 0.05$). Country-wise, only the pair COLOMBIA-GERMANY was significantly different ($t(482) = 2.595$, $p < 0.05$) with COLOMBIA considering augmented humans as less dangerous ($M = 2.926$, $SE = 0.088$).

Privacy The PRIVACY dimension analyzes the extent to which augmented humans are a threat for the observer's privacy, where a lower value means a lower perception of threat for the respondent's privacy. The sampled data show an influence of the TYPE OF AUGMENTATION and the respondent's *country of origin*. Participants rated SENSORY ($M = 3.557$, $SE = 0.116$) and COGNITIVE ($M = 3.477$, $SE = 0.103$) augmented humans as a greater threat to their privacy than MOTOR augmented humans ($M = 2.722$, $SE = 0.102$), with values of ($t(482) = 3.828$, $p < 0.001$) for the pair COGNITIVE-MOTOR and ($t(482) = 5.125$, $p < 0.001$) for the pair SENSORY-MOTOR.

Respondents from GERMANY consistently rated augmented humans as a threat for their privacy in comparison with the rest of the sample (see Table 6.6).

Access The ACCESS dimension analyzes the observer's opinions about regulations of human augmentations, where a lower value means a preference toward universal availability of augmentations. The sampled data did not show any influence of the TYPE OF AUGMENTATION in the respondent judgment of openness or regulation of human augmentation. The COUNTRY OF ORIGIN of the respondent only yielded a significant difference in the pair JAPAN-USA ($t(482) = 2.854$, $p < 0.05$), with USA respondents leaning toward universal availability of augmentations ($M = 2.703$, $SE = 0.117$) and Japan being more conservative than the rest of the sample ($M=3.169$, $SE=0.083$) . However, all the countries remain on the universal availability side.

Table 6.5: Summary of the main effects for all the explored dimensions in terms of Condition and Type of Augmentation, **C** = Cognitive, **M** = Motor, **S** = Sensory

	Disability						Type of Augmentation											
	Yes		No		Yes vs No		Cognitive		Motor		Sensory		C vs M		C vs S		M vs S	
	M	*SE*	*M*	*SE*	*t(482)*	*p*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*t(482)*	*p*	*t(482)*	*p*	*t(482)*	*p*
Perilous	2.827	0.065	3.311	0.061	5.401	**<0.001**	3.310	0.084	2.910	0.071	2.984	0.079	3.161	**0.005**	2.462	**0.038**	-0.722	0.751
Access	2.691	0.056	3.125	0.064	4.912	**<0.001**	2.853	0.072	3.019	0.079	2.868	0.077	-1.700	0.206	-0.139	0.989	1.569	0.260
Motivation	3.960	0.063	4.520	0.075	5.686	**<0.001**	4.211	0.079	4.362	0.105	4.167	0.079	-1.537	0.275	0.340	0.938	1.883	0.145
Privacy	3.015	0.092	3.502	0.087	3.774	**<0.001**	3.477	0.103	2.722	0.102	3.557	0.116	4.838	**<0.001**	-0.265	0.962	-5.125	**<0.001**
Ownership	2.941	0.075	3.450	0.086	3.856	**<0.001**	3.429	0.105	3.230	0.102	2.935	0.092	1.443	0.320	3.419	**0.002**	1.957	0.124
Achievement	2.323	0.092	3.142	0.106	5.659	**<0.001**	3.023	0.129	2.753	0.132	2.435	0.112	0.889	0.648	3.148	**0.005**	2.244	0.065
Personal	4.647	0.108	4.901	0.107	1.613	0.107	4.635	0.133	5.136	0.125	4.576	0.132	-2.444	**0.039**	0.385	0.922	2.839	**0.013**

Table 6.6: Summary of the main effects for all the explored dimensions in terms of the Respondent's Country of Origin, **Per** = Perilous, **Acc** = Access, **Mot** = Motivation, **Priv** = Privacy, **Own** = Ownership, **Ach** = Achievements, **Pers** = Personal Preference

	CO		DE		JP		US		CO vs DE		CO vs JP		CO vs US		DE vs JP		DE vs US		JP vs US	
	M	SE	M	SE	M	SE	M	SE	t(482)	p	t(482)	p	t(482)	p	t(482)	p	t(482)	p	t(482)	p
Per	2.926	0.088	3.161	0.071	3.156	0.132	3.054	0.099	-2.595	**0.048**	-1.907	0.226	-1.294	0.567	-0.031	1.000	0.814	0.848	0.662	0.911
Acc	2.842	0.083	2.968	0.068	3.169	0.083	2.703	0.117	-1.218	0.616	-2.465	0.067	0.755	0.875	-1.654	0.349	1.819	0.266	2.854	**0.023**
Mot	4.246	0.094	4.437	0.080	3.544	0.118	4.310	0.113	-1.855	0.249	4.839	**<0.001**	-0.309	0.990	6.456	**<0.001**	1.231	0.607	-4.637	**<0.001**
Priv	2.978	0.115	3.750	0.093	3.051	0.175	2.759	0.157	-5.862	**<0.001**	-0.518	0.955	0.987	0.757	3.890	**<0.001**	5.962	**<0.001**	1.281	0.575
Own	2.978	0.115	3.750	0.093	3.051	0.175	2.759	0.157	-5.862	**<0.001**	-0.518	0.955	0.987	0.757	3.890	**<0.001**	5.962	**<0.001**	1.281	0.575
Ach	2.651	0.137	2.954	0.114	2.062	0.141	2.891	0.187	-1.775	0.287	2.665	**0.040**	-0.944	0.781	4.125	**<0.001**	0.494	0.960	-3.191	**0.008**
Pers	4.369	0.151	5.268	0.109	4.208	0.181	4.736	0.179	-4.684	**<0.001**	1.180	0.640	-1.152	0.658	4.773	**<0.001**	2.718	**0.034**	-2.017	0.183

Achievement The ACHIEVEMENT dimension analyzes the respondent's perception of the achievements of an augmented human, where a lower value represents a higher validation of the augmented human achievements.

The respondent's *country of origin* and the *Type of Augmentation* had an influence on respondents' perception of achievements of the augmented human. In detail, a SENSORY augmented individual's achievements were regarded as the more legit among the three types of augmentation. Respondents from GERMANY were the most skeptical about achievements attained with the help of augmentations ($M = 2.954$, $SE = 0.114$), followed by those from the USA ($M = 2.891$, $SE = 0.117$), COLOMBIA ($M = 2.651$, $SE = 0.137$), and JAPAN ($M = 2.062$, $SE = 0.114$). Respondents from JAPAN particularly, seemed to validate more augmented human's achievement than the rest of the sample (for more detail, see Table 6.6).

Motivation The MOTIVATION dimension analyzes the respondent's perspective on the motivation that an augmented human had to acquire a given augmentation. It does so in a continuum from social focus (1 in the scale) to individual focus (7 in the scale). How strongly a user of augmentation is perceived to act with an individual or social intention was impacted by the participants' COUNTRY OF ORIGIN but not significantly by the TYPE OF AUGMENTATION.

In this dimension, a clear difference was noted from the respondents from JAPAN regarding the rest of the sample, with social motivation as the perceived motivation. The rest of the sample interpreted a personal motivation (refer to Table 6.6 for details).

Ownership The OWNERSHIP dimension analyzes the extent to which an augmented human is still perceived as having agency over the augmentation (Owning the augmentation) or the augmentation having agency over the human (being a computer, robot, machine). Lower scores represent that the augmented human preserves the agency. In this dimension, our sample yielded that the TYPE OF AUGMENTATION and COUNTRY OF ORIGIN of the respondent impacted the perception of ownership over the augmentation. In detail, COGNITIVE augmentation had the highest impact on ownership perception ($M = 3.429$, $SE = 0.105$). Although respondents of all countries leaned toward augmented humans being the owners of the augmentation, GERMAN respondents were significantly more conservative ($M = 3.750$, $SE = 0.093$) (please refer to Table 6.6 for details on the contrasts).

Personal preference The PERSONAL PREFERENCE subscale addresses a respondent's willingness to acquire a given augmentation; higher values mean higher inclination toward acquiring the augmentation. Based on our data, participants stated a higher interest in obtaining a SENSORY ($M = 4.635$, $SE = 0.133$, $t(482) = 2.839$, $p < 0.05$) or COGNITIVE ($M = 4.576$, $SE = 0.132$; $t(482) = 2.444$, $p < 0.05$) augmentation compared to a MOTOR augmentation ($M = 5.136$, $SE = 0.125$). Country-wise, JAPANESE respondents reported a higher willingness to acquire an augmentation ($M = 4.208$, $SE = 0.181$) in contrast to GERMAN respondents that were the less interested on acquiring one for themselves ($M = 5.268$, $SE = 0.109$; $t(482) = 4.773$, $p < 0.001$); in this regard also, USA ($M = 4.736$ $SE = 0.179$;$t(482) = 2.718$, $p < 0.05$) and COLOMBIAN participants ($M = 4.369$ $SE = 0.151$; $t(482) = 4.684$, $p < 0.001$) followed the trend of JAPAN, being positive toward acquiring an augmentation.

6.6 Discussion

In this section, we first provide answers to our research questions, then we discuss the general perception of augmented humans based on the analysis of our interviews and the results of the online study. We then outline insights following the structure of the dimensions identified. We adapted the MAS questionnaire [349] to measure attitudes toward augmented humans. The questionnaire focuses on three aspects: behaviors, cognition, and affects. The questions concerning behavior focus primarily on avoidance behaviors such as leaving the room the augmented human is in or moving to another space. Cognition mainly focuses on aspects concerning interest in and the first impression of people. This subscale includes questions such as if someone looks interesting or if the participant would like to get to know the augmented human more. Affects focus on affective experiences such as fear, depression, relaxation, or shame.

This study set out to identify factors relevant for the assessment of augmented humans (RQ1). In addition, we analyzed how the type of augmentation and user's disability condition impact the perception of augmented humans (RQ2). With regards to the question *Which factors influence the perception of augmented humans?* (RQ1), our results show that the following six dimensions modulate the perception of augmented humans: peril, privacy, access, motivation, ownership, and achievement . Furthermore, for the question *How do the different augmentation types affect the perception of augmented humans?*, we found that the type of augmentation had an impact on the perception of all dimensions apart from access and motivation (RQ2). Finally for the question *How does it affect the perception of augmented humans whether the augmented human has a disability or not?*, we found that the previous disability CONDITION of the augmented human was the most decisive factor across all our samples; nearly every dimension was impacted depending on whether the individual in question had a disability before acquiring the augmentation (RQ3).

6.6.1 What is The Current Perception of Augmented Humans?

Our analysis, based on multiple data-sources, showcases some interesting tensions. While being generally optimistic about augmentations, respondents reported not wanting an augmentation for themselves. This opinion was shared across every sampled COUNTRY OF ORIGIN with participants from Germany being most skeptical. Based on the adapted MAS questionnaire results (subsections BEHAVIOR and EMOTION), we observed that our sample was mainly positive about the augmented human described in the vignettes regardless of their CONDITION, the TYPE OF AUGMENTATION, and the COUNTRY OF ORIGIN of the participants. However, the COGNITION dimension of the MAS QUESTIONNAIRE showed that our sample tended to have negative thoughts toward augmented humans. Cognitive augmentations were the most controversial augmentation type. Cognitive augmented humans, which elicited the least positive emotions, were seen as the most perilous and the ones reduced the perception of ownership and achievement the most. Further, cognitive augmentation was seen as more dangerous than, for example, motor augmentations, where the augmentation itself could be used to induce physical damage to someone else. Notably, motor augmentations were seen as the least dangerous in terms of privacy and peril, but also the least wanted of the three augmentations. At the same time, motor augmentations elicited the least negative emotions and behaviors. This result could be explained with the high correlation of motor augmentation devices and assistive devices for people with mobility restrictions, and the bias against assistive device adoption [351]. Finally, sensory augmentations seemed to only be perceived less positively when it comes to privacy.

6.6.2 Cross-cultural Aspects

Following a call from Linxen et al. [347], we contribute a study with a geographically diverse sample. The attitude of responses of participants from Japan were significantly different from

German ones in five out of seven dimensions, making their opinions the most contrasted, whereas participants from Colombia and the USA did not present any significant difference.

Although there was a general agreement across all the sample, the extent to which respondents from every country scored augmented humans was significantly different, for example, in the case of privacy, the four countries sampled leaned towards augmented humans not being a threat for privacy, however Germany respondents were significantly less inclined to this judgment (refer to Table 6.6), this behavior also occurred in the Ownership dimension for Germany, in the Achievements dimension for Japan where respondents from Japan were the ones that validated the most augmented human's achievements, and, in the Personal dimension with Germany, where German respondents were the most reluctant to acquire an augmentation for themselves. Interestingly, the only dimension where there was a disagreement is the Motivation to acquire an augmentation; Japanese respondents leaned significantly towards the social use of augmentations while the rest of the sample did it for the personal use. This aligns with the opinions reported in the interview study where an interviewee from Japan mentioned that he cannot imagine human augmentations not being used for social benefit.

It is plausible that elements inherent to the country of origin, which were not accounted for in the main set of control variables, can have an effect. This is particularly intriguing in the case of the education level of participants or the level of exposure to emerging technologies. These things can affect how a person understands the scenarios they are given and, in turn, how they reported their opinions in the survey study.

6.6.3 Design Recommendations

In this section, we assess our findings through the lens of interaction design and give a list of design recommendations, highlighted in bold, along with evidence to support each recommendation.

Our results extend human enhancement literature [352, 353]. Our findings show that safety concerns regarding human augmentations concern two aspects. While previous work showed that the main concerns lie in the safety of the person undergoing an enhancement or intervention, our participants were concerned about the risk associated with getting an augmentation. Furthermore, the augmented human is also regarded as a potential threat to the individuals in their environment. Some participants even suggested that augmentations should be regulated in the same way as guns or weapons. Across all four countries, respondents were more restrictive about the adoption of augmentations by persons without disabilities. In addition, participants reflected on the potential threat of different augmentation types as illustrated by the artificial eye example. Participants speculated that augmented vision could enable individuals to identify physiological reactions that are not evident without the use of technology and, thus, have more information about the people in their environment. While our qualitative analysis revealed ways in which augmented humans can be perceived as a threat to one's safety, our quantitative results showed that in the case of privacy, the four countries sampled leaned toward augmented humans not being a threat for privacy; however, German respondents were significantly less inclined to this judgment (refer to Table 6.6). In general, participants emphasized the need to communicate the purpose of the human augmentation in a clear manner. This is in line with previous work [67]. Uncertainty about the purpose of human augmentations can lead to speculation and fear. **Consequently, communicating the application area or the purpose of augmentation through a clear and unambiguous design could help mitigate the population's concerns about human augmentations.**

The population sampled in this study converged in that augmentations should be available for everyone and not regulated. However, in case access to augmentations is restricted, participants favored prioritizing access to augmentations for people with disabilities, particularly if augmentations extended sensory and motor abilities related to body strength and endurance. The motivation to acquire an augmentation was a recurrent topic. Participants had strong opinions regarding the motivation to get an augmentation; one participant even suggested that people using augmentations for egoistic purposes have criminal intentions. Participants from Japan and Colombia assumed that augmentations are used for a social rather than individual benefit. Moreover, our results showed that the condition and the motivation for getting an augmentation strongly influenced people's attitudes toward human augmentation. Consequently, based on previous work and our analysis, assistive systems seem to be perceived as more acceptable than human augmentations designed for people without previous disability conditions. Therefore, an approachable human augmentation should offer flexible design solutions that can be adapted and used by individuals with different abilities and needs. This is in line with the vision of "assistive augmentations" introduced by Huber et al. [24]. The benefit of this approach could be twofold. People with impairments often reject using assistive technologies to avoid appearing "different." **We propose to design augmented systems that can potentially be used to address a variety of different user needs,. In other words, the design of hybrid augmentations (augmentations that are also built for assistive functions) can be used to address the challenge of making human augmentations acceptable and more inclusive by designing for a spectrum of abilities.**

Moreover, several participants addressed the topic of achievements in the interview study. Perceptions about the weakening of the importance of accomplishments because of the usage of augmentations were extensively discussed in the interviews. Yet, the quantitative data show that respondents tended to judge augmented humans as regular humans and saw their ACHIEVEMENTS as legit regardless of their augmentations. Participants from Japan valued the achievements of augmented humans the most. Participants from Germany valued the achievements of augmented humans the least. Another factor present in the discussion was losing agency after assimilating an augmentation. This recalls Anderson's [354] suggestion that *"Some augmentations may have profound effects on a person's sense of self."* This factor plays a role in the perception of the achievements of the user of the augmentation. There seemed to be a continuum between the joy of having augmented skills, perceiving an augmented human as human (including their achievements), and the need only to receive recognition when one deserves it. While participants expressed interest in experiencing an augmentation, our results indicate that achievements would be worth less if they were achieved while having an augmentation. Participants commented that augmented humans' accomplishments could not be considered their own but instead as the augmentation system achievements. This exposes a trade-off between the system's performance and the level of effort invested by the user. **We recommend paying special attention to navigating effort and effortlessness when designing human augmentations. Reducing the user's effort to the minimum would lead to lowering agency in the user and reducing the importance of accomplishments achieved while using the augmentation.**

6.6.4 Human Enhancement From the HCI Lens: Multifold Dimensions

In contrast with human enhancement literature [352, 353], the threat perception of the intervention (Augmentation, or Enhancement) does not only include the integrity of the individual receiving the intervention but also the individuals in the surroundings. This phenomenon holds for two of our seven dimensions, namely PERILOUS and PRIVACY; During the interviews, our interviewees mentioned the associated of getting an augmentation. For example, in the case of implants, the risk of a wrong intervention, issues with bio-compatibility, or related would indeed impact the individual integrity. However, it unfolded another perspective, which is the threat that a person with augmented skills can pose if these skills are misused, it can be depicted with the case of motor augmentations, where an individual can increase their strength and use it against their peers. Such a situation was mentioned by one interviewee, who even suggested that motor augmentations should be regulated in the same way as weapons, given their potential. This also applies to the Privacy dimension; participants were worried about augmentation manufacturers having access to their data on a more intimate level, given that augmentations would integrate more closely with their bodies and, in a far too futuristic scenario, with their brains, therefore it is at least a reasonable concern. However, what is more interesting is the perceived privacy threat derived of the use of an augmentation; Participants also reported that an augmented human could potentially violate their privacy by making use of, for example, sensory augmentations that can reveal physiological reactions that are not evident without the use of technology.

7 MEASURING ATTITUDES

*"If humans were totally unstructured creatures, they would be... a tool which can properly be **shaped** by outside forces."*

– Noam Chomsky

This chapter is based on the following publications:

Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale - Steeven Villa, Jasmin Niess, Albrecht Schmidt, Robin Welsch - *In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), Volume 7, Issue 3. Association for Computing Machinery.*

Building on the insights from the previous chapter and recognizing the absence of a standardized questionnaire for mapping societal attitudes toward augmented humans, we designed and validated a new instrument to address this gap. This questionnaire offers an ecologically valid method for assessing public perceptions, facilitating the integration of social and other heteronomous perspectives, such as perceived autonomy, into research on human augmentation. By providing this tool, we aim to support researchers in systematically incorporating societal considerations into their work. The development and validation process of the questionnaire is detailed below.

7.1 Scale Formation

The SHAPE scale was developed with the aim of facilitating standardized measurement of attitudes towards human augmentation and performance-enhancing technologies in the field of human-computer interaction (HCI). The SHAPE scale will enable HCI designers to create human augmentation and performance-enhancing technologies that are better aligned with the attitudes and expectations of the general public, thus promoting wider social acceptance and adoption of augmentation technologies.

This study has been fast-track approved by the Institutional Review Board of the University of Munich (LMU Munich).

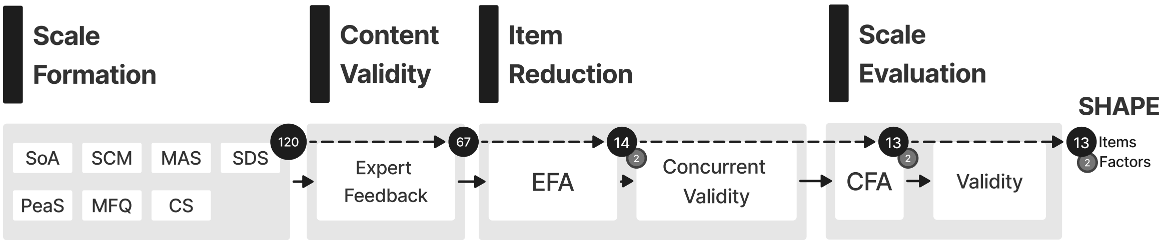


Figure 7.1: Study diagram:In the first stage, Scale Formulation, we searched the literature for intersecting instruments and generated an initial set of items. We then reduced the number of items using expert interviews, and finally, we performed an exploratory factor analysis to reduce dimensionality and discover the underlying structure of the factors. The construct's structure was then assessed using a confirmatory factor analysis. Finally, we ran a series of tests to validate the SHAPE scale psychometrically to establish and validate its final structure.

7.1.1 Item Generation

While several instruments for evaluating perceptions and attitudes towards various technological instances exist [162, 173], a gap remains in the assessment of attitudes towards technologies that blur the boundaries between humans and machines [Het1, 46] as for example Electrical Muscle Stimulation (EMS) to improve reaction times [23], or the use of wearable robotics to control multiple supernumerary limbs at the same time [355]. To address this gap, we constructed the SHAPE scale. As a first step, we conducted an analysis of existing studies and measures in related research fields. This analysis aimed to synthesize the data from instruments with intersecting concepts, such as the sense of agency [135], or attitudes towards assistive technology users [356], and inform the development of the SHAPE scale. The items selected from these instruments were then adapted and grouped to form the initial pool of items for the SHAPE scale. Here, we describe the concepts and instruments selected to design this initial pool of items.

Sense of Agency Scale (SoA)

The relationship between body and action ownership is fundamental to the formation of our self-perception and perception of others [46]. With the proliferation of augmentation technologies, there is an increasing concern that these tools may alter our sense of self and others [46]. Prior research has shown that SoA is especially important in the context of augmented humans, as augmentation technologies may alter an individual's SoA and the amount of effort users invest in a task [Het1, 46].

SoA can be defined as the subjective experience of initiating and controlling one's own actions [357]. It is typically measured by means of self-report through the Sense of Agency Scale [135]. The items extracted from the sense of agency (SoA) scale evaluates an individual's perceived control over their body and actions, providing valuable insights into their subjective experience of agency. Given that the items on the SoA scale are framed in the first person, we modified them to reflect a reference to a third-person perspective, e.g., the item "I am in full control of what I do" changed to "An augmented human is in control of what they do."

Social Stereotype - Stereotype Content Model (SCM)

The SCM is a psychological theory that explains how individuals develop stereotypes about others. It describes stereotypes regarding distinct social groups along two broad dimensions: Warmth and Competence [358, 359]. These factors allow for the prediction of a range of emotions and perceptions, including pride, pity, contempt, and envy towards a distinct social group (in this case towards augmented humans). In HCI, the SCM has been used to describe labeler bias [360], people stereotypes for artificial intelligence systems [361]. In line with [361], we adapted the items of [358] to measure how attitudes vary according to perceptions of competence and warmth for augmented humans. For example, an adapted item would be phrased as follows "In general, augmented humans are perceived as warm."

Multidimensional Attitudes Scale Toward Persons with Disabilities (MAS)

The manner in which an augmented human uses the augmentation technology, the reasons for its use, and whether or not the individual had a prior disability, are among the factors that can evoke emotions in the observer [Het1].

The MAS Scale is a validated measure of attitudes toward people with disabilities. It provides a comprehensive picture of the attitudes of society towards this population. The MAS has been shown to be a valid and reliable measure of attitudes toward people with disabilities [356]. It can be used to identify areas where attitudes may need to be altered and to inform the design of interventions aimed at reducing prejudice and promoting the inclusion of persons with disabilities. The items from the MAS scale provide valuable information about the observer's emotions, thoughts, and behaviors towards augmentation technologies users.

Social Distance Scale (SDS)

The SDS [362] is constructed to measure stigma and is a routine measure in stigma research [363]. As Augmented individuals may face stigma and discrimination, we adapted these items to study how social stigma may affect augmented humans. The original SDS Scale measures how far away from a group (such as people with a specific disease) an individual would like to remain. We adapted the SDS items to fit into the human augmentation context, resulting in items with the following structure "I would have an augmented human as citizens in my country"

Performance Enhancement Attitudes Scale (PeaS)

PeaS [177]focuses on measuring attitudes toward performance enhancement via non-technological means, such as doping. Given the similarities between human augmentation and performance enhancement, this group of items adapts key PEAS components to the context of human augmentation. Items such as "Doping is necessary to be competitive." were adapted to the human augmentation context, resulting in "An augmented human is more competitive."

Moral Foundations Questionnaire (MFQ)

Numerous factors, such as sociocultural context and individual personality traits influence the perceptions of morality. We adapted the MFQ [364] to evaluate how the observer integrates the concept of human augmentation into their personal values, cultural norms, and political ideologies. The MFQ quantifies moral convictions. It assesses an individual's moral sensitivity across five theoretical dimensions, including Harm/Care, Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, and Purity/Sanctity. We extracted and adapted the items relevant to human augmentation. e.g., the item "Whether or not someone violated standards of purity and decency" from the MFQ is reflected in the initial pool of items as " An augmented human would violate standards of purity and decency."

Cross-Sectional Studies in Human Augmentations (CS)

Recent research in human-augmentation and integration technologies has uncovered a variety of factors that influence society's evaluation of adopters of these technologies [Het1, 46, 365]. To better understand how people perceive human augmentation technologies and the reasons behind their assessment, **Villa, Steeven** et al. [Het1] conducted a mixed-method cross-sectional study. They discovered that seven factors influenced people's opinions of human-augmentation users: Privacy, Peril, Ownership, Motivation, Perception of Achievements, and Personal Preference. We added and modified these items to the original set of SHAPE scale items. e.g, the item "I think this person has to disclose the presence of this augmentation in their body to other people." was adapted to be depersonalized as follows "An augmented human has to disclose his augmentation."

Item Reduction

To construct a coherent and consistent initial item set based on the instruments described above, the authors have put forth a set of criteria that would inform the wording and selection of the items. In detail, it was prioritized the use of positive, unambiguous and concise phrasing, the use of depersonalized and hypothetical language whenever possible [309], the use of unemotional language, avoidance of abbreviations and that no prior knowledge is needed for the respondent.

One researcher initially reformulated the initial items according to the established criteria. Subsequently, two researchers separately evaluated the wording of the items independently. A final discussion was then held to address and resolve any disagreements regarding the wording of the items. For all items, a seven-point Likert scale was used to measure agreement (7. Very Much) or disagreement (1. Not at all). In this step, we obtained a total of 120 items

7.1.2 Expert Review

In the subsequent phase, we obtained feedback from six experts who have a record of publication in the domain of human augmentation. The experts provided feedback on each item and suggested eliminating/adding items. Following the expert review, two researchers consolidated and integrated their feedback.

Participants

We invited six experts in human augmentation to participate in the study. Table 7.1 presents the demographic information of the participants. Experts were selected based on publication-record in the field of human augmentation in the Conference on Human Factors in Computing Systems (CHI) and the Augmented Humans (AHs) conference. A total of seven experts were contacted via email, from whom six accepted the call for participation. The interviews and analysis were performed by two researchers, each researcher interviewed three experts. The interviews took place in a period of approximately one month given the availability of the experts. The experts' participation in the study was strictly voluntary and without financial compensation.

Table 7.1: Participants' demographic information: Expert review

Expert	Research	Background
E1	Virtual Reality	Computer Science
E2	Information Transfer	Psychology
E3	Physiological Sensing	Computer Science
E4	Thermal Imaging	Human-computer Interaction
E5	Physiological-based Systems	Computer Science
E6	Neuroscience	Cognitive Science

Procedure

Prior to the interview, the experts were provided with a document containing the initial pool of items to become familiar with the content of the scale. During the interview process, the experts were requested to give feedback on the current set of items, propose new items, and modifications or removal. The interviewers went through each of the 120 items and asked the experts to provide verbal feedback and annotations on the provided document. The annotated documents and interviewer notes were then collected for further analysis.

Analysis

Two researchers participated in the analysis; first, the items suggested for removal by at least one expert were excluded, and the remaining set of items, including those suggested for rephrasing, were discussed. Afterward, the interviewers assessed each item individually and rated the item quality based on the expert feedback on a scale of 1 to 10. Items with high scores (above 6) were retained, items with scores below 3 were excluded, while items with scores between 3 and 6 or were discussed and kept or removed after reaching a consensus.

The expert review started with an initial pool of 120 items sourced from the previously described instruments. The integration of the expert feedback resulted in a reduction of the item pool to 67 partly reformulated items.

7.1.3 Survey #1

In the next stage of our scale development process, we designed a Qualtrics-based online survey to collect data from participants and conducted an exploratory factor analysis and item reduction. Boateng et al. [309], referring to Comrey [366], recommends a sample size of a minimum of 200 participants for studies of this kind and we exceeded this minimal sample size recommendation with a sample size of $n = 302$ participants.

Participants

The sample was composed of 149 female and 153 male participants with a mean age of 44.4 years (SD=13.0). No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. Participants were recruited through the UK-based platform Prolific, with the sample being drawn from the United Kingdom and the United States. All participants reported English as being their mother tongue. Participation was voluntary and compensated by 9 GBP per hour. The participants were informed that the collected data would be anonymized prior to processing. The survey was distributed in an online format and took participants an average of eight minutes to complete ($M = 8.02$, $SD = 4.24$).

Survey Structure

The survey started with an informed consent form, and after participants gave their consent, they read a scenario depicting the journey of an augmented human interacting with a group of people. The scenario was developed based on Findler, Vilchinsky, and Werner [356] and [Het1] work. This scenario was designed to elicit a range of attitudes towards augmentations by incorporating all possible permutations of cognitive, sensory, and motor augmentations. The following is the scenario:

> “Michael went out for lunch with friends to a coffee shop. A man with some technological modifications, with whom Michael is not acquainted, enters the coffee shop and joins the group. Michael is introduced to this person. During the chat, the man tells them that he replaced some of his healthy body parts and replaced them with improved artificial ones: an artificial eye to augment his vision beyond the normal range. Artificial legs to run faster and jump higher than ordinary humans. Additionally, he got a brain implant to

> think faster and have more memory than ordinary humans. Shortly after that, everyone else leaves, with only Michael and the man with the technological modifications remaining alone together at the table. Michael has 15 minutes to wait for his ride home. ”

After this scenario, the participants were presented with a quasi-randomized set of 67 items. Once the participants had responded to all of the questions, their demographic information was collected and the survey concluded.

7.1.4 Exploratory Factor Analysis

For the item analysis, we inverted the negatively worded items, then we examined the densities of all items and eliminated those with high skew and kurtosis. Then, we conducted a Kaiser-Meyer-Olkin (KMO) factor adequacy test, which evaluates the data's suitability for factor analysis. In a KMO test, values close to 1.0 are desired, and our dataset produced KMO Measure of Sampling Adequacy (MSA) = 0.95. Subsequently, we conducted a Bartlett's Test of Sphericity to evaluate the null hypothesis that the inter-correlations among the variables in the dataset are equal to zero, thereby eliminating the possibility of an identity matrix and ensuring that the variables are suitable for factor analysis ($\tilde{\chi}^2(741) = 9953.585$).

We then performed an exploratory factor analysis. The exploratory factor analysis is a statistical procedure that allows determining the number of underlying factors that explain the pattern of correlation of items [322].

Then, we employed parallel analysis [323] and scree plot analysis [324] to determine the optimal number of underlying factors in the data. The inspection of the scree plot indicated that a two-factor solution was the optimal solution which amounts to extracting factors with an eigenvalue > 1.83.

We then used varimax rotation similar to Woźniak et al. [162]. A varimax rotation produces independent factors; it is an orthogonal rotation method used in factor analysis to maximize the variance of the variable factor loadings while minimizing the number of variables with high factor loadings [367].

From this model, we eliminated all items with loadings below 0.40 and those that were loaded on multiple factors. We merged items with high similarity as a final step. The scale encompassed fourteen items distributed in two factors; seven items per factor. The model had a good fit, KMO $MSA = 0.85$, Tucker Lewis Index of factoring reliability $TLI = 0.842$, and, $RMSEA = 0.104$. Table 7.2 presents the results of exploratory factor analysis. The first factor is related to Social Threat (ST) [109], indicating that augmented humans will pose a threat to oneself and society. The second factor, Agency (AG), is characterized by a focus on control and includes items that assess the perceived agency of the augmented human over their augmentation. Internal consistency as indicated by Cronbach's alpha was $\alpha = 0.852$ for

Table 7.2: The revised version of the SHAPE scale consisted of fourteen items grouped in two factors: Social Threat (ST) and Agency (AG), with item loadings and their respective sources reported.

Item	ID	Source	Factor	
			ST	AG
An augmented human is a threat to society.	I1	CS	0.85	
An augmented human would be dangerous.	I2	CS	0.84	
An augmented human is intimidating.	I3	MAS	0.64	
An augmented human would conform to the traditions of society.	I4	MFQ	0.58	
An augmented human has to disclose their augmentation.	I5	CS	0.58	
An augmented human would do something cruel	I6	MFQ	0.55	
An augmented human is more competitive than a non-augmented human.	I7	PeaS	0.53	
The actions of the augmented human do not match their intentions.	I8	SoA		0.70
An augmented human is not the author of their own actions.	I9	SoA		0.68
An augmented human is just an instrument of something or somebody else.	I10	SoA		0.66
An augmented human does things without any intention.	I11	SoA		0.66
An augmented human suffering through their augmentation should get help.	I12	MFQ		0.50
If an augmented human were to suffer through their augmentation, I would have compassion.	I13	MFQ		0.48
An augmented human is in full control of what they do.	I14	SoA		0.47

ST and $\alpha = 0.834$ for AG and thus can be regarded as good internal consistency of the scales [326].

7.1.5 Content Validity

Warmth and competence have been used to structure stereotypical attitudes towards human augmentation [368]. To establish that our novel measure relates to an established measure, we have correlated the ST-scale and the AG-scale to each warmth and competence of the SCM. We observed that the perceived warmth correlates with both SHAPE factors, meaning that a decrease in perceived threat and an increase in control of augmentation technologies users increase the perceived warmth. Similarly, we found that competence correlates with both ST-scale and AG-scale control factor, see Table 7.3. This indicates that an increase in the perceived control over the augmentation and a decrease in the perceived threat increases the perceived competence. These results are consistent with the findings of Meyer and Asbrock [368].

Table 7.3: Correlations between the SHAPE scale factors, Social Threat (ST) and Agency (AG), and the Warmth and Competence scale. degrees of freedom for all the tests are $df = 300$

Factor	ST			AG		
	r	t	p	r	t	p
Warmth	-0.581	-12.392	$< .005$	-0.491	-9.766	$< .005$
Competence	-0.205	-3.634	$< .005$	-0.392	-7.380	$< .005$

7.2 Scale Validation

After building the factor structure of the scale, we continued with the evaluation of the SHAPE scale. We performed a confirmatory factor analysis to test the fit of the structure to novel data. Subsequently, various correlational tests were conducted to assess the scale's content validity and reliability. In this section, we report the first version of the SHAPE scale and evaluate its consistency. We then refine the scale and construct its final version.

7.2.1 Survey #2

We designed a Qualtrics-based online survey to collect data from participants and conducted a confirmatory factor analysis (CFA) during this phase of the research. It is important to note that the structure of the questionnaire at this stage is identical to that described in subsection 7.1.3, with the exception that the set of items has been replaced with those obtained from the exploratory factor analysis described in subsection 7.1.4.

Participants

For this stage, we recruited a sample of $n = 297$ participants, in accordance with the recommendations by Comrey [366] that posit confirmatory factor analysis requires at least 200 participants. The sample consisted of 150 females and 147 males with a mean age of 44.4 (SD = 13.9) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The sample was composed of individuals from the United Kingdom and the United States who were recruited through the British platform Prolific. All participants were native English speakers. The participants were compensated with 9 GBP per hour. All participants were informed of the voluntary nature of their participation and provided with the option to withdraw at any time if they felt uneasy. Participants were also informed that the collected data would be anonymized prior to processing. The survey was distributed online and took respondents an average time of three minutes to complete ($M = 3.45$, $SD = 1.86$).

7.2.2 Confirmatory Factor Analysis (CFA)

In order to assess the validity of the SHAPE scale's structure, we conducted a Confirmatory Factor Analysis (CFA). This statistical procedure allowed us to confirm the dimensionality of our proposed factor model. The solution had two factors, see again Table 7.2. The results of the model fit assessment indicated a sub-optimal fit, as evidenced by the Root Mean Square Error of Approximation (RMSEA) value of greater than 0.1, a Comparative Fit Index (CFI) of 0.93, and a Standardized Root Mean Square Residual (SRMR) of 0.08. Detailed examination of the data revealed high correlations between two items in the Agency factor (items I12 and I13).

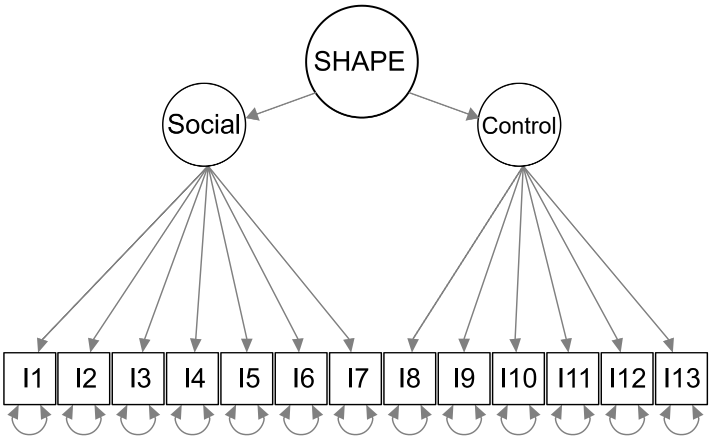


Figure 7.2: The findings of the confirmatory factor analysis indicated a two-factor model for the SHAPE scale, comprising two inter-correlated subscales.

Also, item I13 was identified as dissimilar due to its wording and was removed to improve the coherence of the SHAPE construct.

We conducted another CFA using the reduced set of items. The confirmatory factor analysis revealed an RMSEA of 0.08, which falls within the acceptable range. Additionally, the CFI was calculated to be 0.97 and the SRMR was determined to be 0.063. Both values are within the desirable bounds, with CFI values above 0.95 and SRMR values below 0.08 being considered indicative of a good fit of the data to the model. Again Cronbach's Alpha for ST was $\alpha = 0.808$ and for AG was $\alpha = 0.809$ can be deemed a good level of internal consistency.

The final compositions of the factors is reported in Table 7.4 and visualized in Figure 7.2; A two-factor model consisting of thirteen item; 6 for ST, 7 for AG.

7.3 Test-retest Reliability and Construct Validity

In this step, we evaluated the construct validity of the SHAPE scale through three methods: (1) Reliability: conducting a test-retest reliability study. (2) Content validity: analyzing the correlation between the SHAPE scale and the willingness to acquire an augmentation, and (3) Convergent validity and discriminant validity: examining the correlation between the SHAPE factors and subscales of the Technology Readiness Index (TRI) that bears subscales that conceptually relate to our measure and subscales that do not [369].

7.3.1 Data Collection

To gather data and evaluate the three aforementioned points, we developed two online surveys using Qualtrics software. The surveys included the final thirteen items of the SHAPE scale and were completed by a total of n = 103 participants in the first round and n = 78 participants in

Table 7.4: The final version of the SHAPE scale consisting of thirteen items. Internal consistency Cronbach's alpha values are displayed on top of their respective item group. Each item is answered in a 7 point Likert scale ranging from (1) Not at all to (7) Very Much. (*) denotes that the item is inverted.

Item	ID	Source
Social Threat $\alpha = 0.808$		
An augmented human is a threat to society.	**S1**	CS
An augmented human would be dangerous.	**S2**	CS
An augmented human is intimidating.	**S3**	MAS
(*) An augmented human would conform to the traditions of society.	**S4**	MFQ
An augmented human has to disclose their augmentation.	**S5**	CS
An augmented human would do something cruel	**S6**	MFQ
(*) An augmented human is more competitive than a non-augmented human.	**S7**	PeaS
Agency $\alpha = 0.809$		
The actions of the augmented human do not match their intentions.	**S8**	SoA
An augmented human is not the author of their own actions.	**S9**	SoA
An augmented human is just an instrument of something or somebody else.	**S10**	SoA
An augmented human does things without any intention.	**S11**	SoA
An augmented human suffering through their augmentation should get help.	**S12**	MFQ
(*)An augmented human is in full control of what they do.	**S13**	SoA

the second round. The surveys were distributed with a minimum interval of 15 days between assessments ($M = 16.52$, $SD = 0.63$, $min = 15.66$, $max = 18.31$).

Survey #3: Test-Retest first sample, Technology Readiness Index (TRI), and Willingness to acquire an augmentation

The survey started with an informed consent process; following this, participants viewed the same scenario from the first survey (see subsection 7.1.3 for details). Participants were then presented with the thirteen-item SHAPE scale, and upon completion, participants were asked a binary question regarding their willingness to acquire an augmentation, "I would like to get an augmentation for myself," with response options of "Yes" or "No." Finally, we administered the Technology Readiness Index (TRI) before concluding the survey by collecting demographic data.

Participants

For this stage, we recruited a sample of n = 103 participants using Prolific, The sample consisted of 51 females and 52 males with a mean age of 45.5 (SD = 13.1) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The recruiting, compensation and consent scheme were similar to the previous two studies. The survey was distributed online and took respondents almost six minutes to complete ($M = 5.88$, $SD = 3.15$).

Survey #4: Test-retest second sample

Most of the questions from Survey #3 were re-invited to Survey #4, with the Technology Readiness Index (TRI) being the only exception.

Participants

About 80% responded again, n = 78, using Prolific. The sample consisted of 44 females and 34 males with a mean age of 47.3 (SD = 13.9) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The compensation and consent scheme was the same as in the previous study. The survey was distributed online and took respondents an average of four minutes to complete ($M = 4.19$, $SD = 6.80$).

7.3.2 Test-retest Reliability

Temporal stability refers to the ability of a scale to produce consistent results when administered to the same participants at different time points [309]. We conducted a test-retest reliability evaluation to assess the temporal stability of the SHAPE scale construct. This psychometric evaluation is commonly used in the scale development process (e.g. Woźniak et al., Bentvelzen et al. [162, 308]) to estimate reliability based on temporal stability.

Similar to Woźniak et al. [162], we calculated calculated a two-way Single-measurement intraclass correlation coefficient (ICC) for consistency and agreement. The ICC quantifies the degree of agreement between two or more continuous measures, values close to 1 indicate a perfect agreement whilst values close to 0 indicates no agreement at all. The ICC, for each subscale[1], indicated good reliability for ST and AG in terms of consistency (ST $\kappa = 0.735$, AG $\kappa = 0.715$) and agreement (ST $\kappa = 0.736$, AG $\kappa = 0.709$), see also Table 7.5. Additionally we computed Spearman correlations for each subscale, indicating a high correlation between samples; namely for AG we found that $r_s = 0.68$, $p < .005$, and for ST $r_s = 0.707$, $p < .005$.

Table 7.5: The two-way single-measurement intraclass correlation coefficients (ICC) calculated for both the Social Threat (ST) and Agency (AG) factors of the SHAPE scale.

Factor	ST			AG		
	κ	F	p	κ	F	p
Consistency	0.735	F(76,76) = 6.54	$< .005$	0.715	F(76,76) = 6.02	$< .005$
Agreement	0.736	F(76,76.8) = 6.54	$< .005$	0.709	F(76,74.2) = 6.02	$< .005$

To further determine the absolute reliability of the SHAPE scale, we analyzed the data using the Bland and Altman method [331]. Each participant's mean difference between the initial

[1] We calculated the Intraclass Correlation Coefficient (ICC) for each subscale by averaging the responses of the Likert items on the subscale. Therefore it is important to note that this computation was performed on interval data, as opposed to ordinal data.

test and the retest was plotted as a function of the means of both test sessions using Bland-Altman plots. The dashed horizontal lines in the plots represent the limits of agreement, which correspond to the 95% confidence interval surrounding the mean difference between the test sessions. These limits indicate the range within which 95% of the values are likely to fall [331, 332].

Figure 7.3: Bland and Altman plots: difference in SHAPE scores obtained from two surveys (Described at the beginning of this section) as a function of the average score of both test sessions for individual participants, the data is segregated based on ST and AG categories. The mean bias is indicated by the black line, while the 95% limits of agreement are represented by the gray lines.

In the plot (see Figure 7.3), the mean difference close to zero, (dotted line) indicates that the SHAPE scale has absolute temporal stability on average and the distribution around zero is indicative of reliability not being related to the mean score, thus, demonstrating that it can be reliably administered at different time points and is suitable for use in between-groups or repeated-measures designs.

7.3.3 Concurrent Validity

In this step, we wanted to investigate the extent to which the factors of the SHAPE scale could predict an individual's inclination to obtain augmentation technologies to show concurrent validity. We measured this inclination in Survey one with the response options of "yes" or "no" to the question "I would like to get an augmentation for myself.". We calculated Spearman correlation, for the ST and AG and the above-mentioned question. We found a negative association for the ST, r_s = -.40, p < .001, and AG, r_s = -.31, p = .001, concerning their indication of willingness to acquire an augmentation technology. The less threat and the more control they attribute to augmented humans in general, the more likely participants are to indicate they would want to use an augmentation technology themselves.

7.3.4 Convergent & Divergent Validity

Utilizing a methodology similar to that of Schepman and Rodway [369], we assessed the convergent validity of the SHAPE scale by applying the Technology Readiness Index (TRI). The TRI scale comprises 18 items and is frequently used due to its sound psychometric properties [370]. The TRI scale has four subscales: Innovativeness, Optimism, Discomfort, and Insecurity. The scale has demonstrated the ability to predict user interactions with technology products [369]. The Innovativeness sub-scale is correlated with the tendency to be a thought leader, Optimism with a positive view about technology, discomfort, with the feeling of being overwhelmed by technology and Insecurity, with distrust in technology. We expect Innovativeness to be conceptually independent of ST and AG and discomfort and insecurity to overlap with ST and AG .

To evaluate the internal consistency of the TRI, we determined the Cronbach alpha for each sub-scale. The resulting alpha coefficients were $\alpha = 0.813$ for Innovativeness, $\alpha = 0.698$ for Optimism, $\alpha = 0.725$ for Discomfort, and, $\alpha = 0.792$ for Insecurity. The obtained metrics reflect an acceptable to good performance for each sub-scale. We then obtained the sub-scale values by computing the average of the corresponding items.

The correlations of the SHAPE scale and the TRI factors are presented in Table 7.6. The correlation analysis indicated that the Social Threat and Agency factors of the SHAPE scale were strongly correlated with the Discomfort and Insecurity scales of the TRI. The less Discomfort and Insecurity experienced in response to technological advancement, the less they perceived augmented humans as threatening and the more control they attributed to them. Thus, we can show convergent validity to negative aspects of technology readiness concepts.

Table 7.6: Correlation with Technology Readiness Index. degrees of freedom for all the test are $df = 101$

Factor	ST			AG		
	r	t	p	r	t	p
Innovativeness	-0.119	-1.206	0.230	-0.132	-1.340	0.183
Optimism	-0.153	-1.559	0.122	-0.272	- 2.848	< 0.01
Discomfort	0.277	2.902	< 0.005	0.410	4.523	< 0.005
Insecurity	0.213	2.197	< 0.05	0.468	5.326	< 0.005

7.4 Validation of SHAPE Scale in the Context of Disabilities

As a final step, we further explored the fit of the two-factor thirteen item structure of the SHAPE scale to the assessment of augmentation technologies when the user of such technology is an individual with a previous disability condition. We conducted a new Confirmatory

Factor Analysis with a modified vignette to reflect a scenario where the technology user is enhancing their skills to compensate for a disability.

7.4.1 Survey #5

We developed a new online survey using Qualtrics software. The survey included the final structure of the SHAPE scale and was completed by a total of $n = 216$ participants, in accordance to Comrey [366]. The sample consisted of 123 females 91 males and 2 individuals that preferred not to disclose their gender, the mean age of participants was 43.9 ($SD = 11.66$) years. No participants self-identified as non-binary or other. The sample was comprised of individuals from the United Kingdom and the United States who were recruited through the platform Prolific. All participants were native English speakers. Participants were compensated with 9GBP per hour. All participants were informed of the voluntary nature of their participation and provided with the option to withdraw at any time without the need for further explanation. Participants were informed about the data collection and the anonymization policy prior to processing. The survey was distributed online and took respondents an average time of three minutes to complete ($M = 3.55$, $SD = 1.82$).

7.4.2 Confirmatory Factor Analysis for the Disabilities Context

We performed a new Confirmatory Factor Analysis utilizing a modified vignette that accounts for a scenario where users are augmenting their skills to offset disability-related impairments instead of augmenting to increase their skills. The original two-factor thirteen items model revealed a sub-optimal fit to this new scenario; the Root Mean Square Error of Approximation (RMSEA) value was greater than 0.1 and the Comparative Fit Index (CFI) had a value of 0.96 with a Standardized Root Mean Square Residual (SRMR) of 0.06. Given these values, we examined the correlations between items using modification indices [371] , which indicate a potential change in χ^2 when adding or removing items, We subsequently deleted S2, S11, and, S12 which were highly skewed and thus had a reduced variance for a disabilities context (more positively valenced responding patterns).

We then performed a CFA with these three items removed. The new analysis revealed an RMSEA of 0.07 which can be considered as a reasonably good fit. For the CFI, we found a value of 0.98, with a SRMR value of 0.043. Both values being considered as a good fit of the data to the model. The Cronbach's Alpha for ST $\alpha = 0.773$ and for AG $\alpha = 0.802$ which can be interpreted as a good level of internal consistency.

In addition, we compared the reduced scores of Survey 5 (validation for disability scenarios) and Survey 3 (test-retest on the initial sample) for both subscales using an unpaired *t*-test. The results indicated that the score for the non-disability scenario was significantly higher overall (M_{nd} = 3.87, SD_{nd} = 0.76) than the disability scenario (M_d = 3.40, SD_d = 0.74, $t(192.21) = 5.10$, $p < 0.005$). In addition, the subscales exhibited a similar pattern, for example for Social Threat we found (M_{ST-nd} = 4.19, SD_{ST-nd} = 0.91) and (M_{ST-d} = 3.63, SD_{ST-d} = 0.90, $t(196.91) = 5.25$, $p < 0.005$), while for Agency we found (M_{AG-nd} = 3.54, SD_{AG-nd} = 0.83) and (M_{AG-d} = 3.18, SD_{AG-d} = 0.82, $t(194.31) = 3.57$, $p < 0.005$).

Based on the findings of this confirmatory factor analysis, the scoring system for the context of disabilities is described more in detail in subsection 7.5.1.

7.5 Discussion

In this section, we provide an overview of our approach, the necessary details for administering the SHAPE scale as well as information on how to use it. In addition, we discuss the limitations of our approach and opportunities for further developments.

In this paper, we introduce the development and validation process of a brief 13-item measure, the SHAPE scale, which was designed to measure attitudes towards humans using digital technologies that enhance human abilities.

We identified a two factorial structure that encompassed attitudes that we summarized under the Social Threat factor, which measures threat to oneself and others, as well as a factor that we summarized under the Agency factor, which describes agency and support for augmented humans. The SHAPE sub-scales were validated and refined in confirmatory factor analysis, showing a good fit based on several fit indices, excellent internal consistency and good test-retest reliability. Also, medium test-retest reliability indicates that attitudes toward augmented humans might be susceptible to changes over time and can thus be used to investigate how attitudes toward augmented humans evolve in the future.

We have evaluated the validity of SHAPE across studies. In Survey #1, we could show that threat and competence relate to the Stereotype-content model; people that attribute low threat to augmented humans perceived them as warmer, while competence of augmented humans was increased for low social threat and more control. This aligns with the findings of Meyer and Asbrock [368], who discovered that individuals with bionic prostheses were perceived as competent without a reduction in perceived warmth.

On the other hand, in Survey #3, we demonstrated construct validity. There is convergent validity in terms of correlation with the technology readiness index that addresses discomfort and insecurity about technological developments but discriminant validity in terms of innovativeness. Therefore, the scale covers both stereotypes' attributes on the perception of augmented humans and technological attributes. In Survey #3, we could also show concurrent validity in that attitudes toward augmented humans can predict whether participants are willing to use augmentation technologies themselves. Therefore, positive attitudes regarding threat and control in augmented humans are associated with acceptance of the technology. This mirrors the recent call in HCI [149] to integrate negative aspects of social acceptability into technological acceptance models.

So far, research in the area of social attitudes toward augmented humans has been limited due to the lack of assessment tools. Work that considered attitudes towards augmented humans was mainly conducted using qualitative methods [Het1, 372]. Quantitative studies in the domain have adapted conventional scales ,e.g. from the SCM [368, 373, 374], at the expense of interoperability and specificity to the domain of human augmentation. SHAPE now gives researchers in the domain of human augmentation a tool to quantify attitudes in terms of Social Threat and Agency, which adds a quantitative tool to the repertoire of researchers in the domain of human augmentation. We envision that the scale can meaningfully complement qualitative approaches and thereby enable holistic and impactful insights into the field of human augmentations. In this respect, the scale can be a particularly valuable addition when it comes to comparative long-term studies and studies that are concerned with the attitudes of different samples towards human augmentation technologies (e.g., users from different countries).

According to [Het1], new augmentation devices should be designed with a focus not only on the artifact itself but on the human that would be integrating it into their life/body and their social environment. Our scale development process showed that the assessment of the social human factor is comprised of two aspects: Social Threat and Agency, which should be considered when evaluating augmentation technologies and other types of performance-enhancing technologies.

In the final set of items of SHAPE , there is no explicit reference to privacy threats, which is interesting given that only one item related to privacy was removed, while the remaining items underwent filtering in the EFA. The absence of explicit representation of privacy concerns among the filtered items may suggest that we considered them to be less relevant compared to other factors, such as the agency of the augmented human or the perceived threat it poses to the observer. Furthermore, we acknowledge that the subscale "Social Threat" may not specifically target any particular type of threat, including privacy threats. Therefore, it is possible that this subscale captures certain aspects of privacy concerns, even in the absence of explicit references to privacy threats.

Our study provides valuable information on the social perception of augmented humans. In the initial item pool, we had a sizable number of items that corresponded to a benevolent or positive view of augmented humans,e.g., "An augmented human is interesting." or "An augmented human is friendly." from the MAS scale. However, none of these items surfaced in the exploratory factor analysis to correspond to a factor. We thus suggest that in our sample, attitudes mainly revolved around a negative view of augmented humans. This aligns with recent scale developments such as the Creepiness of Technology Scale (PCTS) [162] where the authors reported three subscales, all of them negatively valenced. Nevertheless, it will be important for future research to investigate measurement invariance of the SHAPE scale as attitudes differ across cultures [Het1]. This resonates with the fact that beliefs and attitudes toward innovative technologies are ever-evolving. To illustrate this point, the TRI was updated after only a little more than a decade [375, 376] to cover novel aspects of technology readiness. Likely SHAPE might need to be revised when augmentation technologies are more broadly used. This limitation also points to the research opportunity to investigate with the SHAPE scale how attitudes evolve and change over time. The SHAPE scale was built to be unspecific concerning the disability status and the type of augmentation, covering sensory, motor and cognitive augmentations alike; future studies may piece apart how attitudes differ as a function of augmentation characteristics and person characteristics. In order to enable this, we validated the disabilities scenario and discovered that SHAPE can also be utilized effectively in the context of disabilities by ignoring the non-descriptive items. The three non-descriptive items for the disability case pertain to situations that may have been affected by the observer's forgiveness of individuals with prior disabilities. This aligns with previous work that has found that observers find more acceptable the use of some technologies when the user has a disability condition [377, 378, 379]. In subsection 7.5.1 we provide a tailored scoring system for this specific case.

The final version of the SHAPE scale is available at `posthci.com`. This website has long-term support planned and is available to distribute the scale easily. The website is planned to serve as a reference point to evaluate the evolution of the attitudes toward human augmentation and performance-enhancement technologies. The anonymized collected data in the website along with translated versions of the SHAPE scale will be made available for researchers to further advance the field.

7.5.1 Scoring

The SHAPE scale is scored on a seven-point Likert scale from Not at All (1) to Very Much (7). Items S4, S7 and S13 are reverse-scored. Higher scores indicate higher aversion towards augmentation technology's users:

Full Scoring System

In the full scoring system of the SHAPE scale, it is advisable to calculate the arithmetic mean of all the items to obtain the overall score, or to compute the mean of the items corresponding to each subscale if the reader seeks insights into specific dimensions. This approach is feasible because both subscales possess equal valence; higher scores indicate a greater degree of aversion towards Augmented Humans or Performance Enhancing technology users.

$$
\begin{aligned}
SHAPE &= \frac{\mu_{ST} + \mu_{AG}}{2} \\
with\ \mu_{ST} &= mean(S1 + S2 + S3 + S4_R + S5 + S6 + S7_R), \\
and\ \mu_{AG} &= mean(S8 + S9 + S10 + S11 + S12 + S13_R)
\end{aligned}
$$

Disability Scenarios Scoring System

In the context of disability scenarios, we suggest using a scoring system similar to the full scoring system. However, instead of utilizing the entire set of items, we suggest excluding the items that exhibit a pronounced skew in opinions toward individuals with disabilities. This adjustment is intended to improve the reliability and validity of the scoring procedure.

$$SHAPE = \frac{\mu_{ST} + \mu_{AG}}{2}$$
$$with\ \mu_{ST} = mean(S1 + S3 + S4_R + S5 + S6 + S7_R),$$
$$and\ \mu_{AG} = mean(S8 + S9 + S10 + S13_R)$$

Reflections on Part III

This part explored the perception, social attitudes, and ethical considerations surrounding augmented humans. Through a series of empirical studies, we identified key dimensions that influence how augmented humans are evaluated by society, developed a validated measurement tool for assessing these attitudes, and provided design recommendations to guide the responsible development of augmentation technologies.

We first investigated the perception of augmented humans, revealing complex tensions between general optimism about augmentation and personal reluctance to adopt it. Our findings demonstrated that factors such as peril, privacy, access, motivation, ownership, and achievement play critical roles in shaping public opinion. Notably, cognitive augmentations were viewed with the most skepticism, raising concerns about fairness, safety, and societal impact. Additionally, cross-cultural differences emerged, particularly between German and Japanese participants, highlighting the influence of cultural values on augmentation acceptance.

Building on these insights, we developed the SHAPE scale, a validated instrument for quantifying attitudes toward augmented humans across two key factors: perceived social threat and agency/control. Our validation studies confirmed its reliability and predictive power, showing that positive attitudes toward augmented humans correlate with greater acceptance of augmentation technologies. This scale provides a much-needed quantitative tool for researchers investigating societal attitudes toward human enhancement and its long-term evolution.

Finally, we translated our findings into design recommendations aimed at fostering more inclusive and acceptable augmentation technologies. Our recommendations emphasize the importance of clear communication about augmentation purposes, hybrid augmentation designs that serve both assistive and enhancement functions, and balancing effort and agency to maintain a sense of personal achievement. Additionally, our results suggest that privacy concerns extend beyond personal data protection to fears about how augmented humans might use enhanced sensory capabilities in social interactions.

IV

Experiencing and Simulating Human Augmentation

8 HUMAN AUGMENTATION VERSUS LEARNING

This chapter is based on the following publications:

Understanding the Influence of Electrical Muscle Stimulation on Motor Learning: Enhancing Motor Learning or Disrupting Natural Progression? - Steeven Villa, Finn Jacob Eliyah Krammer, Yannick Weiss, Robin Welsch, Thomas Kosch - *2025 CHI Conference on Human Factors in Computing Systems (CHI '25). Association for Computing Machinery.*

Assisting Motor Skill Transfer for Dance Students Using Wearable Feedback - Steeven Villa, Jasmin Niess, Bettina Eska, Albrecht Schmidt, Tonja-Katrin Machulla *In: Proceedings of the 2021 ACM International Symposium on Wearable Computers (ISWC '21). Association for Computing Machinery.*

After discussing autonomy and heteronomy in the previous sections, this part focuses on the experience of being augmented, including the potential to explore augmentations in virtual reality. The first chapter investigates motor augmentation in the context of motor learning, followed by a use case involving wearable motor augmentation for dance instructors.
This chapter specifically examines Electrical Muscle Stimulation (EMS), a motor augmentation technology commonly used to enhance physical performance. While EMS is widely adopted, its effect on motor learning remains unclear. According to motor learning theory, consistent practice and movement awareness are essential for skill acquisition, and disruptions to sensorimotor representation during training can hinder long-term retention. Since EMS externally actuates the body, it alters sensorimotor mapping, which may lead to difficulties retaining skills after the assistance is removed. Conversely, some research in Human-Computer Interaction (HCI) suggests EMS can support motor learning. To resolve this tension, this chapter presents a multi-session study evaluating the impact of EMS on motor learning outcomes.learning outcomes.

8.1 Neural Basis of Motor Learning

Motor skills are developed through different cognitive stages [380], with information moving from short-term to long-term memory, which includes explicit and implicit types [381]. Explicit memory is conscious while implicit memory is not, and it is challenging for individuals to articulate it in detail. In motor skill learning, explicit memory forms in the early stages and is later consolidated into implicit memory, requiring less conscious attention. Motor skills can be retained for years. Various models explain the neurobiological processes

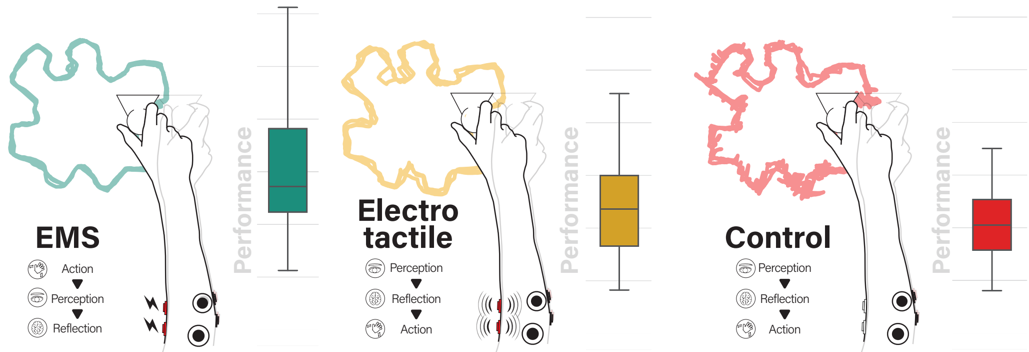


Figure 8.1: In this paper, we compared Electrical Muscle Stimulation (EMS) and electrotactile feedback against a no feedback control condition for evaluating motor learning consolidation. Our results show that EMS enhances motor skill acquisition despite a lower initial learning rate.

behind implicit memory storage [382, 383, 384, 385, 386] and brain region contributions during learning [387], such as the medial temporal lobe in fast learning and cortical motor regions in slow learning. Among the various motor learning models studied, the framework proposed by Doyon et al. [383] is particularly relevant to our research. This model outlines five phases of motor skill acquisition and retention: (1) **Fast (early) learning**, where rapid improvements occur; (2) **Slow (later) learning**, characterized by gradual performance gains; (3) **Consolidation**, during which learned skills are stabilized; (4) The **automatic phase,** where skilled behavior becomes more effortless and consistent; and (5) **Retention**, which describes the maintenance of skills after long periods without practice.

The **fast (early) learning phase** is characterized by significant improvements in motor behavior. This stage requires high levels of attention and generates substantial cognitive workload, especially when a task is encountered for the first time [388]. Fast learning is typically observable from the first session, and research shows that proper feedback provides significant benefits during this phase [389]. Additionally, error correction plays a critical role in early learning, being more important at this stage than in later phases [385].

The slow learning phase, spanning multiple sessions, is characterized by a deceleration in progress as motor skill performance stabilizes and becomes more consistent. Motor learning is inherently time-dependent, with **consolidation** serving as an intermediate process between practice sessions. During this phase, explicit knowledge of the motor skill transitions into implicit memory. Notably, evidence suggests that sleep plays a crucial role in motor memory consolidation [383, 390, 391]. Beyond sleep, factors such as interest, motivation, attentiveness, vigilance, and levels of distraction also significantly influence how well a memory is retained [381, 392]. **Motor consolidation** is key to embedding the skill into the body's memory, eventually leading to the **automatic execution phase**. This last stage occurs when the task can be performed without conscious effort, indicating that the motor skill has become automatic. **The retention phase** is achieved when this skill can be recalled after a significant period without practice, remaining intact in long-term memory.

Most HCI studies on motor learning primarily focus on the fast-learning phase. A variety of interfaces have been proposed to support motor learning during this phase, often through augmented feedback mechanisms, such as vibrotactile feedback. These mechanisms aim to enhance awareness of movements and errors, supporting reflection and adjustment. Such approaches fit coherently within motor learning frameworks, as they improve awareness by providing additional information, enabling users to reflect on their actions, adjust motor behavior, and thus learn more effectively.

8.2 Experimental Design and Hypotheses

EMS introduces a new paradigm by reversing the typical motor learning sequence. Traditional approaches focus on deliberate practice and user adjustments [Exp2], following a perception-reasoning-action cycle. However, EMS-augmented actions occur before reflection, in the order of action, perception, and reasoning [393]. This can hinder reflective practice and sensorimotor learning Proteau, Marteniuk, and Lévesque [394]. For example, Tatsuno, Hayakawa, and Ishikawa [395] found that participants trained with EMS in a wrist rotation task compensated for EMS-induced movements after removing stimulation, although this compensatory effect diminished over time. Similarly, Nishikawa et al. [396] recently found that EMS use during hand gesture learning led to higher errors.

We operationalized motor learning through the Mirror Drawing (MD) task across two distinct sessions, a common task that has been employed in previous research [397, 398, 399]. The MD task measures motor learning through both within-session performance (Post-Training 1 and 2 Assessments - **fast learning phase**) and across-session performance (Consolidation Assessment - **consolidation phase**). We assessed learning using two metrics: (1) the distance traced within a fixed time and (2) the total time to trace a complete shape. Participants show motor learning by tracing longer distances in the given time and completing shapes faster. To examine learning transfer, we introduced an unfamiliar shape at the end of the second session, evaluating the participants' ability to apply their acquired motor skills to a new context. To assess learning rates and gains, we use an exponential decay model. Based on this experimental design and in light of previous work, we derived the following hypotheses:

H1 There will be a performance difference between the ELECTROTACTILE, EMS, and CONTROL conditions during the motor learning task, with ELECTROTACTILE showing better results than the CONTROL and EMS conditions.

H2 ELECTROTACTILE stimulation will result in significantly better motor skill consolidation performance compared to both EMS and the CONTROL condition.

H3 Learning rate will be higher with ELECTROTACTILE stimulation compared to EMS, as indicated by the α parameter of the exponential decay model.

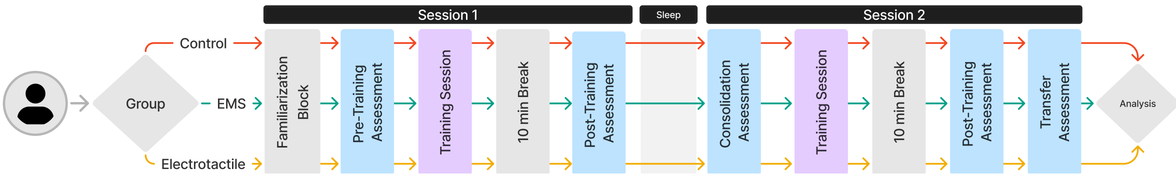


Figure 8.2: Procedure of the experiment. Our participants were assigned to their condition and got familiarized with the setup and task. Then, the participants started with the first assessment, a break, and a second assessment. After a long-term break, participants participated in a third and fourth session. Finally, we assessed transfer learning using a transfer assessment.

H4 ELECTROTACTILE stimulation will lead to greater overall learning, as measured by the A parameter of the exponential decay model, compared to both EMS and the CONTROL condition.

H5 The EMS condition will show significantly lower performance in the learning transfer task compared to the CONTROL group.

Having these hypotheses, we tested the following conditions, following the procedure outlined in Figure 8.2:

- **EMS**: Participants in this group received Kinaesthetic EMS feedback. This type of feedback provides electrical impulses that stimulate muscle contractions. Thereby applying a corrective movement in the correct direction through participant's muscle exertion.
- **Electrotactile**: Participants in the ELECTROTACTILE group received electrotactile feedback. This approach involves the delivery of electrical stimuli directly to the skin to create vibration sensations without inducing muscle contractions.
- **Control**: The CONTROL group did not receive any form of haptic feedback. This group served as a baseline to compare the effects of the two haptic feedback methods against no feedback at all.

8.2.1 Participants

A total of 36 participants (N = 36; 14 males, 20 females, 2 non-binary) participated in this study, each completing two sessions, resulting in 72 sessions. The participants had an average age of 25 years (M = 25.91, SD = 4.86). This sample size aligns with those typically employed in motor learning research [398, 399, 400, 401, 402, 403] and EMS studies [393]. Participants were compensated €6 per 30 minutes of participation. Each participant attended two sessions on separate days, with an average interval of 6 days between sessions (M = 6.73, SD = 4.60). No significant differences were observed between groups regarding the time between sessions. Each session lasted approximately 75 minutes, totaling 2.5 hours per participant. Participation was voluntary, and participants were informed that participation could be terminated.

Group Allocation Check:

To control for potential effects of group allocation, we collected data on participants' dominant hand and their drawing skills, measured by the average hours spent drawing per week (self-reported). Additionally, we gathered information on the average hours spent playing video games per week (self-reported). Out of all participants, two reported being left-handed, and they were assigned to different groups; previous work revealed that left-handed individuals did not significantly differ from right-handed individuals in the MD task [402]. As the hours per week self-reported data was found not to be normally distributed, we conducted a Kruskal-Wallis test to determine if there were statistically significant differences between the distributions of CONTROL, EMS, and ELECTROTACTILE groups. The test did not show a statistically significant difference between the groups for hours drawn per week $H(2) = 0.593$, $p = 0.743$, nor for hours playing games $H(2) = 0.059$, $p = 0.970$. Additionally, to control for a priori motor skills of participants, we assessed performance before the first training session consisting of 3 trials. We conducted a Kruskal-Wallis test to compare the means of CONTROL, EMS, and ELECTROTACTILE groups. The analysis showed that there was no statistically significant difference between the groups, $H(2) = 4.38$, $p = 0.112$. These results suggest that the differences between groups at the start of the experiments are not significant and provide a ground for further statistical differences to be influenced by the interventions made during the sessions.

8.2.2 Experimental Design

We conducted a between-subjects study to evaluate the effectiveness of the different haptic feedback methods mentioned above. Participants were randomly assigned to one of three experimental groups: the EMS group, the ELECTROTACTILE group, and the CONTROL group. The study consisted of two sessions; each session involved two assessment stages (i.e., at the beginning and 10 minutes after the training session) and a training session; each training session involved 30 trials, while the assessment involved 3 trials each. Additionally, the first session included a familiarization stage for the participants to understand and ask about the task and the setup. The last session included a Learning Transfer Test stage for assessing how well participants transfer the knowledge to a different shape.

8.2.3 Feedback Rendering

The feedback was rendered using an FDA-approved Sanitas 41 generator connected to a "Let Your Body Move" toolkit [31]. We used Axion EMS/TENS 32mm diameter round electrodes for easier placement and muscle targeting. The electrodes were adhesive and adhered to the user's arm.

EMS Feedback

In this condition, EMS was applied to participants' arms to provide kinaesthetic feedback, guiding them to correct their movements. The feedback was designed to influence muscle contractions, helping participants stay on the intended path without overshooting and avoiding additional corrections. Therefore, corrective actuation was triggered whenever participants deviated from the specified path, which is a common approach in HCI research on motor skill transfer and learning using EMS [11, 404, 405, 406]. The frequency and pulse width of the EMS were set to 150 Hz and 100μs, respectively, and the intensity was calibrated before the study.

Electrotactile Feedback

In this condition, Electrotactile stimulation was applied through the same set of electrodes that participants in the EMS group received feedback. Participants in this group received electrical pulses directly to the skin, which created a tingling sensation. The stimulation intensity was adjusted to be noticeable, but without exerting any movement on the participant, the electrode location was similar to that of the EMS group. This type of feedback is analogous to traditional vibrotactile feedback, which serves as a notification indicating that a correction is necessary. Yet, it does not actuate the participants' bodies but makes them aware and lets them correct themselves; this type of feedback is also typical in HCI [405, 407].

Control

Participants in the CONTROL group received no feedback while performing the task. They completed the task without any external cues, relying solely on their proprioception and observation.

8.2.4 Mirror Drawing Task

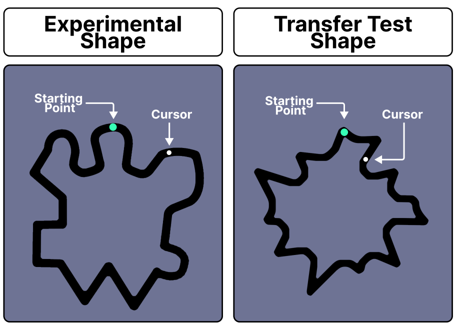


Figure 8.3: Shapes used in the MD task. **Left:** Shape extracted from [408] and used in the main experiment. **Right:** Shape generated for the transfer test. The starting point was the same as the endpoint and was depicted in the interface as a green point; the participant cursor was displayed as a withe circle.

The MD task, a well-established method for studying skill learning since 1910 [397, 398, 399], involves participants tracing a shape (typically a polygon, such as a star, diamond, square, or triangle) while remaining within the boundaries of a double line. The key challenge is that participants can only see an inverted reflection of their hand through a mirror or, in modern setups, a mirrored input mapped in the screen. his setup allows researchers to study how new associations are formed between visual input and corresponding arm movements [409].

The MD task utilized in this experiment is an implementation of the original MD task by Snoddy [410], further developed by Stratton et al. [411], and more recently adapted for delta robot input by Sullivan et al. [408]. In this experiment, participants were asked to repeatedly trace an abstract shape displayed on a computer monitor as quickly and accurately as possible. They interacted with the system using a Novint Falcon delta robot, with position data sampled at 200 Hz. A stiff virtual spring was applied along the Falcon's third degree of freedom (DOF) to constrain movement to a vertical plane parallel to the computer screen. This setup ensured that the horizontal and vertical movements of the Falcon were directly translated to the corresponding movements of the on-screen cursor. However, the horizontal axis was inverted: moving the Falcon to the left caused the cursor to move right, and vice versa.

Shape Selection

In this study we used two shapes in the MD task; First, a *Test Shape:* In previous research, Saquares or star shapes have often been used in the MD task due to their simplicity [398, 402, 403, 412]. Yet, for healthy users, this shape can be overly simple in healthy adults. To introduce a higher level of complexity for our experiment, we selected a shape that has been validated in the literature as sufficiently complex within the context of motor learning [408]. Second, for *Motor Transfer*, we needed a shape that participants had not encountered before [403]. Consequently, we designed a new geometrically irregular shape. Both shapes are illustrated in Figure 8.3.

8.2.5 Apparatus

The experimental setup utilized a Novint Falcon device for input, constrained to one dimension, allowing users to move the robot's end-effector within a 2D plane, similar to the setup described by Sullivan et al. [408]. The device was connected to a Dell G5 laptop running Windows 11, and the experiment was programmed using Unity 3D version 2024.1. To provide different feedback modalities, we used two FDA-approved EMS signal generators (Sanitas 41) and two "Let Your Body Move" toolkits, initially reported by Pfeiffer, Duente, and Rohs [31]. This configuration allowed us to utilize four EMS channels. Additionally, a Manfrotto armrest was employed to prevent participant fatigue and minimize using non-target muscles during the task. The complete setup is illustrated in Figure 8.4.

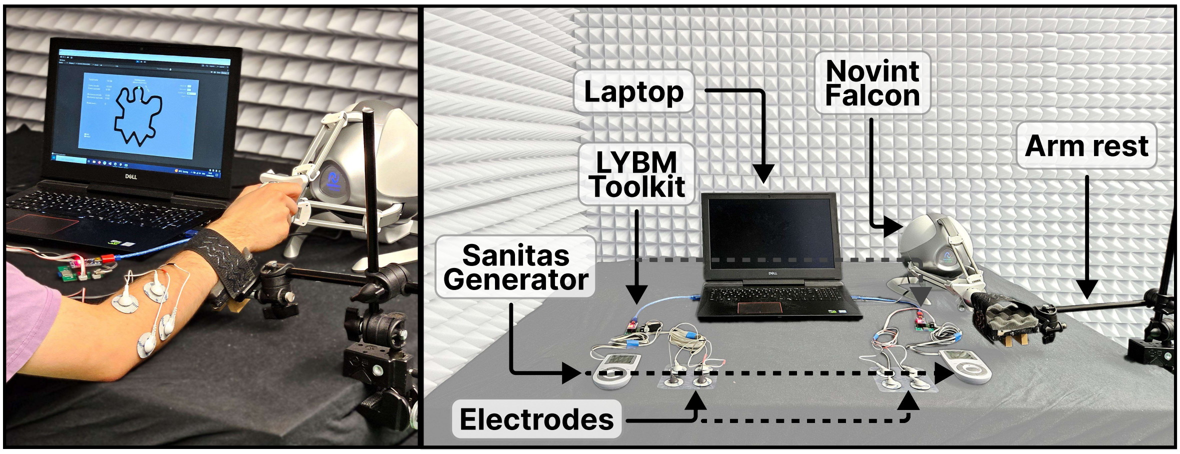


Figure 8.4: Left, User performing the experimental task. Right, Experimental setup featuring a Novint Falcon device constrained to one dimension for input, connected to a Dell G5 laptop running Windows 11, and programmed using Unity 3D version 2024.1. Feedback modalities were provided using FDA-approved EMS signal generators and "Let Your Body Move" toolkits, utilizing four EMS channels. A Manfrotto armrest was employed to prevent participant fatigue and ensure proper muscle usage.

8.2.6 Electrode Placement

To achieve control of two-dimensional movement in the vertical plane, we focused on four key muscle groups at the forearm (which is the most common location for EMS actuation in HCI [393]) that facilitate wrist motion; Radial deviation (leftward movement) is controlled by the Extensor Carpi Radialis (ECR), while Ulnar deviation (rightward movement) is driven by the Extensor Carpi Ulnaris (ECU). Wrist flexion (downward movement) is primarily managed by the Flexor Carpi Ulnaris (FCU), and wrist extension (upward movement) is enabled by the Flexor Carpi Radialis (FCR) [413, 414]. We selected the electrode placement following the setup described by Lopes et al. [415], effectively supporting this range of motion. The specific electrode placements are illustrated in Figure 8.5. For more detailed information on the electrode placement, please refer to [415, 416].

8.2.7 Calibration Procedure

We first focused on the muscle groups mentioned above for the calibration procedure. The participants were instructed to tense the target muscle in the desired direction, and the experimenter positioned two electrodes in the skin over the muscle. The EMS device was then incrementally adjusted, increasing the intensity step by step until either movement was observed or the participant reported mild discomfort.

Once the movement was successfully induced, we transitioned to the computer, where the calibration scene was prepared. We initially set the EMS generator to the intensity at which movement was first observed. The participant's arm was then positioned and secured using the armrest. They moved the mouse to center the cursor on the screen, where a green circle

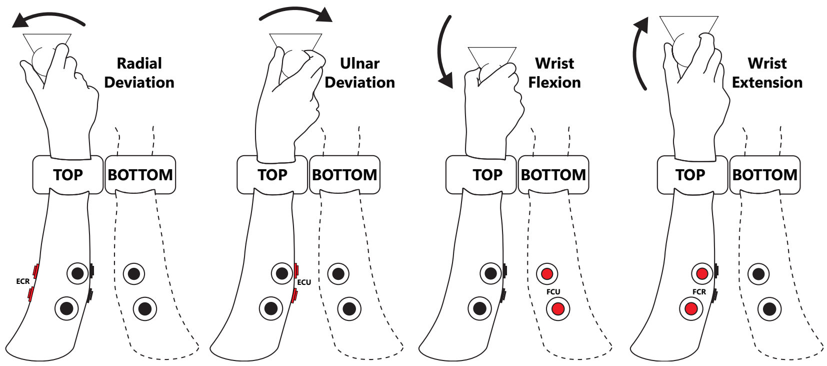


Figure 8.5: Illustration of wrist movements and associated muscle activations. Radial deviation (left) is facilitated by the Extensor Carpi Radialis (ECR), while Ulnar deviation (right) is driven by the Extensor Carpi Ulnaris (ECU). Wrist flexion (downward movement) is primarily controlled by the Flexor Carpi Ulnaris (FCU), and wrist extension (upward movement) is enabled by the Flexor Carpi Radialis (FCR). The diagram indicates the specific muscle groups responsible for each directional movement

appeared at the start of the test. After informing the participants of the upcoming stimulation, they were instructed to remain still while the EMS was active.

When the participant kept the cursor inside the green circle for 3 seconds, the target muscle was stimulated for one second, after which the EMS was deactivated. This process allowed us to assess the effect of EMS on wrist movement under experimental conditions. The intensity was adjusted accordingly if the movement was too pronounced or absent. The goal was to achieve minimal yet observable movement to avoid overcorrections during the task.

8.2.8 Experimental Procedure

Participants attended two experimental sessions. Upon arrival at the first session, they were informed about the study's purpose and provided with an informed consent form. Participants were informed of their right to withdraw from the study without explaining or impacting their compensation. After providing consent, each participant was randomly assigned to an experimental group and seated before a screen. Feedback calibration was performed, and their arm was positioned on the armrest with the elbow resting on the table. The chair height was adjusted for comfort, and the armrest was positioned to support the arm and prevent fatigue, ensuring minimal muscle use during the task. The distance from the Novint Falcon device was also adjusted for comfortable wrist movements.

We explained the MD task to the participants and instructed them to complete the trials "as fast and accurately as possible," following the practice from previous studies [403, 408]. They were given three practice trials without feedback to familiarize themselves with the setup and could ask questions before beginning the experiment. Once the participant confirmed their understanding, they completed a 3-trial pre-intervention motor skills assessment. This and all subsequent assessments were conducted without feedback across all groups.

Following the assessment, the first training session began, consisting of 30 trials with feedback provided based on the participant's group. After the training session, participants took a 10-minute break before performing a second assessment. The first session concluded afterward, and participants were dismissed.

In line with motor learning theory, which emphasizes the importance of sleep for consolidation [390], participants returned for the second session after at least one night of sleep. The second session began with a third assessment and another training session with group-specific feedback. After a 10-minute break, participants completed a fourth assessment and a motor transfer test using a different MD shape. Participants completed NASA TLX [189] and System Usability Scale[317] assessments during both sessions. An overview of the experimental process is illustrated in Figure 8.2.

8.2.9 Measures

In this study, we investigate the impact of EMS feedback on motor learning using the well-established MD task [403]. This task is frequently used to assess motor learning, with two notable variations: measuring the time after completing a fixed length (i.e., time to trace the full shape) [398, 408, 417] or path length/number of shapes achieved within a fixed timeframe [418, 419, 420]. We analyze both metrics across the three primary motor learning assessments (Post-training assessments 1 and 2 and Consolidation assessments) and during a motor transfer assessment conducted in the second session. The following section provides detailed descriptions of these measures.

Path Length:

We measured the distance a participant could accurately trace along a shape's path in 5 seconds. The metric reflects the path length, adjusting for errors where the tracing deviates outside the shape's boundaries. Only correctly traced portions within the borders and in the clockwise direction are counted, with higher values indicating better performance.

Total Time:

The total time, measured in seconds, that a participant takes to complete a shape. This metric indicates the efficiency of the participant's performance, with shorter completion times being better.

Path Exits:

The number of times a participant crosses the boundaries of the shape, specifically when they leave the main body of the shape, is counted. However, the times when the participant re-enters the shape are not included in this count.

Learning Assessments:

To evaluate motor learning in participants, we assessed their motor skills using the metrics outlined above (Path Length, Total Time, and Path Exits) at different time points: after each training session (Post-Training), and after learning consolidation at the beginning of the second session. All assessments were conducted without providing any feedback across the three groups.

- **Post-Training 1:** This assessment took place 10 minutes after the training trials at the end of the first session. It evaluates motor learning during the Fast Learning stage. A significant performance improvement is expected compared to the Pre-training assessment.
- **Post-Training 2:** A similar assessment was conducted at the end of the second session. As with Post-Training Performance 1, a substantial performance improvement is expected compared to the Pre-training assessment.
- **Consolidation:** This assessment occurred at the beginning of the second session, after participants had completed the first training session and had a night of sleep but before undergoing any further training. This session evaluates the consolidated motor skills in the Slow Learning phase. While performance is expected to be better than the Pre-training assessment, it may not surpass the Post-Training assessment, as participants rely on the knowledge consolidated from the previous session, which may not encompass all the gains achieved during the session.

Motor Transfer Assessment

To evaluate the generalizability of the acquired motor knowledge to different motor tasks, we performed a motor transfer assessment, consisting of the Mirror Tracing task with a previously unseen shape.

Learning Rate

We employed an exponential decay function with an asymptote to model the learning across the three feedback groups [421]. Exponential decay models are frequently used to quantify learning rates in motor learning processes [422, 423]. We fitted a three-parameter model to the training data from both sessions in sequence to capture this learning process. Specifically, we concatenated the trials from both training sessions, allowing us to account for trial-level learning throughout the training period. The fitted model was initially proposed by Newell and Rosenbloom [424], which is the following:

$$E(RT) = A + Be^{-\alpha N} \tag{8.1}$$

Where $E(RT)$ is the expected value of the Response Time (RT) under evaluation on practice trial N; A is the expected value of the RT after practice has been completed (asymptote parameter). This parameter can also be viewed as the minimum response time that can be achieved after all the practice trials; B is the change in the expected value of the RT from the beginning of practice to the end of practice (change score parameter); α is the exponential learning rate parameter [421].

Additional Measures:

We assessed the perceived usability of the system using the System Usability Scale (SUS). Participants completed the SUS questionnaire after the study. We also measured task load using the NASA Task Load Index (NASA-TLX [189]). Participants filled out the NASA-TLX questionnaire at the end of each session.

8.3 Results

To determine the appropriate statistical tests for analyzing the variable of interest, we first assessed the normality of the data using the Shapiro-Wilk test across all groups. For each group, we computed the test statistic and the corresponding p-value. If all groups were found to be normally distributed ($p > 0.05$), we proceeded with a one-way ANOVA to evaluate the differences between the groups, followed by Tukey's Honest Significant Difference (HSD) test for post-hoc analysis in cases where a significant effect was observed. However, if at least one group violated the normality assumption ($p < 0.05$), we employed the non-parametric Kruskal-Wallis test instead. When the Kruskal-Wallis test indicated a statistically significant difference between groups, Dunn's post-hoc test with Bonferroni correction was used for pairwise comparisons.

8.3.1 Path Length

We evaluated the PATH LENGTH across the three main assessments: *Post-Training 1*, *Consolidation*, and, *Post-training 2* across the three experimental conditions (CONTROL, EMS, and ELECTROTACTILE). We report the results in the following.

Post-Training 1

To evaluate differences in PATH LENGTH in the Post-Training 1, we conducted a Kruskal-Wallis test, as the data did not meet the normality assumptions required for parametric tests. The test revealed a statistically significant difference between the groups, $H(2) = 6.51$, $p = .03$. Subsequent pairwise comparisons using Dunn's post-hoc test with Bonferroni correction indicated that the EMS condition significantly outperformed the CONTROL condition ($p = .03$). No statistically significant differences were found between the other pairs (all $p > .05$). The

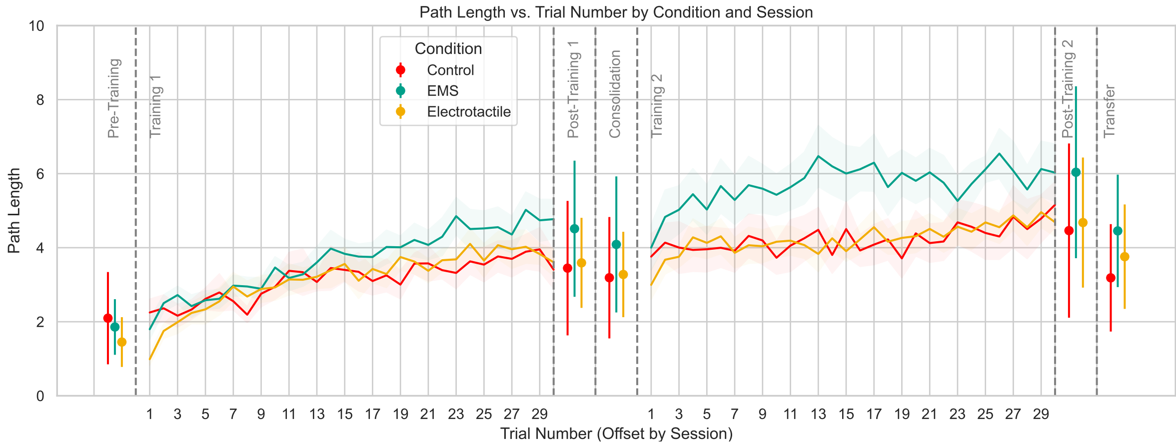


Figure 8.6: Path Length plots throughout the experiment, from the pre-training assessment to the transfer test. This figure offers a trial-by-trial overview, illustrating the progression of the participant's motor skills, as reflected by changes in Path Length over time.

performance ranking, based on mean values, suggests that participants in the EMS group (M = 4.51, SD = 1.84) achieved the highest performance, followed by the ELECTROTACTILE group (M = 3.59, SD = 1.22), and the CONTROL group (M = 3.45, SD = 1.82). These results indicate that the EMS feedback led to longer PATH LENGTH compared to the CONTROL feedback. See Figure 8.6 for an overview.

Consolidation

We conducted a Kruskal-Wallis test to evaluate PATH LENGTH in the Consolidation assessment. The test did not reveal a statistically significant difference in *Path Length* between the CONTROL, EMS, and ELECTROTACTILE groups, $H(2) = 5.81$, $p = .05$. Despite the lack of statistical significance, the mean performance values suggest a trend where participants in the EMS group (M = 4.08, SD = 1.83) performed better on average than those in the ELECTROTACTILE (M = 3.27, SD = 1.15) and CONTROL (M = 3.18, SD = 1.63) groups. The median values further support this trend, with EMS showing the highest median performance (3.61), followed by ELECTROTACTILE (3.20) and CONTROL (2.85). The groups were ranked accordingly, with EMS achieving the highest rank, followed by ELECTROTACTILE and CONTROL.

Post-Training 2

We conducted a Kruskal-Wallis test to evaluate the differences in PATH LENGTH in Post-Training 2. The results revealed a statistically significant difference between the groups, $H(2) = 11.911$, $p = .003$. Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction indicated that the EMS group significantly outperformed the CONTROL group ($p = 0.002$). At the same time, no significant differences were observed between the ELECTROTACTILE and CONTROL groups ($p = .98$) or between the EMS and ELECTROTACTILE groups ($p = .05$). Based on mean performance values, the EMS condition ranked highest (M = 6.04, SD = 2.32), followed by the ELECTROTACTILE condition (M = 4.68, SD = 1.76), and finally the

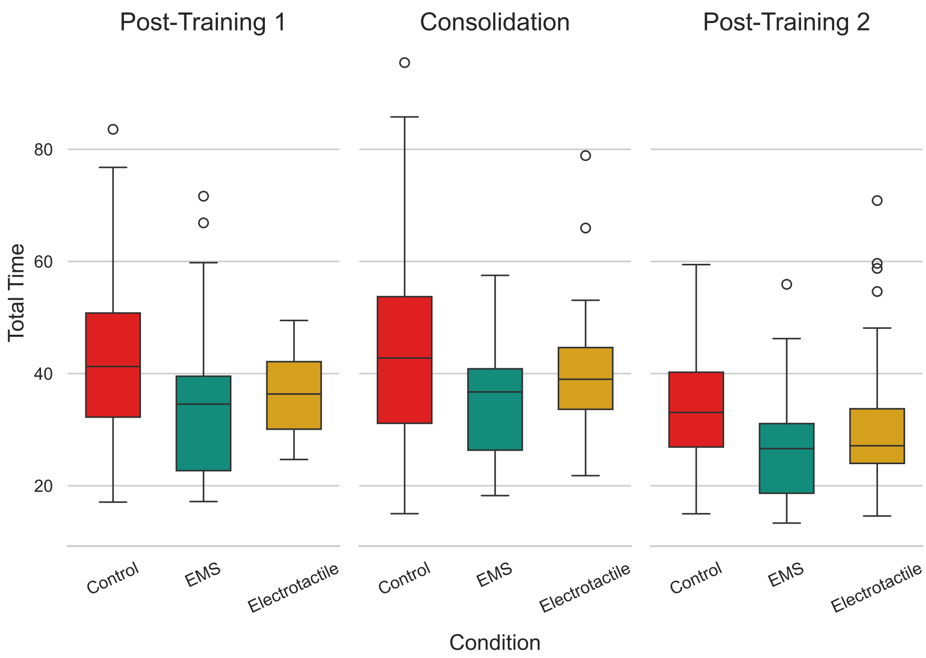


Figure 8.7: Total time across the three main assessments, comparing EMS and ELECTROTACTILE conditions. The EMS and ELECTROTACTILE conditions generally showed better performance in all assessments.

CONTROL condition (M = 4.46, SD = 2.35). These findings suggest that the EMS feedback led to superior overall performance compared to the other feedback groups.

8.3.2 Total Time

We evaluated the TOTAL TIME across the three main assessments: *Post-Training 1*, *Consolidation*, and, *Post-training 2* across the three experimental conditions (CONTROL, EMS, and ELECTROTACTILE). The results are as follows:

Post-Training 1

To investigate the differences in TOTAL TIME in Post-Training 1, we performed a Kruskal-Wallis test due to violating normality assumptions in at least one group. The Kruskal-Wallis test revealed a statistically significant difference between the groups, $H(2) = 9.15$, $p = .01$. We conducted Dunn's post-hoc test with Bonferroni correction to identify the specific group differences. The results showed a significant difference in TOTAL TIME between the CONTROL and EMS groups ($p = .007$), while no significant differences were observed between the other pairwise comparisons (all $p > .05$).

The group performance ranking, based on mean values, indicated that the EMS condition had the lowest mean TOTAL TIME (M = 26.97, SD = 9.74, median = 27.27), followed by the ELECTROTACTILE condition (M = 31.70, SD = 12.48, median = 27.88), and the CONTROL condition had the highest mean Total Time (mean = 34.43, SD = 11.39, median = 33.16). These results suggest that the EMS condition led to significantly faster completion times than the Control condition. In contrast, the Electrotactile condition did not differ significantly from either the Control or EMS conditions.

Consolidation

We evaluated the differences in Total Time in the Consolidation assessment using the Kruskal-Wallis test due to the non-normal distribution of the data. The test revealed a statistically significant difference between the groups, $H(2) = 6.79$, $p = .034$. Dunn's post-hoc test with Bonferroni correction was conducted to investigate these differences further. The results indicated a significant difference in Total Time between the Control and EMS groups ($p = .02$), while no significant differences were found between the other group pairs (all $p > .05$).

The mean Total Time for the EMS group was 35.86 seconds (SD = 10.68), followed by the Electrotactile group at 40.73 seconds (SD = 11.21), and the Control group at 45.89 seconds (SD = 18.07). Ranking the groups based on mean values, the EMS group performed the best, followed by the Electrotactile group and the Control group. These results suggest that the EMS condition led to a significantly lower Total Time compared to the Control condition in the consolidation test.

Post-Training 2

To examine the differences in *Total Time* in Post-Training 2, we performed a Kruskal-Wallis test due to the non-normal distribution of the data. The Kruskal-Wallis test revealed a statistically significant difference between the groups, $H(2) = 8.98$, $p = .01$. Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction indicated that the EMS condition significantly differed from the Control condition ($p = .008$). At the same time, no significant differences were found between the other pairs (all $p > .05$).

Ranking the group performances based on mean *Total Time*, the EMS condition had the shortest mean time ($M = 35.14$, SD = 13.49), followed by the Electrotactile condition ($M = 37.08$, SD = 6.64), and the Control condition had the longest mean time ($M = 44.22$, $SD = 15.81$). These results suggest that EMS was the most efficient condition in terms of total time in Post-Training 2, while the Control condition required the most time on average.

8.3.3 Path Exits

We evaluated the Path Exits across the three main assessments: *Post-Training 1, Consolidation,* and, *Post-training 2* across the three experimental conditions (Control, EMS, and Electrotactile). The results are as follows:

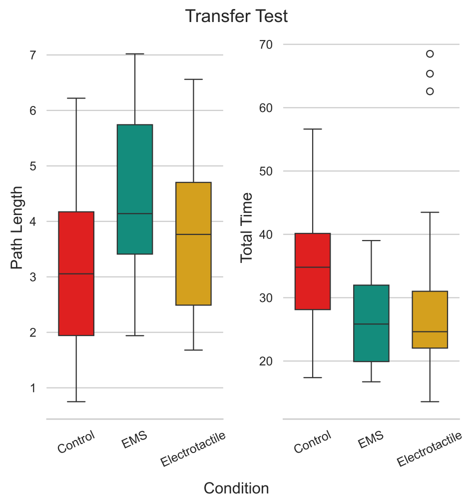


Figure 8.8: Comparison of Path Length and Total Time across conditions during the Motor Transfer test. The EMS condition resulted in superior performance in both metrics.

Post-Training 1

To evaluate the differences in PATH EXITS in Post-Training 1, we conducted a Kruskal-Wallis test due to the non-normal distribution of the data. The test revealed a statistically significant difference between the groups, $H(2) = 10.33$, $p = .006$. Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction indicated a significant difference between the CONTROL group and the EMS group ($p = .006$). Although no significant differences were found between the other group pairs (all $p > 0.05$), the ranking of group performance based on mean values showed that the EMS condition (M = 2.42, SD = 2.36) had the highest number of PATH EXITS, followed by the ELECTROTACTILE condition (M = 2.03, SD = 2.56), and the CONTROL condition (M = 1.14, SD = 2.27) had the lowest. These results suggest that the EMS condition led to a significantly higher number of PATH EXITS compared to the CONTROL condition.

Consolidation

We conducted a Kruskal-Wallis test to evaluate the differences in PATH EXITS in the Consolidation, as the data did not meet the assumptions for parametric testing. The Kruskal-Wallis test revealed no statistically significant difference between the groups, $H(2) = 3.42$, $p = .18$.

Post-Training 2

To analyze differences in the dependent variable *Path Exits* in the Post-Training 2, we conducted a Kruskal-Wallis test due to the non-normality of the data. The test revealed a statistically significant difference between the groups, $H(2) = 6.11$, $p = .04$. Post-hoc pairwise comparisons using Dunn's test with Bonferroni correction showed no significant pairwise differences between any two conditions (all $p > 0.05$).

8.3.4 Transfer Tests

We analyzed the Motor transfer with a different shape in the MD task at the end of the last training session. Here we analyze the motor transfer performance across the main motor learning operationalization of the MD task: Path Length, Total Time, and Path Exits; the results are as follows:

Path Length

To assess the feedback impact on *Path Length* for motor transfer, we conducted a one-way ANOVA followed by Tukey's (HSD) post-hoc test. The ANOVA revealed a statistically significant effect of condition on *Path Length*, $F(2, n) = 6.81$, $p = .002$. Post-hoc comparisons using Tukey's HSD test showed that the mean difference between the CONTROL and EMS groups was significant (MD = 1.26, $p = .001$), with the EMS group demonstrating a higher mean PATH LENGTH. However, no significant differences were found between the other group pairs (all > $p = 0.05$). The ranking of group performance based on mean values was as follows: EMS (M = 4.45, SD = 1.52), ELECTROTACTILE (M = 3.76, SD = 1.41), and CONTROL (M = 3.19, SD = 1.45). These results suggest that the EMS condition led to a significantly higher *Path Length* compared to the CONTROL condition.

Total Time

To evaluate the differences in TOTAL TIME in the motor transfer test, we conducted a Kruskal-Wallis test, as the assumption of normality was not met. The results indicated a statistically significant difference between the groups, $H(2) = 13.73$, $p = .001$. Post-hoc comparisons using Dunn's test with Bonferroni correction revealed significant differences between the CONTROL group and both the EMS group ($p = .004$) and the ELECTROTACTILE group ($p = .004$), while no significant difference was found between the EMS and ELECTROTACTILE groups ($p > .05$).

The mean *Total Time* values for each group ranked the EMS condition as the fastest ($M = 26.88$, $SD = 6.92$), followed by the ELECTROTACTILE condition ($M = 28.73$, $SD = 12.86$), and the CONTROL condition being the slowest ($M = 34.48$, $SD = 9.86$). These results suggest that both the EMS and ELECTROTACTILE conditions resulted in significantly faster task completion times compared to the CONTROL condition, with the EMS condition being the most efficient overall.

Path Exits

To evaluate the differences in PATH EXITS in the motor transfer test, we conducted a Kruskal-Wallis test due to the non-normal distribution of the data. The Kruskal-Wallis test revealed a statistically significant difference between the groups, $H(2) = 9.06$, $p = .01$. Subsequent pairwise comparisons using Dunn's post-hoc test with Bonferroni correction indicated a significant difference between the CONTROL and ELECTROTACTILE conditions ($p = .012$), while the comparisons between CONTROL and EMS and between EMS and ELECTROTACTILE were not statistically significant.

Group performance rankings, based on the mean values of *Path Exits*, indicate that the CONTROL group had the lowest mean (M = 1.06, SD = 1.82), followed by the EMS group (M = 2.03, SD = 2.01), and the ELECTROTACTILE group with the highest mean (M = 2.61, std = 2.72). These results suggest that participants in the CONTROL condition experienced fewer path exits compared to those in the ELECTROTACTILE condition, with the EMS condition showing intermediate performance.

8.3.5 Learning Model Parameters

Using nonlinear least squares (NLS) regression, we fitted an exponential decay model to the response times across the three experimental conditions on a population level [421, 423]. We dynamically estimated starting values for the model parameters to improve the fitting process. We then extracted the coefficients (A, B, and α) from the fitted models. A visualization of the fitted models is shown in Figure 8.9, and the resulting parameters are presented in Table 8.1

Table 8.1: Model coefficients

Condition	A	B	α
Control	33.019	25.813	.046
EMS	24.802	34.602	.053
Electrotactile	34.553	29.297	.094

Based on the extracted coefficients of the Exponential Decay model, the CONTROL condition shows a higher asymptotic response time ($A = 33.019$) compared to EMS ($A = 24.802$) and ELECTROTACTILE ($A = 34.553$), indicating that participants in the EMS condition achieve the fastest minimum response time after practice. The EMS condition also demonstrates the most significant change in response time from the beginning to the end of practice ($B = 34.602$), suggesting an improvement in performance over time. Interestingly, the ELECTROTACTILE condition exhibits the highest learning rate ($\alpha = .094$), implying that participants in this group adapted more quickly during practice, even if their final performance (as reflected in A) was not as low as in the EMS condition.

8.3.6 Task-Load

To evaluate the differences in TASK-LOAD Across conditions, we conducted Kruskal-Wallis test across the six subscales of the NASA-TLX questionnaire, as the data did not meet the assumptions for parametric testing. The Kruskal-Wallis test revealed no statistically significant difference between the groups in any subscale (all $p > .05$).

8.3.7 Usability

To analyze differences in the System Usability Scale (SUS) scores across the three conditions a one-way ANOVA was conducted. The ANOVA revealed a statistically significant difference between the groups, $F(2,36) = 3.976$, $p = .028$. Post-hoc analysis indicated a significant difference between the CONTROL group and the ELECTROTACTILE group (MD = 14.90, $p = .023$), with the ELECTROTACTILE group demonstrating higher SUS scores. No significant differences were found between the other group pairs (all $p > .05$).

8.4 Discussion

In this paper, we present an empirical evaluation of the effects of EMS on motor learning, comparing it to two other conditions: an ELECTROTACTILE feedback condition, representing the state-of-the-art feedback type, and a CONTROL condition with no feedback intervention as a baseline. Our study examined motor learning across three key phases: fast learning, consolidation, and motor transfer. We aimed to explore the tension between recent HCI research, which suggests that EMS can enhance motor learning, and traditional motor learning theories that emphasize the importance of repeated practice for the creation of sensorimotor representations through perception, reflection, and correction of motor actions. Our findings reveal that the EMS group outperformed both the ELECTROTACTILE and CONTROL

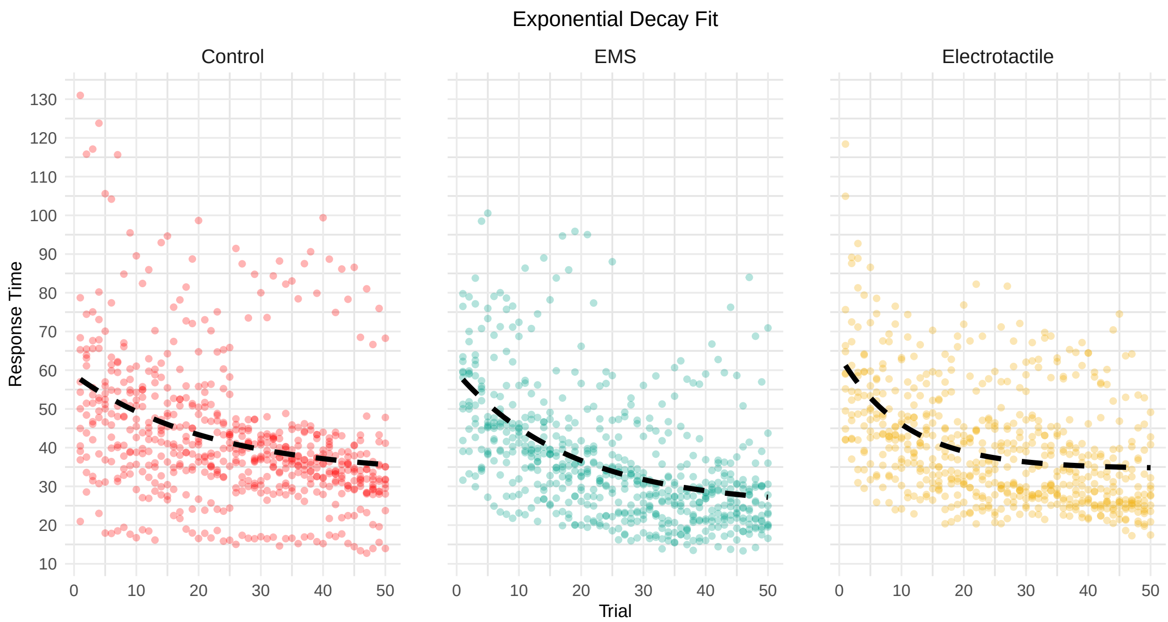


Figure 8.9: Exponential decay model fitted to the data for the experimental groups. The lines represent the model's predicted values, while the scatter points indicate the actual TOTAL TIME data recorded for each trial under each condition. The TOTAL TIME to complete the shapes was used as the Response Time variable in this analysis.

groups across all phases of motor learning, with the ELECTROTACTILE condition yielding intermediate results. However, there are important considerations when selecting feedback mechanisms for motor tasks, which we discuss in detail in this section:

8.4.1 EMS as an augmentation technology for motor skills:

Previous research has consistently shown the efficacy of EMS in temporarily enhancing motor skills, such as reaction time and posture correction. While EMS has been associated with improved motor skill transfer, questions remain as to whether these improvements reflect genuine learning or are merely temporary augmentations dependent on active stimulation. In this paper, we present empirical evidence confirming that EMS-supported motor learning can result in lasting skill acquisition, extending beyond temporary performance gains. Our results demonstrate that EMS outperforms traditional haptic feedback during sessions, corroborating the immediate augmentation potential reported by Kasahara, Nishida, and Lopes [23] and Tatsuno, Hayakawa, and Ishikawa [395]. However, we show that EMS not only enhances immediate performance but also promotes motor skill learning, suggesting a deeper connection between augmentation and learning than we previously assumed. *In this sense, **H1** could be rejected, as during the motor learning task,* EMS *showed better results than* ELECTROTACTILE *and* CONTROL.

8.4.2 Typical feedback modalities remain useful for motor learning

Our experimental results show that while ELECTROTACTILE feedback led to lower performance in some assessments, it significantly outperformed the CONTROL condition, making it a viable alternative for learning. According to the exponential decay model, ELECTROTACTILE feedback yielded a higher learning rate than EMS, reaching stable performance faster, though less intensively than EMS. Despite this, ELECTROTACTILE feedback demonstrated its effectiveness. Additionally, it offers practical advantages: it is easier to implement, requires less exhaustive calibration, and is more suitable for a wider range of users, as EMS can cause discomfort for some and targeting specific muscles can be challenging. Thus, ELECTROTACTILE feedback is not necessarily inferior to EMS but may be better suited for different scenarios depending on user goals. *However, we reject **H2**, as* ELECTROTACTILE *feedback did not result in better motor skill consolidation than* EMS. *Nonetheless, the results support **H3**, as* ELECTROTACTILE *feedback led to a higher learning rate than* EMS.

8.4.3 EMS for motor learning: Does it support the learning process?

Our results provide empirical evidence that EMS not only enhances motor skills during use but also facilitates motor learning, viewed as the sustained improvement of a skill even after the removal of the EMS device—both immediately following training and after a delay of one or more days. Furthermore, the findings demonstrate that ELECTROTACTILE feedback also supports motor learning, consistent with prior research suggesting similar feedback mechanisms can promote learning. Contrary to our initial hypotheses, which were informed by the literature, EMS in this experiment led to greater overall motor learning compared to the CONTROL condition across various assessments and metrics, although the rate of learning was lower than that observed with ELECTROTACTILE feedback alone. *In consequence, given that* EMS *led to a higher overall learning, we reject* ***H4****.*

EMS and Electrotactile Feedback for Motor Transfer

Similarly, the effects of EMS extended to new motor tasks, indicating that the learning was not confined to the original motor skill but had been sufficiently internalized to transfer across different contexts. ELECTROTACTILE feedback demonstrated comparable performance under the conditions reported in this experiment. *Given this evidence,* ***H5*** *does not hold, as, contrary to our initial hypothesis,* EMS *did not lower the performance but instead resulted in a higher performance than the other two conditions.*

8.4.4 Potential Neurophysiological Mechanisms of EMS for Supporting Motor Learning

Learning models emphasize the importance of awareness during training to enhance the effectiveness of trial and error. Faltaous, Koelle, and Schneegass [393] report that EMS first induces action, followed by perception and reflection, which could influence the learning process. However, as our results demonstrate, this did not hinder learning; in fact, EMS outperformed other conditions. We attribute this to the alignment of participant intention with EMS actuation throughout the experiment. Specifically, participants aimed to correct their path, and EMS facilitated this by actuating the wrist, potentially contributing to a heightened sense of agency. Kasahara et al. [406] similarly found that participants exhibited increased reaction times after EMS actuation, but only when agency was sufficiently present, whereas conditions with no actuation or agency did not show this effect. Our findings suggest that, beyond the action-perception-reflection sequence, the sense of agency—particularly how participant intention aligns with EMS stimulation—plays a critical role in the learning process. Therefore, future research should explore how varying levels of agency affect motor skill training with EMS support.

Another possible explanation for this effect is that, contrary to the sequential model proposed by Faltaous, Koelle, and Schneegass [393], action and perception may occur simultaneously. In this case, users might be learning while EMS is stimulating their body, in addition to subsequent reflection. Supporting this, Hagert et al. [425] demonstrate that EMS stimulation on the wrist triggers a proprioceptive response, which could provide supplementary feedback alongside kinaesthetic information.

8.4.5 Implications for Motor Learning in HCI using EMS

Despite previous claims regarding the potential of EMS in supporting motor learning, concrete empirical evidence has been lacking. Prior research primarily focused on short-term effects, often limited to a single session, leaving the broader impact of EMS on long-term motor learning unclear. Moreover, the distinction between EMS merely augmenting motor performance and genuinely supporting motor learning has not been fully established.

Our findings provide empirical evidence that EMS not only enhances immediate motor performance but also contributes to long-term motor learning. However, our results indicate that traditional feedback mechanisms [408]—which provide users with additional information and allow them to make their own corrections—still lead to higher learning rates. Thus, while EMS is effective, it may not entirely replace more conventional feedback approaches for motor learning, especially when it comes to fostering independent error correction and self-guided improvement.

Nevertheless, EMS shows significant promise in reducing the learning ceiling often observed with traditional [408] feedback methods. By offering direct physical guidance, EMS can speed up the learning process, especially for tasks where users struggle to make appropriate corrections independently. Future research should explore the potential of EMS to complement, rather than replace, traditional feedback systems in motor learning tasks, particularly for users with different learning capabilities.

8.4.6 Recommended Practices for EMS in Motor Learning and Augmentation

Based on the result presented in this manuscript, and previous works the following best practices are recommended to ensure the validity and reliability of results when using EMS in motor learning or augmentation studies:

- **Differentiate Learning vs. Augmentation:** Clearly distinguish between experiments aimed at enhancing motor learning and those focused on augmenting motor performance. Learning should be assessed over time, whereas augmentation effects can be measured immediately after, or during EMS intervention.
- **Allow Time for Learning:** When testing for learning outcomes, provide adequate time for participants to consolidate motor skills, either during or after the session. For motor learning studies, schedule sessions with sufficient time in between, ideally with a sleep interval, as this supports skill consolidation [390]. To accurately assess motor learning, design studies ideally would involve at least two sessions, separated by a period of sleep. This helps isolate the long-term effects of EMS on learning from short-term performance enhancements.

- **Ensure Agency:** Guarantee a sufficient sense of agency during EMS interventions. Participants should feel that their intentions and actions are aligned with EMS stimulation, as agency is a crucial factor in effective motor learning and performance augmentation.
- **Skill Assessment Before Intervention:** Assess participants' baseline motor skills before introducing EMS. This will provide a clearer understanding of the effects of EMS on motor learning or augmentation and allow for a more personalized approach to EMS intensity and feedback.
- **Avoid Repeated Measures for Learning Assessments, Introduce rest intervals for Augmentation Assessments:** When testing motor learning, avoid using repeated measures designs that could confound results with practice effects. If augmentation is being assessed, introduce a rest interval between blocks to account for the immediate effects of EMS, as demonstrated in the findings of Kasahara et al. [406]

8.4.7 Limitations

While this work aims to comprehensively address the effects of EMS feedback on motor learning, several limitations of the current setup must be acknowledged. First, although we cover multiple phases of motor learning, we do not extend to the most advanced phases, such as automatic execution. This would require a significantly higher number of sessions, which, given the three conditions and participant numbers, would be resource-intensive. However, these post-consolidation stages are equally important for the motor learning process, particularly for participants aiming to achieve high levels of skill in a given task. Additionally, our measurements were limited to behavioral responses, lacking physiological data such as EEG or fMRI that could provide further insights into the mechanisms of EMS in motor learning. Finally, we did not evaluate participants' sense of agency, which could offer valuable information about the impact of agency levels on learning effectiveness.

8.4.8 Next Steps in Motor Learning using EMS

In advancing our understanding of motor learning with EMS, several avenues for future work are identified; **Exploring the underlying mechanisms of EMS in motor learning:** While the current research demonstrates the potential of EMS in enhancing motor learning, a deeper investigation into its physiological mechanisms is necessary. Utilizing tools like EEG or fMRI could provide valuable insights into how EMS influences neural pathways and motor control systems during learning. **Determining the ceiling effect of EMS in motor learning:** It remains unclear whether there is a point at which EMS reaches a threshold of effectiveness in motor learning. Future studies should aim to identify whether there is a diminishing return in skill acquisition with prolonged EMS exposure or if it continues to offer incremental benefits over time/sessions. And, finally **Agency as a critical factor:** Preserving a sense of agency remains a key consideration in EMS-based interventions, as highlighted by Kasahara et al. [406]. Understanding how different levels of agency affect motor learning outcomes will be crucial in designing more effective EMS applications, and a logical next step in light of the results presented in this paper.

8.5 Wearable Vibrotactile Feedback For Motor Skill Transfer in Dance Learning

Dance plays a crucial role in human well-being and expression. To learn dance, transferring motor knowledge across humans is relevant. Several technologies have been proposed to support such knowledge transfer from teacher to student. However, most of such systems applied a pragmatic approach focused on the feedback and the quality of the feedback system and not necessarily on the human mechanisms behind the dance learning process. In contrast, we inquire about the teacher-to-student motor knowledge transfer from the neural perspective to design motor learning wearable systems. We conducted interviews with dance students and teachers using vignettes based on motor learning theory as a discussion base. We derived insights about dance learning and identified a series of requirements for motor skill transfer-focused wearable devices. Based on our results, we present a prototype that reflects the minimum functional setup for effectively supporting motor learning.

8.5.1 Requirements Analysis

We aimed to design a system that supports motor learning in dance. To that end, we conducted a user-centered design process. As a first step, to build an understanding of the requirements of such a system, we conducted an in-depth literature analysis of existing motor learning theories. More precisely, we explored how the neurobiological theory behind motor learning can support wearable devices expected to support dance motor skill transfer. As a second step, we conducted semi-structured interviews with dance students and dance teachers. During the interviews, we discussed vignettes that were based on motor learning theory with the participants. This approach allowed us to blend the empirical knowledge from literature and practical knowledge from dance experts.

8.5.2 Interviews

Learning a motor skill such as dancing usually involves (at least) two parties, the dance student and the dance teacher. These two parties have different perspectives regarding the dance learning process. We conducted semi-structured interviews (duration 45 minutes on average) to understand these different perspectives, thus gathering information from dance experts in the motor learning context. Due to the COVID-19 pandemic, the interviews were conducted via video conferencing software and audio-recorded after participant’s informed consent. The interviews were conducted in the participant's native language, transcribed verbatim, and translated by a professional transcription and translation service. Based on the analysis of the expert interviews combined with the literature analysis, we established the requirements for a technology that supports knowledge transfer in the dance context.

Based on the four stages of motor learning by Doyon et al. [383], we developed a series of vignettes. The vignettes described an imaginary student going through the four stages of motor learning. We focused on the first four stages of motor learning as they depend on the active practice of a motor skill. We discussed every vignette in detail with each participant. This was followed by an open discussion about technology support for dancing. We adapted the interview protocols slightly to the needs of the respective user group (teachers, students). The vignettes and the interview protocols are provided in the supplementary material.

Participants

We interviewed 11 dance experts (2 male, 9 female). Five participants were dance teachers, aged 24–56 (referred to as $T[1-5]$). Six participants were dance students, aged 21–54 (referred to as $S[1-6]$). The participants practiced diverse dance styles, with six practicing a Ballroom dance (e.g. Tango, Jive, Waltz). The remaining five practicing a performing dance style (e.g. Ballet, Jazz dance, Modern dance, Belly dance). To provide meaningful insights, we required the level of the students to match at least phase 4 (automatic phase of skilled behavior). The participants were compensated with 10€ per hour.

8.5.3 Results

We applied an iterative analysis process combining open coding and thematic analysis in line with Blandford et al. [30]. As a first step, two authors coded 35% of the material (i.e. 2 student/2 teacher interviews). Through a set of iterative discussions, we established an initial coding tree used by one of the authors to code the remaining material. We then used

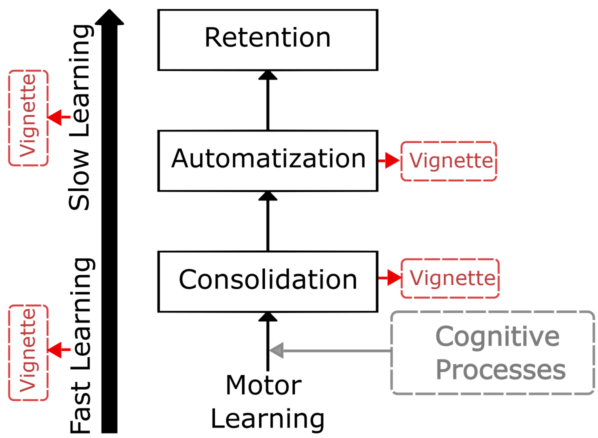


Figure 8.10: Motor learning process (adapted from [383]).

thematic analysis to derive a more abstract understanding of the data. We identified five themes in the interview data: *Teaching and learning style*, *Difficulties*, *Technique and health aspects*, *Feedback*, and *Vision* (for an overview see Table 8.2). Based on our analysis of the interviews and the theoretical background, we derived four central design requirements.

Takeaway:
The system should:

- be a lightweight, wearable system;
- offer personalized, implicit feedback for single dance figures;
- offer the possibility to combine the system with an explicit, audio-visual introduction to a dance figure, thus mirroring the structure of a classic dance lesson;
- integrate sonification elements.

The importance of having a **lightweight system** that does not hinder the variety of different dance moves was addressed by both dance students and dance teachers. The social context of many dance events further reinforces this need, as illustrated by the following statement of a dance student:

> “On the one hand, sure, it's practical. On the other hand, it would probably be too much effort for me. Because either I have to carry something, I have to set something up, or attach something so that it works in a precise way. That would then probably be too much work for me. I'd have to take care of it. Most of the time I go dancing with friends, which is Friday nights where you go out together. If I want to use it actively then, I would have to make sure that I wear my sensors and at the same time have the program running, which measures it, and besides, make sure that I get the feedback. I don't think I would want that. (S1)”

Another requirement that became apparent during our analysis was the need for **personalized, implicit feedback focusing on single dance figures**. However, relying on personalized feedback from technology without the involvement of dance teachers was still difficult to imagine for some dance students:

> “What I would like from such a system is the possibility to get individual feedback from the teacher. This would imply that I have sensors somewhere on the corresponding parts of my body. [The data] arrives visually processed at my teacher so that they can give some kind of feedback in this regard, e.g.

Table 8.2: Overview of identified code groups from the interviews with explanations and examples.

Group	Short Description	Example
Teaching and learning style	Approaches that the participants use when they learn or teach a new movement	Showing the entire figure at the beginning
Difficulties	Common problems and classic mistakes when learning new dance movements	Complex dance moves
Technique and health aspects	Aspects that are considered relevant to learn the correct dance technique in a healthy way	Posture relevant
Feedback	How the feedback is transmitted and perceived	Explicit feedback after dance
Ideas for setup/ helpful tool	Approaches and technology the participants imagined for a dance learning assistant tool	Sensor positioning

> hands, feet, whatever you would need. I imagine a replication of the human body that is somewhere with my teacher and does what I do, and then they can explicitly say, "Hey, take your shoulders back a bit" or something like that. (S3) ”

Going beyond technology-mediated person-to-person feedback, many participants expressed interest in **implicit feedback**. For instance, one dance student reflected on feedback in the form of vibrations during dance practice:

> “ I can imagine the haptic feedback quite well. I think that would be the most intuitive for me now, I would say, or what I would like to have as a support, for example, that when I know, I have to move the right foot backward faster here, that I then get a slight vibration or something that [shows me the direction]. (S4) ”

Both dance teachers and dance students **envisioned recreating their dance classes** at home. However, while students mainly envisioned situations where the technology takes on the role of their dance instructor, dance teachers were more prone to retaining their expert role and using the technology to provide a new perspective:

> “ There could be an app that has two screen pages. One screen shows me what to do, and the other allows me to record myself. Then I can compare my recording one-to-one with the example. (T4) ”

Music and dance are deeply intertwined activities. In this respect, it is natural that **sonification** emerged as one of the key requirements. For instance, music can serve as an orientation point for training certain dance sequences:

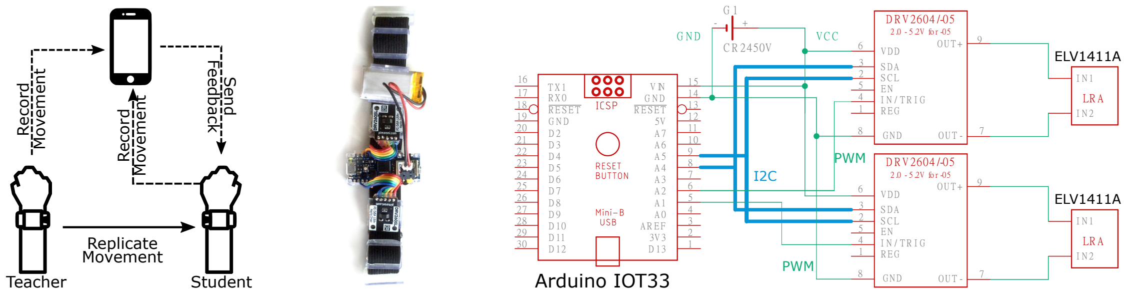


Figure 8.11: System architecture, prototype, and schematic.

> Picking out the parts of the music where you got it wrong and playing them several times so that you can practice them more often might also be a possibility. That you are told afterward, 'You didn't get that part quite right. That's how it would be correct. Now try it again', and then it's played back x number of times. (S3)

On the other hand, the variation of certain musical elements could serve to gradually learn to dance in harmony with the music:

> That the beats become clearer. But I think that would be a way for someone to passively get used to dancing to the right beat, because I've noticed that the longer people dance, the less you have to count in, the more they have a sense of what's right and what's wrong. And of course, you could influence that in such a way that they don't even notice that they've made a mistake. (T1)

To reiterate, based on our literature analysis and the expert interviews (using theory-based vignettes), the design requirements called for a lightweight, wearable system that offers personalized, implicit feedback for single dance figures and integrates sonification elements. Furthermore, the system should offer the possibility to combine it with an explicit, audio-visual introduction to dance figures, thus mirroring the structure of classic dance lessons.

8.5.4 Design

We developed a prototype inspired by the recommendations provided by the teachers and students and summarized in subsection 8.5.1. The system (Figure 8.11) includes a mobile app and a set of wearable bands with vibrotactile feedback and Inertial Measuring Units. We focus on transferring knowledge through implicit feedback; avoiding explicit instructions as communicating dance movements can be challenging for both teacher and student. Our prototype addresses this point in two ways by (a) providing vibrotactile feedback in real-time when the student is replicating the movement and by (b) providing auditory feedback through sonification. We make this possible by recording the teacher's movements and later recording and calculating a correlation between the student's and teacher's movements. When the student follows the movement accurately, no feedback is provided. When the movement differs, the vibrotactile bands vibrate based on the correlation value, and the reverb of the song is altered accordingly. The user's goal is to keep the song playing and avoid the activation of the bands.

8.5.5 Discussion and Implications

The interviews showed that practice, feedback, and teaching the technical details of a movement are the most critical factors for improving dancing skills. Participants highlighted the role of feedback in early stages since it is harder to correct movements that already reached the automatic phase. The importance of repetitive practice is often stressed in the dance context. Therefore, the system should support continuous practice across motor learning stages. Also, approaches that prioritize implicit over explicit feedback can better support dance students. Furthermore, approaches such as sonification or real-time haptic feedback are well received by teachers and students [426]. In contrast, summaries or statistics after a session or the execution of a movement are less valued.

Besides, we observed a difference in attitudes towards non-dance-related training. These differences in attitudes were also reflected in the opinions about a technical setup. While T2, T3, and T5 value additional exercises to reduce the risk of injury and allow healthy dance training, for T1, it does not belong to the hobby field of dancing but competitive dancing. One explanation for this may be the different dance styles. The participants who practice a performing dance type often shared the same opinion that differed in part from participants who practiced a Ballroom dance style.

9

VIRTUAL REALITY AS A PLATFORM FOR AUGMENTATION EXPLORATION

This chapter is based on the following publications:

Exploring Virtual Reality as a Platform for Early-Stage Design for Human Augmentation Technologies - Steeven Villa, Yannick Weiss, Robin Neuhaus, Marc Hassenzahl - *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia 2024*

After exploring the relationship between learning and augmentation, this chapter shifts focus to the use of Virtual Reality (VR) as a platform for prototyping augmentations, addressing the challenges of developing real-world Human Augmentation Technologies. VR offers a flexible environment to design and test augmentations, which are often complex to develop in physical form.
In this chapter, I investigate the potential of VR for early human augmentation technology development within an educational framework. Through a semester-long course, students designed virtual augmentations in a VR environment, resulting in three VR applications. Interviews with four participants revealed both opportunities and limitations: while VR facilitated creative exploration and immersive experiences, issues like simulator sickness and the lack of haptic feedback constrained the development of physical augmentations. The findings underscore the need for guidelines based on best practices for virtual augmentations and highlight VR's potential as a tool for early-stage prototyping and research in human augmentation.

9.1 Virtual Reality as A Platform for Understanding Human Augmentation

Human augmentation technologies improve human performance to a level that would not be possible otherwise [22, 23]. These technologies could change how people interact with their surroundings and execute tasks that require specific physical, mental, or sensory skills [35, 68]. Recent advances in artificial intelligence (AI) [427], augmented reality (AR) [428], prosthetics [429], robotics [430], and wearables [Exp8], among other technologies [431], have made tangible the vision of enhancing human skills. However, despite these advancements, the inherent complexity and functional diversity of augmentation technologies pose significant challenges in their prototyping, development, and testing [Het1].

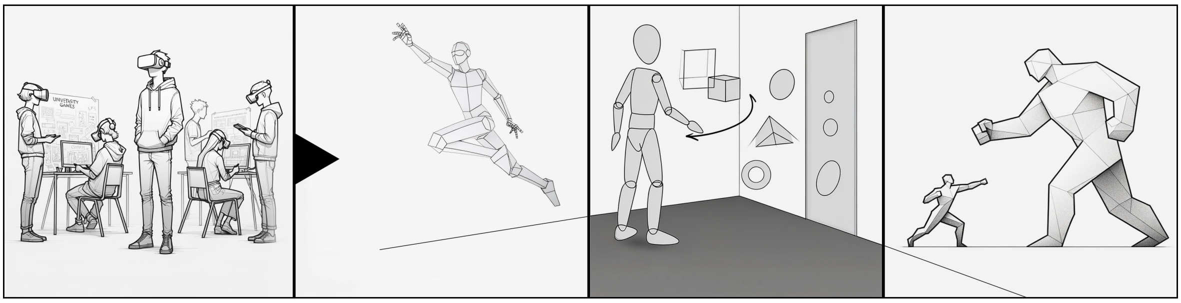


Figure 9.1: This work investigates the potential of Virtual Reality (VR) as a platform for early-stage design and exploration of augmentation technologies. A semester-long VR development course was conducted with 19 participants divided into three groups. Each group designed and implemented VR prototypes focusing on specific human augmentation technology functionalities: jumping augmentation, object manipulation augmentation, and super strength augmentation.

In response to these challenges, recent research has explored the potential of Virtual Reality (VR) as a platform to prototype augmentation technologies by leveraging the concept of "superpowers" within VR [432]. This approach turns VR's inherent limitations in replicating the physical world into an advantage by shifting the focus from replicating reality to creating experiences otherwise impossible [433]. Thus, by intentionally exceeding these limitations, VR-based superpowers can explore functionalities that are currently unrealizable in the real world, potentially serving as inspiration for future human augmentation technology development. Furthermore, VR-based prototyping offers a controlled environment, mitigating the complexities and risks associated with real-world human augmentation technology testing. This establishes VR as a plausible test bed for early-stage human augmentation technology development and exploration.

Therefore, in this work, we leverage these frameworks to explore the ideation and development process within a VR educational setting. As part of a VR development course, students designed and developed human augmentation technology concepts in VR. Following the course, we interviewed a subset of students about their experiences developing human augmentation technology concepts in VR.

9.2 Methods and Procedure

As part of a university course on the topic of VR that spanned one semester, we conducted a week-long practical exercise to investigate and prototype Human Augmentation (HA) in VR. The course was composed of 19 students from master's and bachelor's programs in media informatics and computer science. The course structure comprised bi-weekly sessions throughout the semester, featuring theoretical lectures on VR and HA alongside practical exercises for implementation. In the practical exercise described here, student groups were tasked with developing a game in VR making use of virtual augmentation. The prototypes were required to utilize Electromyography (EMG) for user input, incorporating multiplayer functionality, enabling interaction between participants within the VR environment and the possibility of non-verbal communication. The initial sessions introduced the concept of virtual augmentation. This foundation was established through a lecture drawing upon the framework of human augmentation presented by Raisamo et al. [35]. Subsequent sessions employed facilitated brainstorming activities to encourage creative ideation. This fostered the generation of novel human augmentation technology concepts. Drawing on the acquired theoretical and practical knowledge, student groups followed an iterative ideation process, progressively refining their VR augmentation concepts. A block week in the lecture-free period was dedicated to finalizing prototypes, after which the students presented their applications.

In the weeks after the presentation and after grades for the seminar had been finalized, we reached out to participants of the practical exercise for their availability to voluntarily take part in an interview about their experience developing the VR game. We conducted semi-structured interviews with four participants (2 identified as female, 2 as male, aged 25-29 years) comprising three different project groups. Two participants were master's students in Human-Computer Interaction, one a master's student in IT, and one a bachelor's student in computer science. Two participants reported having previous experience developing VR applications, while the other two had no experience.

In the interviews, after an introduction, we asked participants about the game they developed, its key features, and the ideation process. In the next section of the interview, we asked about the game's development process. Finally, we asked about virtual and human augmentation in general, and participants elaborated on potentials and challenges based on their experience in the exercise. To conclude, we allowed participants to share further thoughts or ask questions.

After obtaining permission, we audio-recorded and transcribed the interviews for analysis. Then, we conducted a content-analysis as described by Mayring [434]. The first 3 authors coded the data separately and then collated the results.

(a) JumpiFlumpi (b) PortalGhost (c) GuidingLight

Figure 9.2: Screenshots of the prototypes developed by the course participants showcasing the Virtual Environments

9.3 Virtual Augmentation Prototypes

The course culminated in the development of three distinct human augmentation technology prototypes. In this section, we provide a short overview of the projects, game mechanics, and EMG functionalities of each prototype:

Jumping Augmentation: JumpiFlumpi is a VR game that combines physical activity with competitive gameplay. Players compete to achieve the highest score by jumping within the virtual environment. *Mechanics:* The core mechanic revolves around jumping. Players physically squat in real life to initiate jumps within the game. The duration of the squat determines the jump height. *EMG Functionality:* The game utilizes physical movement (squatting) as the primary input mechanism (see Figure 9.2a).

Object Manipulation: PortalGhost is a game that requires teamwork and puzzle-solving to escape a laboratory complex. One player assumes the role of a trapped human character within the VR environment, while the other acts as a ghostly entity controlled by a desktop PC. *Mechanics:* The core gameplay revolves around collaboration between the two players. The ghost can pass through walls and objects and create portals that share the ghost's vision with the VR player. This allows the VR player to perceive and interact with objects beyond their physical reach. *EMG Functionality:* EMG sensors worn on the biceps and forearm detect muscle contractions, allowing the player to control in-game telekinetic abilities. Contracting the biceps pulls objects closer while extending the arm pushes them away (see Figure 9.2b).

Super Strength: GuidingLight is a game where two players collaborate to cleanse a sacred temple of zombies. One player takes on the role of a warrior within the VR environment, while the other acts as a supporter in the real world, influencing the VR player's abilities using EMG. *Mechanics:* The VR player utilizes hand controllers to fight zombies directly, while the other player utilizes EMG on their arms to charge the VR player's abilities. *EMG Functionality:* EMG sensors worn on both arms detect muscle contractions, allowing the player to directly influence the VR player's abilities. Stronger muscle contractions translate to increased power for punches, longer slow-down durations, and greater teleportation range (see Figure 9.2c).

9.4 Ideation and Implementation of Virtual Augmentations

All participants described a typical group ideation process: they performed brainstorming sessions to collect quick ideas and then discussed and selected together. Coming up with initial ideas for virtual augmentations was generally regarded as easy (P3, P4). However, because they had little experience with virtual augmentation, they could not yet imagine how easily the ideas could be realized and what the experience would be like:

> *"[...] in our fantasy, I think it was easy to imagine what it [the virtual augmentation] should look like. I think it wasn't 100 percent clear at the start what we could get done in that time, that is fun. [...] I mean, that was the hard part, I would say."* - P4.

Notably, all participants describe different starting and focus points during their ideation. While P2 mentioned that their group started by brainstorming ideas for possible virtual augmentations very openly, P4 described that their group initially settled on a game mechanic and only proceeded to ideate and select augmentations afterward. In contrast, P3 mentioned that their group took the required input method as the basis for their ideation. When deciding on their virtual augmentations, the groups attempted to implement abilities that would go well beyond what humans can achieve in the real world, even by training. Ideally, the virtual augmentations would go even further and represent abilities that are impossible or too dangerous in the real world, even with considerable expenditure (e.g., traveling through portals or performing super jumps).

We found two layers in the groups' intentions when ideating and developing the virtual augmentations. First, multiple participants (P1, P4) explained that they aimed to make the augmented abilities as intuitive as possible and enable players to feel like the ability is really a part of their body. The virtual augmentations should be understandable, easy, and fun to use. Second, participants outlined the experience their groups intended to create for players. Key motives all participants mentioned were to create stimulating experiences that allow players to explore new abilities and feel empowered or unconstrained by the new virtual abilities. For instance, P1 and their group implemented an ability through which flexing the players' muscles would charge a strong punching ability and make the lights in the virtual environment flare up. This was meant to make the player feel powerful but also encourage exploration:

> *"I think it [the virtual augmentation] also was a motivation to just like try around because a different hand would make a different light and a different ability. And I think that also like motivated the player to just, like, play."* - P1

While the first layer can be considered the usability for virtual augmentation to make sure that the experience is smooth and working well, the second layer describes what should make the experience positive and valuable for the players. As this was the first time the participants engaged with virtual augmentation, it is unsurprising that the described intended experiences revolve around being new and surprising for players. In the future, we believe that other motives are possible as well.

All participants mentioned that the development of their games with a focus on virtual augmentation was not fundamentally different or more challenging than other development projects they had done before in Unity3D. However, one difference was that specifically implementing the augmentations felt new and the participants could not fall back on many best practices or examples to copy from other existing games or applications. While this may sound like a difficulty, P4 explained how it was a welcome challenge to expand their horizon as a developer:

> *"[. . .] let's say you develop a game, oftentimes you'll start out by just pretty much googling how similar games were developed [. . .]. But with this setup, there was nothing, right? This is something new or something that hasn't been done that much [. . .]. So you have to think it all up by yourself. [. . .] I was really excited to go there and work on it because it was something new, something we ourselves could define how it worked and how we did it."* - P4

9.5 Virtual and Human Augmentation

All participants described both working with and experiencing virtual augmentation as positive. Some participants (P3, P4) expressed that wielding superpowers, magic, or other impossible abilities is fun and stimulates by introducing novel experiences that offer a departure from mundane routines. Further, multiple participants felt that experiencing different and impossible abilities in VR could generate interest and attention (P1, P2) and motivate people to use VR.

Interestingly, P3 explained that they felt that well-implemented virtual augmentations could help make a VR experience more immersive and realistic:

> *"I think there are advantages of it because it just feels more realistic. Also the way you can interact [. . .] feels more natural [. . .]. So I think it's more, yeah, more immersive."* - P3

Moreover, abilities in the virtual world are not fixed and can be tailored to users with different requirements. P2 mentioned how they see the potential of this design approach when designing for users with disabilities:

> *"It is also one method to let [users with] disabilities play the game."* - P2

In some of the games, the groups specifically focused on creating an experience in which players with different (virtual) abilities cooperate to reach the goal.

Participants also mentioned the drawbacks and challenges of virtual augmentation. Virtual augmentation utilizing body movements can quickly lead to physical exhaustion and restrict who can use it and for how long (P1, P4). Participants (P3, P4) reported that especially virtual augmentations that affect movement or other physical abilities are often at risk of creating simulator sickness. Further, abilities such as increased force are especially difficult to implement, as you typically do not feel resistance in VR. Consequently, audio-visual feedback must be fine-tuned to create satisfying and engaging experiences (P1).

In the interviews, the participants reflected on the relationship between human augmentation in the real world and virtual augmentations. Looking at the virtual augmentations from the games, most are fully or almost impossible to transfer into real life and would cause serious safety concerns (P2, P4). Even when it is possible to transfer virtual augmentations to the real world, the technological effort required and the danger associated with abilities like jumping seem too large to make it worth it (P4). In VR, more or different abilities are possible (e.g., teleportation) (P2, P3) and the virtual environment is a safe space to explore enhanced abilities in the absence of many security concerns. However, the virtual space also comes with some drawbacks, as the lack of physical feedback and resistance make it harder to implement physical augmentations believably, and the augmented movement quickly causes simulator sickness (P3, P4) or requires complicated technological setups (e.g., for natural movement).

9.6 Discussion and Future Work

The results of the interviews offer a detailed account of the experience of ideating and implementing virtual augmentations in VR. While on a surface level, generating initial ideas for virtual augmentations was deemed relatively straightforward and development is not fundamentally different or more demanding than in other projects, specific challenges emerged. As the design approach of focusing on virtual augmentations is not widely established, typical practices such as looking at published examples for inspiration, common especially in VR development [435] are not possible. Currently, little formal knowledge exists about what makes for a good user experience when it comes to virtual augmentations. The participants outline starting points from their experience in the exercise, revealing the need to first ensure that augmenting interactions work well, are easily understood, and feel like they are part of the user (as in usability). On top of that, to create a positive experience, e.g., be stimulating and encourage users to explore the new ability. However, with the experiences mentioned by the participants, there is a risk that virtual augmentations are merely positive and fun because they are new. To create lasting positive experiences, other goals should be explored, such as creating virtual abilities that are satisfying to use in themselves (i.e., induce a state of flow) or, especially for non-game applications, offering users new perspectives and insights through abilities that are only possible in VR, relevant to a goal they pursue.

In the practical exercise, the groups were free to decide what kind of augmentation technologies they wanted to develop. Intuitively, the groups focused primarily on augmentations that affect physical or movement abilities. While especially the term "superpowers" initially brings to mind such augmentations, they are difficult to realize with current VR technology. While VR allows for physical and spatial interactions, new forms of movement run the risk of triggering simulator sickness or requiring elaborate equipment such as treadmills. Additionally, an augmented force can feel empty, as the primary feedback players receive from current VR systems is visual - lifting a large, heavy object in VR, for example, feels the same as lifting a small, light object. Traditionally, VR is more suited to audio-visual illusions. In the future, the ideas and development of virtual augmentations could be more focused on enhancing perception rather than action. Another approach could be to extend the VR system in which augmentation is experienced through haptic feedback.

Overall, virtual augmentation was positively perceived, offering stimulating and fun experiences and immersive interactions. However, participants recognized challenges such as physical exhaustion and simulator sickness, which could exclude many potential users. At the same time, designing and playing with the virtual abilities of users bears the potential to include users with disabilities, e.g., by transforming senses in the virtual environment or tailoring virtual abilities to specific users.

While we acknowledge the inherently small sample size of this study, it is important to emphasize that our focus was on exploratory qualitative work. Our primary interest lies in the existence of phenomena rather than their incidence. Therefore the subjective insights provided by our participants serve as a valuable starting point, offering rich insights that could be further developed in future research. Additionally, it is important to note that our primary focus was on the prototyping process itself, and as such, we did not conduct a user evaluation of the applications. Future studies could build upon our findings by incorporating larger sample sizes and user-testing phases to further validate and expand on the insights generated here.

We believe that virtual augmentation can be utilized in different ways. First, it shows the potential to contribute to a positive user experience in VR when implemented well. Beyond this study, different authors have elaborated on its potential when designing VR applications (e.g., [432]). Further, an experimental vignette study suggests that augmentation experience positively contributes to good user experiences in VR [433]. A first design tool attempts to guide designers leveraging this potential [436]. Other than that, virtual augmentation can also be used to prototype and try out human augmentations intended for the real world.

10

ENHANCING SENSORY EXPERIENCES IN VIRTUAL ENVIRONMENTS

This chapter is based on the following publications:

An Examination of Ultrasound Mid-air Haptics for Enhanced Material and Temperature Perception in Virtual Environments - Steeven Villa, Yannick Weiss, Niklas Hirsch, Alexander Wiethoff *In: Proceedings of the ACM on Human-Computer Interaction, Volume 8, Issue MHCI. Association for Computing Machinery.*

Extended Mid-air Ultrasound Haptics for Virtual Reality - Steeven Villa, Sven Mayer, Jess Hartcher-O'Brien, Albrecht Schmidt, Tonja-Katrin Machulla - *In: Proceedings of the ACM on Human-Computer Interaction, Volume 6, Issue ISS. Association for Computing Machinery.*

Touch It Like It's Hot: A Thermal Feedback Enabled Encountered-Type Haptic Display for Virtual Reality - Steeven Villa, Kenji Ishihara, Moritz Ziarko, Sebastian Günther, Florian Müller - *In: 23rd IEEE International Symposium on Mixed and Augmented Reality (ISMAR 24). IEEE*

Building on the previous concepts, this chapter explores methods to enhance sensory experiences in virtual reality (VR), focusing on improving haptic rendering. While the primary goal is to advance human augmentation experiences, the techniques discussed have broader applications for visuo-haptic interactions in VR. This chapter examines the use of mid-air ultrasonic haptics to convey object properties, strategies to extend the range of ultrasonic haptic feedback, and approaches for rendering temperature sensations using encountered-type haptic devices.

Using Ultrasonic Mid-air Visuo-haptic presentation to render Objects Properties in VR

Ultrasound Mid-air Haptics (UMH) has been established as a viable technology rendering a diverse array of haptic sensations. However, the question of how congruent UMH feedback is in representing various object properties remains an open area of investigation. In particular, objects in different states of matter (solid, liquid, gas) present unique haptic characteristics.

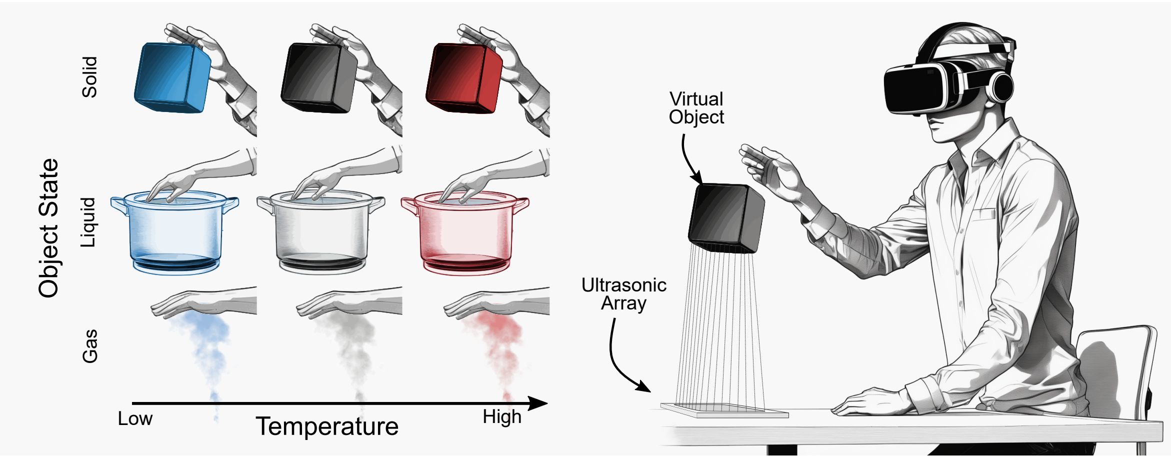


Figure 10.1: In this section we explore the efficacy of Ultrasonic Mid-air Haptics (UMH) in enhancing the perceived congruency of virtual objects with differing material states and color-temperature associations. Specifically, participants interacted with virtual objects rendered in three distinct physical states—Solid, Liquid, and Gas—while also exposed to three hue-temperature associations: Blue (cool), White (neutral), and Red (warm). The study compared these experiences in two conditions: with the presence of Ultrasonic Mid-air Haptic feedback and with Visual Only (no haptic feedback).

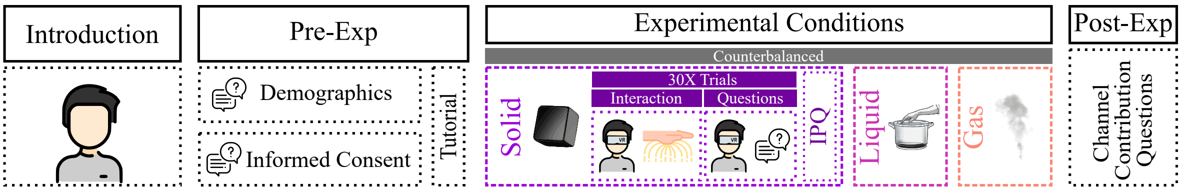


Figure 10.2: Experimental Design Overview: Our study involved three distinct blocks, each aligned with a specific material state. We used the Latin Square design to counterbalance presentation order, mitigating potential order effects. Within each block, we randomized combinations of temperature color and haptics, each repeated five times. Each block included 30 trials, leading to a total of 90 trials.

To explore the capabilities of UMH under these varying conditions, we selected these three object states for examination.

Further, to augment the potential of UMH in simulating temperature, we incorporated the Hue-Temperature Association Hypothesis. According to this hypothesis [437], cooler colors are associated with a colder temperature perception, while warmer colors evoke sensations of heat. Previous studies have explored this association without the addition of tactile cues [438]. UMH presents an intriguing opportunity in this context: it offers a touchless method to stimulate palm mechanoreceptors without necessarily engaging thermoreceptors. This additional layer of tactile stimulation may potentially strengthen the existing Hue-Temperature association.

Given this context, we propose two hypotheses:

H1: The congruency of haptic feedback rendered through UMH will vary significantly across different object states.

H2: The addition of tactile cues from UMH will strengthen the pre-existing hue-Ttemperature association, leading to a more pronounced or significant shift in perceived temperature based on color cues

10.1 Experimental Design

We conducted a laboratory experiment using a within-subject design to explore how the combination of UMH and HIs can enhance the representation of material properties in virtual reality (VR). Specifically, our investigation explored the effects of ultrasonic rendering on three distinct states of matter (gas, liquid, and solid) and how this rendering impacts the overall perception of material consistency. Additionally, we examined how the combination of ultrasonic rendering influences the perception of object temperatures within the VR environment. Accordingly, we manipulated three variables: MATERIAL STATE, TEMPERATURE COLOR, and HAPTICS. MATERIAL STATE had three levels: Gas, Liquid, and Solid, TEMPERATURE COLOR also had three levels: Blue, White, and Red. Finally, HAPTICS had two levels: Active and Inactive. We organized our experiments into three distinct blocks, with each block corresponding to a specific MATERIAL STATE. To minimize any potential order effects, we used a Latin Square design to counterbalance the order of presentation for these blocks. Within each block, we randomized the combinations of TEMPERATURE COLOR and HAPTICS, ensuring each combination was repeated five times. In total, each block consisted of 30 trials, resulting in a total of 90 trials. Participants were able to see their hands during all the experiment.

10.2 Participants

We recruited a total of thirty participants (N = 30) through the university's mailing lists from which fifteen identified as female, fifteen as male, and no participants indicated a self-specified gender. Our sample had an average age of twenty-five years (M = 25.47, SD = 8.69). Participants reported a low familiarity with haptic devices in general (M = 1.33, SD = 0.60), and a low to medium level of familiarity with VR (M = 2.53, SD = 1.00). Participants were compensated 6 euros/30 min for their involvement. The study was approved by an ethics committee (Grant Nr. <removed>).

10.3 Apparatus

We used a Meta Quest 2 headset [1] along with an Ultraleap STRATOS Explore device [2] to provide Visuo-Haptic feedback. The VR environment was developed using Unity3d and executed on an ACER Predator Helios computer [3]. Additionally, we incorporated a 64-electrode R-net EEG system [4] from which we used 32 electrodes. To ensure synchronization between the VR environment and the EEG recording device, we employed the Labstreaming Layer[5].

10.4 Stimuli

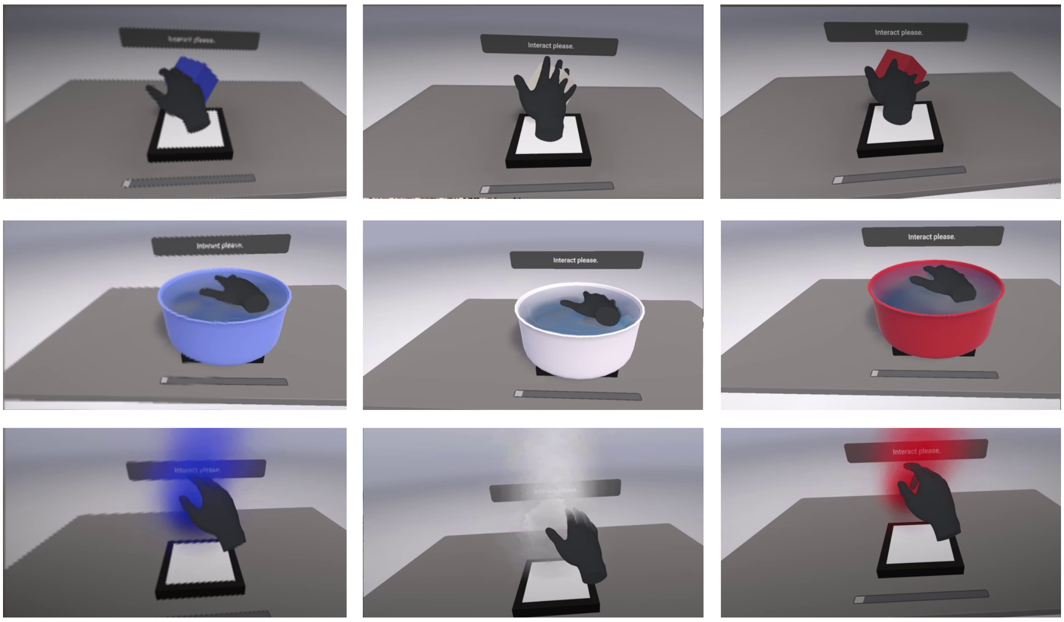


Figure 10.3: Example stimulus combinations in VR: The solid object was represented using a cube, the Liquid using a water shader in a colored pot, and the gas using a particle system

Participants were given instructions to interact with a virtual object for a fixed duration of 5 seconds. Following this interaction period, they were asked to provide responses to two questions and continue with the next object. Participants were allowed to touch the object multiple times during this phase, the object's presentation was a combination of the following factors:

[1] https://www.meta.com/de/en/quest/products/quest-2/

[2] https://www.ultraleap.com/company/news/press-release/stratos-platform/

[3] https://www.acer.com/de-de/predator/laptops/helios/helios-300

[4] https://www.brainproducts.com/solutions/r-net/

[5] https://github.com/sccn/labstreaminglayer

10.4.1 Hue-Temperature Association

The Hue-Temperature Hypothesis posits that specific colors are subjectively linked to temperatures; i.e., blue is often perceived as colder than red [437]. Evidence supporting this can be found in physiological research [439] and human-computer interaction studies [438]. We applied it to VR objects to simulate temperature; we rendered three colors: blue (`#2F48C5`), neutral (white) (`#D1D1D1`), and red (`#AB3737`). We included white as a control point given its lower color temperature association [439]. Based on beta-tester feedback during the design process, in the Liquid condition, we colored the container itself, given that fluids often do not exhibit a color change in response to increased temperature (Figure 10.3).

10.4.2 Material State

We rendered the material state in three forms: a solid cube to symbolize the solid state, a pot incorporating a fluid shader to depict the liquid state, and a steam particle system to represent the gas state (Figure 10.3).

10.4.3 Presentation

We provided UMH feedback using the ultraleap STRATOS Explore phased array. We employed spatiotemporal modulation a method detailed in previous research [440]. We calculated collision points where fingers interacted with virtual objects. Subsequently, we created a curve connecting these points and moved the focal point rapidly along this curve trajectory [441].

10.5 Measures

We gathered data from multiple sources; two per-trial questions, a questionnaire after each block of trials, and, brain signal data from an EEG device [6]. To elaborate further, we measured:

10.5.1 Perceived Temperature

We assessed participants' subjective temperature perception using the Bedford Thermal Scale, specifically the extended University of California Berkeley (UCB) model with nine distinct levels. This instrument employed the question, "Please rate your thermal sensation," on a scale that spanned from "Very Cold" to "Very Hot." [442].

[6] `https://www.brainproducts.com/solutions/r-net/`

10.5.2 Perceived Congruency

We measured congruency by asking the following question: *How much did your experiences in the virtual environment seem consistent with your real-world experiences?* on a 9-point Likert scale once per trial ranging from *Very Inconsistent*, to *Very Consistent* [443].

10.5.3 Presence

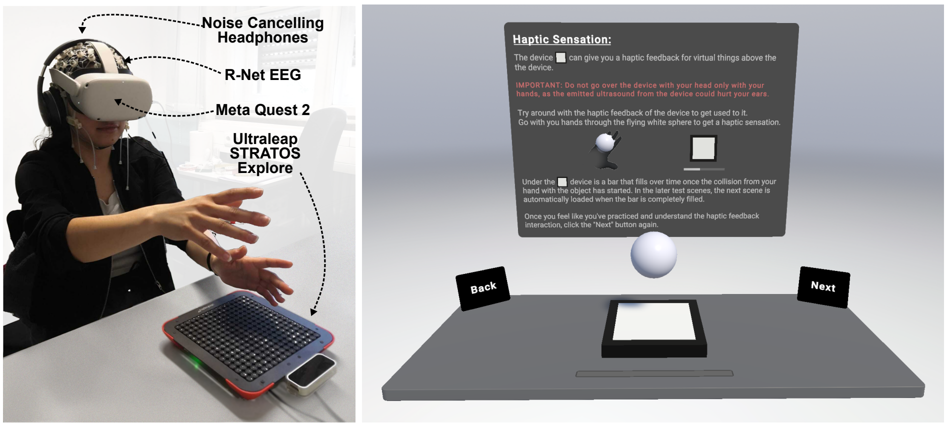
We used the Igroup Presence Questionnaire (IPQ), the IPQ comprises four subscales: Sense of Being There, Spatial Presence, Involvement, and Experienced Realism. We administered the IPQ to participants after the completion of each block of trials.

10.5.4 Perceived Channel Contribution to the Overall Experience

We evaluated participants' perceptions of the relative contributions of the sensory channels (visual and haptic) to both the perceived temperature and congruency of the object using four questions at the end of the experiment: **Q1:** *"My perception of temperature was mainly influenced by the tactile feedback."*, **Q2:** *"My perception of temperature was mainly influenced by the visual feedback."*, **Q3:** *"My perception of the material state was mainly influenced by the tactile feedback."*, and **Q4:** *"My perception of the material state was mainly influenced by the visual feedback."*

10.5.5 Processed Congruency

We quantified congruency through a metric called Mismatch Negativity, derived from the EEG signal. This metric was introduced by Gehrke et al. [444] as an alternative to subjective evaluations of haptic inconsistencies [7]. When users encounter a visuo-haptic mismatch, the negativity of the resulting Event-Related Potential (ERP) signals at the forehead (FCz location) decreases (the signal becomes more positive). This means that a more negative value represents a more matching stimulation. Throughout the experiment, we continuously monitored EEG signals and identified the first instance of contact between the participant's hand and the object in each trial to assess the mismatch negativity. In our study, we explored this recently introduced association to study congruency from the neurophysiological perspective. Yet, we want to emphasize that his metric has not been extensively validated by Gehrke et al. [444]; therefore, it is for informative purposes only. For this reason, we do not center our analysis on the neurophysiological metrics of congruency but on the self-reported metrics.



(a) User exploring a virtual object in VR

(b) Snapshot of one of the tutorial screens pre-experiment

Figure 10.4: Experimental setup and tutorial screen: Our setup included Noise Canceling headphones to mask environment sounds and the Ultrasonic Array noise, We used a combination of Meta Quest 2 for Visual rendering and Ultraleap Stratos for Haptic rendering, additionally we included an R-Net EEG device to measure participant's brain responses. Please note that the spherical shape was only used during the tutorial phase.

10.6 Procedure

Participants then received an introduction to the VR hardware and system functionality and were asked to provide informed consent. Next, an experimenter securely attached the EEG headset, ensuring that electrode impedance remained below 50k Ω. To minimize external disturbances, participants were instructed to wear noise-canceling headphones, which masked environmental sounds and any noise generated by the active haptic array. Subsequently, participants were assisted in putting on the VR headset, and then they had to complete a tutorial to acquaint themselves with the experiment's flow, including interactions and questions (see Figure 10.4 for an example screen). Following, the formal experiment commenced, with participants being assigned a specific order of experimental conditions. Over the course of the study, participants were required to complete a total of 90 interactions. Following each interaction, they answered questions regarding perceived temperature and congruency. After completing each block of trials, participants also filled out the IPQ questionnaire. Notably, all questionnaires were administered within the VR environment to minimize disruptions to participant immersion. The entire study, including the setup of EEG equipment and questionnaire completion, lasted one hour.

[7] It is important to emphasize that this metric does not have formal validation. Therefore, it should not be regarded as a ground-truth measurement.

10.7 Results

10.7.1 Linear Mixed Models

We employed Linear Mixed Models (LMMs) for our data analysis, which are statistical models designed to manage correlated data. LMMs encompass both fixed effects (predictors) and random effects (room-temperature, and humidity in our case), making them suitable for analyzing intricate datasets with nested or repeated measurements, as in our case. we accounted for the non-independence resulting from multiple responses from the same participant, as recommended by Winter [445]. We followed the model selection process outlined by Zuur et al. [446] using a top-down strategy. Inspection of residual plots did not reveal deviations from homoscedasticity or normality. P-values were obtained through likelihood ratio tests, comparing the full model with the predictor in question to a simplified model without the predictor. Additionally, we utilized the *lmerTest* package [447], which approximates degrees of freedom for t- and F-tests using the Satterthwaite method and provides p-values for the fixed effects.

10.7.2 EEG Analysis

For EEG data analysis, we utilized the Python MNE library. We applied a high-pass filter at 1 Hz and a low-pass filter at 15 Hz, following the methods described in previous studies [191, 192]. The data was re-referenced to the average of all channels, and a notch filter was employed to eliminate the 50 Hz powerline noise. Subsequently, we segmented the epochs into time blocks spanning from -0.3 ms to 0.7 ms, with 0.0 ms indicating the stimulus onset. The period between -0.3 ms and 0.0 ms served as a baseline for the measured stimulus signal. To identify and discard epochs likely to contain noise, we employed the Autoreject Library [193].

10.7.3 Control Variables

We measured the temperature and humidity of the experimental room at the start and end of each experiment. We used this information as control variables in all the models presented in this section. Additionally, we report the mean values and variations during the experiment here. The average room temperature was 26 degrees Celsius ($M = 26.00$, $SD = 1.01$), with an average variation (end temperature minus start temperature) of 0.79 degrees Celsius ($M = 0.79$, $SD = 0.27$). The average relative humidity of the room was 42 ($M = 42.48$, SD = 6.16), which is within the recommended values for room temperature ($Min = 30$, $Max = 60$). The average variation in relative humidity was 1.64 ($M = 1.64$, $SD = 1.97$).

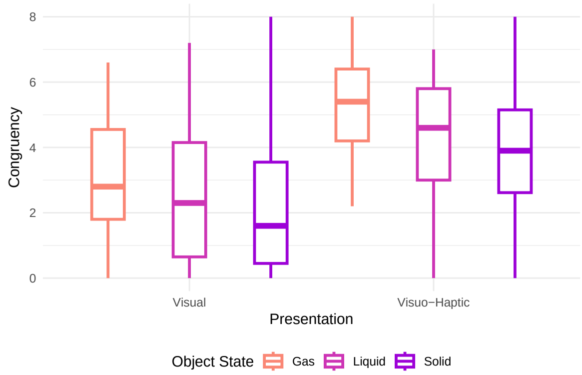


Figure 10.5: Perceived Congruency per Object State and Presentation, Overall, Visuo-Haptic objects were perceived as more congruent than Visual objects. Further, Gas was generally perceived as more congruent than Liquid and Solid, the latter being the one perceived as least congruent

10.7.4 Object Congruency Perception

In order to determine the contribution of PRESENTATION to the explanatory power of the CONGRUENCY model, we compared a *reduced* and a full model. The reduced model included only MATERIAL STATE as a fixed effect, along with random effects for Participant and HUE-TEMPERATURE. The full model additionally incorporated PRESENTATION as a fixed effect.

A likelihood ratio test indicated that the inclusion of PRESENTATION led to a statistically significant improvement in model fit ($\chi^2(1) = 184.61, p < 0.001$). This was corroborated in the goodness-of-fit measures by lower values of both the Akaike Information Criterion (AIC_f = 2051.7, full model vs. AIC_r = 2234.3, reduced model) and the Bayesian Information Criterion (BIC_f = 2081.7 full model vs. BIC_r = 2260.0, reduced model). Indicating PRESENTATION (presence or absence of haptics) as a significant predictor for accurate data representation.

Considering the influence of MATERIAL STATE on CONGRUENCY ratings under two PRESENTATION modalities: VISUO-HAPTIC and VISUAL.

We found that for VISUO-HAPTIC PRESENTATION, MATERIAL STATE was a significant predictor of CONGRUENCY ratings ($\chi^2(2) = 61.75, p < 0.001$). The goodness-of-fit measures further supported this: $AIC_f = 944.59$ was lower than $AIC_r = 1002.34$, and $BIC_f = 966.18$ was lower than $BIC_r = 1016.74$. Within this modality, we found that Solid had the lowest mean CONGRUENCY rating ($M = 3.85, SD = 1.86$), followed by Liquid ($M = 4.35, SD = 1.78$), and Gas had the highest ($M = 5.34, SD = 1.47$).

Similarly, in the VISUAL PRESENTATION, MATERIAL STATE was also a significant predictor for CONGRUENCY ($\chi^2(2) = 20.76, p < 0.001$). Here, the full model had an $AIC_f = 981.39$ and $BIC_f = 1003.19$, both lower than the reduced model's $AIC_r = 998.15$ and $BIC_r = 1012.5$. The mean CONGRUENCY rating for Solid was $M = 2.18, SD = 2.06$, for Liquid it was $M = 2.59, SD = 1.98$, and for Gas it was $M = 3.04, SD = 1.92$.

In the VISUO-HAPTIC PRESENTATION, the average scores for each MATERIAL STATE were higher compared to Visual Presentation (see Figure 10.5). These results lend empirical support to **H1**, indicating that objects with lower kinaesthetic complexity are perceived more congruently in a virtual reality environment compared to objects with high kinaesthetic complexity. Furthermore, our data reveals that the incorporation of tactile cues—specifically, UMH—significantly enhances the overall perception of object congruency.

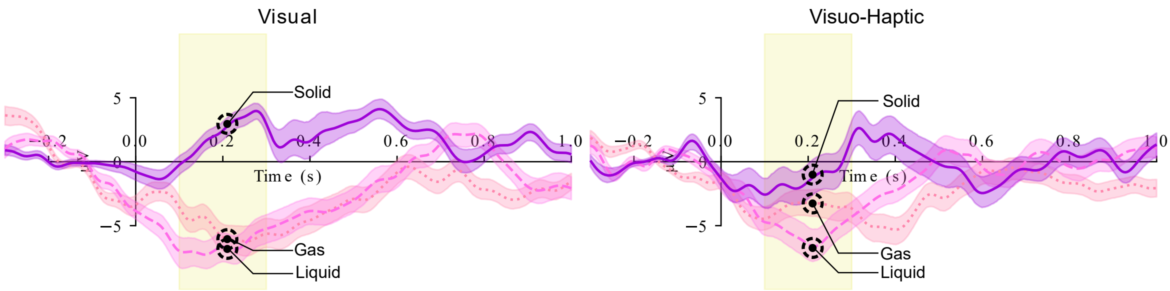


Figure 10.6: Mismatch Negativity (Prediction Errors) in ERP response , Lower values are associated with more congruent stimulus, A higher value is associated with a more unrealistic VR interaction [444].

ERP Analysis and Mismatch-Negativity

We observed differences in Event-related Potentials (ERP) at the FCz electrode based on Material State and Presentation. Figure Figure 10.6 illustrates the negativity response elicited by the exploration stimulus. According to the metric proposed by Gehrke et al. [444], higher negativity is indicative of a more congruent virtual reality experience. [8]

Consistent with the Congruency ratings we reported earlier, the Solid State had the lowest levels of congruency. On the other hand, the Liquid and Gas states performed better in terms of congruency across both Presentation modalities. Interestingly, Figure 10.6 shows that the negativity for the Solid State is heightened during Visuo-Haptic Presentation. This suggests that participants found the Solid State to be more congruent when experienced in the Visuo-Haptic mode also at a physiological level. However, additional studies focusing on Mismatch Negativity are needed to confirm these findings.

10.7.5 Object Temperature Perception

To examine H2, we employed a three-step approach. First, we developed a model to investigate the perceived temperature in relation to the Hue-Temperature Association, aiming to validate if participants associated the object color with the object temperatures. Second, we assessed whether the Presentation modality influenced this color-temperature association. Finally, we conducted a level-specific analysis to determine if the inclusion of haptic cues amplified the Hue-Temperature Association effect

[8] It is important to emphasize that this metric does not have formal validation. Therefore, it should not be regarded as a ground-truth measurement.

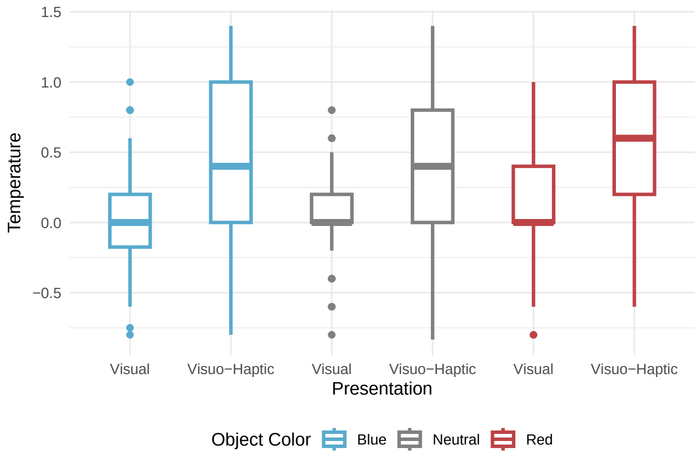


Figure 10.7: Perceived Temperature, Haptic feedback consistently shifted the perceived temperature across all Hue-Temperature associations.

To assess the influence of HUE-TEMPERATURE on the Perceived TEMPERATURE of virtual objects, we employed both a reduced and a full model. The reduced model incorporated MATERIAL STATE and PRESENTATION as fixed effects, with Participant as a random effect. In contrast, the full model extended this framework by including HUE-TEMPERATURE as an additional fixed effect.

A likelihood ratio test revealed HUE-TEMPERATURE as an statistically significant improvement in model fit ($\chi^2(2) = 11.945, p < 0.005$). The model had a goodness-of-fit (AIC_f = 609.28 vs. AIC_r = 617.23, BIC_f = 643.19 vs. BIC_r = 642.66). Indicating HUE-TEMPERATURE as a significant predictor for accurate representation of the data. This provides evidence that participants associated the temperature of the presented object with the color of the object. Specifically, blue was perceived as cooler (M = 0.22, SD = 0.53) compared to red (M = 0.34, SD = 0.50). However, the neutral color was rated even colder than blue (M = 0.19, SD = 0.45). Suggesting a consistent heat association for Red, but a less clear one for Blue and Neutral.

Next, to explore the influence of PRESENTATION on the Perceived TEMPERATURE of virtual objects, the reduced model incorporated MATERIAL STATE and HUE-TEMPERATURE as fixed effects, with Participant as a random effect. In contrast, the full model extended this framework by also including PRESENTATION as an additional fixed effect (Similar to the full model for HUE-TEMPERATURE association).

A likelihood ratio test indicated that the inclusion of PRESENTATION led to a statistically significant improvement in model fit ($\chi^2(2) = 96.17, p < 0.001$). The model had a goodness-of-fit (AIC_f = 609.28 vs. AIC_r = 703.45, BIC_f = 643.19 vs. BIC_r = 733.12). Therefore, PRESENTATION was identified as a significant predictor for data representation, suggesting a substantial influence on the perceived TEMPERATURE of the virtual object. Specifically, objects in the VISUO-HAPTIC PRESENTATION were rated as warmer (M = 0.44, SD = 0.56) compared to those in the VISUAL PRESENTATION (M = 0.07, SD = 0.34). Importantly, this difference in perceived

TEMPERATURE was more pronounced than the variance attributed to the HUE-TEMPERATURE association.

At the level of individual colors, we observed that objects were consistently rated as warmer in the VISUO-HAPTIC PRESENTATION compared to the VISUAL PRESENTATION. Although both PRESENTATION and HUE-TEMPERATURE Association serve as significant predictors for perceived temperature, their impact diverges from our initial hypothesis. Specifically, VISUO-HAPTIC PRESENTATION does not amplify the Hue-TEMPERATURE association; rather, it uniformly elevates the perceived TEMPERATURE (see Figure 10.7). This trend remains consistent across all combinations of Material State and Hue-Temperature Association, as further detailed in Table 10.1.

Table 10.1: Temperature Associations across all Object States, Hue-Temperature Associations, and Presentation

	Solid - M (SD)		$Liquid - M(SD)$		$Gas - M(SD)$	
	Visual	Visuo-Haptic	Visual	Visuo-Haptic	Visual	Visuo-Haptic
Blue	0.02 (0.30)	0.23 (0.56)	0.05 (0.36)	0.37 (0.61)	0.10 (0.40)	0.56 (0.65)
Neutral	0.04 (0.28)	0.34 (0.51)	-0.04 (0.31)	0.36 (0.55)	0.07 (0.27)	0.43 (0.50)
Red	0.14 (0.41)	0.49 (0.46)	0.10 (0.36)	0.59 (0.64)	0.17 (0.32)	0.60 (0.48)

10.7.6 IPQ

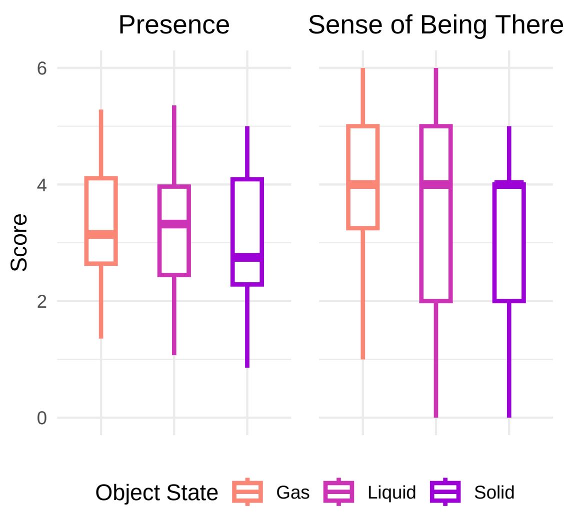


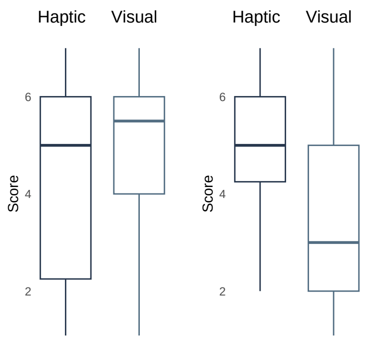
Figure 10.8: IPQ General presence and Sense of Being There subscale: We found significant differences in the Sense of Being There subscale for Gas and Solid states, yet, no differences were observed in the remaining subscales

We measured the overall impact on presence at the end of each block, Figure 10.8 presents an overview of the data for the four sub-scales (*Sense Of Being There*, *Spatial Presence*, *Involvement*, *Experienced Realism*) as well as the composite *Presence* score. As the data did not follow a normal distribution, we employed the *Friedman test* in this case.

Only the *Sense Of Being There* sub-scale exhibited a significant difference across object states, as indicated by a $\chi^2(2) = 11.877, p < 0.05$. Post-hoc test showed significant differences between the GAS and SOLID states (p < 0.05). No significant variations were observed in the general *Presence* score across different object states.

10.7.7 Perceived Channel Contribution to The Overall Experience

As the data did not follow a normal distribution, we applied the *Wilcoxon signed-rank test* to assess differences between paired questions: Q1 vs. Q2 and Q3 vs. Q4. Statistical analysis indicated a significant difference between questions Q1 and Q2 (V = 311.5; $p < 0.05$) and between questions Q2 and Q4 (V = 35.5; $p < 0.05$), see Figure 10.9.



(a) Object State (b) Temperature

Figure 10.9: Reported contribution to Object State and Temperature perception per channel: Participants rated that the visual channel had a bigger contribution to their perception of the material state, while haptics had a bigger impact on the perceived temperature than the visual channel.

10.8 Discussion

Our analysis showed that Presentation, Hue-Temperature Association, and the target Object State serve as significant predictors for both Perceived Object Congruency and Temperature. Notably, the observed effects of Hue-Temperature and Presentation diverged from our initial hypotheses. This section elaborates on these findings.

10.8.1 Effects of Visuo-Haptic Presentation on Object Congruency Perception

Simulating the haptic characteristics of virtual objects in VR remains a challenge. Yet, our findings indicate that Ultrasound Haptics can substantially improve the perceived congruency across all evaluated object states. While simulating solid objects poses particular difficulties, we observed that the mere presence of tactile stimuli via Ultrasound Haptics significantly enhances users' perception of congruency. We also found that Gas and Liquid VR objects are already perceived as more congruent in the Visual Presentation modality; however, even the lower-rated Visuo-Haptic Object State already outperforms the best Visual Object State. This supports the importance of considering multisensory experiences in VR and not relying only on visual information as a feedback channel.

10.8.2 Effects of Visuo-Haptic Presentation on Physiological Responses to Object Consistency

Our physiological measures were in line with self-reported experience metrics. Specifically, in the event-related potentials (ERP) data, Solid objects demonstrated a larger shift in negativity compared to Gas and Liquid objects, which remained relatively consistent. In a visual-only setup, the negativity associated with Solid objects was low, however, this value increased significantly when tactile stimuli were introduced through visuo-haptic interaction. It is important to approach these physiological findings with caution, as the ERP correlate for experience congruency is a recent addition to the literature and lacks extensive validation. Nevertheless, our results are consistent with the initial findings reported in the original study.

10.8.3 Effects of Visuo-Haptic Presentation on Thermal Perception

Our data revealed that the incorporation of tactile cues did not amplify the Hue-Temperature association, as initially hypothesized. Instead, it led to an overall shift in the perceived temperature of the objects. This effect was consistently observed across all states of matter and color conditions. This suggests that UMH may influence temperature perception, even though there is currently no empirical evidence to show that it specifically stimulates thermoreceptors.

An additional observation was that, within the context of visuo-haptic presentation, objects in the Gas state received higher warmth ratings compared to those in the Liquid and Solid states. This implies that the perceived temperature of an object could be influenced by its state. Literature has reported an effect of specific visualizations impacting temperature perception, including objects in gas states such as rain clouds [448] and steam [449]. However, due to these explorations targeting very specific objects, the concrete effect of an object's state on perceived temperature requires further investigation.

10.8.4 Per channel contribution to Temperature and Congruency

Our findings reveal that participants placed greater sensory importance on UMH when evaluating the temperature of virtual objects. In contrast, they relied more on visual cues for determining the objects' state (e.g., Solid, Liquid, or Gas). Although our modeled data supports these observations, it's noteworthy that both UMH and visual cues significantly influenced the perception of object state congruency.

10.8.5 Implications for Design

In this section, we present new insights into how UMH rendering affects how users perceive the congruency and temperature of objects. These insights offer practical guidance that haptic designers and researchers can apply. Here's a summary of our key takeaways:

Enhanced Congruency with Visuo-Haptic Presentation:

Our results indicate that integrating UMH in visuo-haptic presentations improves the perceived congruency of virtual objects.

Optimal Rendering Scenarios For Ultrasound Mid-air Feedback:

The efficacy of this haptic feedback is more pronounced for objects in gaseous or liquid states. This suggests that designers should tailor the choice of haptic feedback based on the specific type of object being rendered. Alternatively, they might opt for attributing gaseous or liquid properties to virtual objects and interfaces to maximize the congruency effect.

Shift in Perceived Temperature:

Our study strongly indicates that ultrasound haptic feedback alters the perceived temperature of virtual objects. Designers should take this into account, especially in scenarios where accurate temperature perception is crucial.

Design Opportunities with Temperature Shift:

The observed shift in perceived temperature also presents an opportunity for designers. Ultrasound haptics can be utilized to intentionally modulate the perceived temperature of virtual objects, thereby adding a new dimension to object rendering without requiring additional hardware.

Implications for VR prototyping

Utilizing UMH enhances the presence and congruency of VR design explorations, particularly in the early stages. This technology allows for the simulation of diverse object properties, such as soft/warm or hard/cold surfaces, within the same virtual model. Such haptic feedback can be coupled with Hue-Temperature associations to offer a more comprehensive sensory experience. This approach has not yet been incorporated into VR industrial design iterations and has the potential to significantly influence design decisions and project directions:

- Seamless Sensory Integration: Facilitates haptic explorations in VR without requiring user-worn devices, creating a holistic experience.
- Enhanced Materiality: Enables industrial designers to explore material affordances in VR, from texture to temperature, broadening the design palette.

- Perceptual Consistency: The introduction of UMH elevates the overall congruency of virtual objects, even outperforming visual-only simulations.
- Tailored Haptic Feedback: Our findings indicate optimal use cases for different object states, allowing for more targeted design choices in rendering gas, liquid, or solid objects.

Extended Mid-air Ultrasound Haptics for Virtual Reality

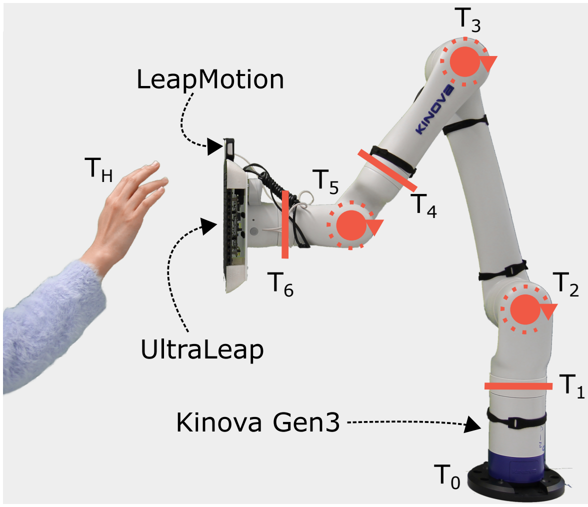


Figure 10.10: The setup comprises an ultrasound array (Ultraleap), a six-degree-of-freedom Kinova ®cobot, and a VR laptop (not visible in the picture). The final position of the haptic array is given by the transformations (T) in every joint of the robot. The transform (3d Position information) of the hand T_H depends on the cameras of the LeapMotion attached to the Ultraleap.

Real-world interactions are bare-hand and involve large volumes of interaction; we do not restrict ourselves to touching objects within a bounding box in our daily lives. Unfortunately, state-of-the-art haptic experiences tend to break these conditions in one way or another. For example, wearable haptics requires users to attach devices to their hands or bodies to exert a force or induce a vibrotactile stimulus, cf. Fang et al. [450]; force feedback devices require users to be in contact with the end-effector of a robot. In contrast, mid-air haptics offers bare-hand interaction but is constrained in workspace. We propose to render ultrasound mid-air haptics in motion. The proposed setup's dynamic nature allows stimulation from different angles, allowing haptic experience designers to explore a broader range of possibilities.

10.9 Setup

Figure 10.10 illustrates the system setup; we used a six-degree-of-freedom serial robot manufactured by Kinova robotics (Boisbriand, Canada) for driving a 256 ultrasonic transducer array (Stratos Explore) by Ultraleap (Bristol, United Kingdom) that features a LeapMotion tracker for hand tracking. To attach the haptic array to the robot, we 3D printed a lightweight interface that fits both the casing of the ultrasonic display and the robot's end-effector interface. The robot is integrated with the graphics engine Unity3D using the toolbox reported in [Exp7].

10.10 Tracking

The hand position guides the robot's end-effector. We divided the tracking into two levels. 1. short-range: guided by LeapMotion and 2. long-range, guided by a vibe tracker (but could be replaced by any tracking device).

For the short-range tracking, we converted the hand position from the LeapMotion Coordinate System (LCS) to the World Coordinate System (WCS) using the forward kinematics of the robot, which can be calculated using the transformations for every joint of the robot. Equation 10.1 describes the expression for moving from LCS to WCS, where the last transformation corresponds to the rotations and translations detected by LeapMotion and T_0 to T_6 can be found in Kinova's gen3 manual[9]. The encoders in the robot can measure the values of such transformations. In Figure 10.10 we illustrate the position of every transformation (T_N) in the robot.

$$T_{0,H} = T_{0,1}T_{1,2}T_{2,3}T_{3,4}T_{4,5}T_{5,6}T_{6,H} \tag{10.1}$$

For the long-range tracking, we used a VIVE tracker attached to the wrist of the user. This allows us to prevent collisions and guide the robot whenever the short-range tracking fails to capture and interpret the user's hand. In addition, this allows us to overcome the limited tracking space and quality of the LeapMotion. The VIVE tracker is only a fallback option when the primary tracking fails, so it can be replaced by alternatives that do not require the user to wear any device (following the rationale of mid-air haptic interaction), for example, Realsense Cameras or AI-based IK solvers that can provide information about the user's arm location with relation to the robot.

10.11 Guidance Strategy

The hand is the target point of the end-effector. We first calculate the end-effector's target position (Ultraleap device) and then the rotations. We trace a vector from the robot's base to the hand to calculate the robot's target position. Then we determine an offset H_{off}, which is especially important since it can determine the ultrasonic rendering's intensity and quality.

[9] Kinova Gen3 User Manual

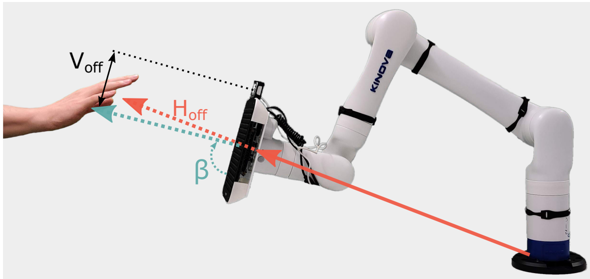


Figure 10.11: We guided the haptic array to follow the user's hand closely. The tracking strategy involves tracing a vector from the robot base to the palm and calculating an offset H_{off}. We also added a vertical correction factor V_{off} to center the hand in front of the haptic array instead of the leap motion coordinate system.

The β angle offset can be used to define the array's rendering direction (facing downwards or upwards). Finally, the value V_{off} is the height correction necessary to keep the hand in the center of the ultrasonic array instead of the LeapMotion sensor. This set of parameters must be tuned depending on the VR scenario to keep the hand within the tracking volume. These values equally impact the short-range and long-range tracking of the hand. It is important to highlight that the rotations of the end-effector are faster than the translation movements. For this reason, we calculate the rotations of the end-effector based on the vector from the end-effector to the hand instead of the vector from the base to the hand. This allows the system to act quickly to adjust the haptic feedback within the space in front of the ultrasonic array, allowing the user to explore an object quickly.

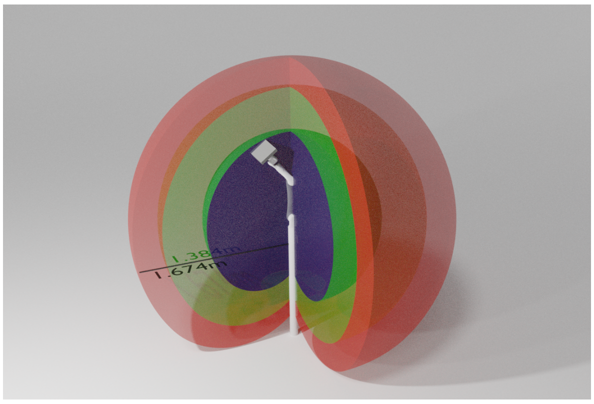


Figure 10.12: Rendering reachable volumes: nominal space of the robot 3.29m^3 (Blue), space within ideal rendering quality 7.42 m^3 (Green) and maximum rendering space 8.27m^3 (lower quality, in Red). This is 19.98 m^3 Volume in total.

10.12 Improvement in workspace Dimensions

The proposed system can drive the array in a spherical radius of 93.4 cm (length from T_2 to the Ultraleap end-effector Figure 10.10).

Table 10.2: Comparison in workspace of different mid-air haptics systems.

Author	Strategy	Rendering Volume [m^3]	Gain
Ultraleap (measures from [452])	Static Array	0.055	1
Howard et al. [452]	Pan-Tilt	0.95	17.2
Brice, McRoberts, and Rafferty [453]	Switch Array Positions	1.5	27.2
Suzuki et al. [454]	Multiple Static Arrays	2.	36.3
Proposed System	**Robotic Driven Array**	19.98	363.2

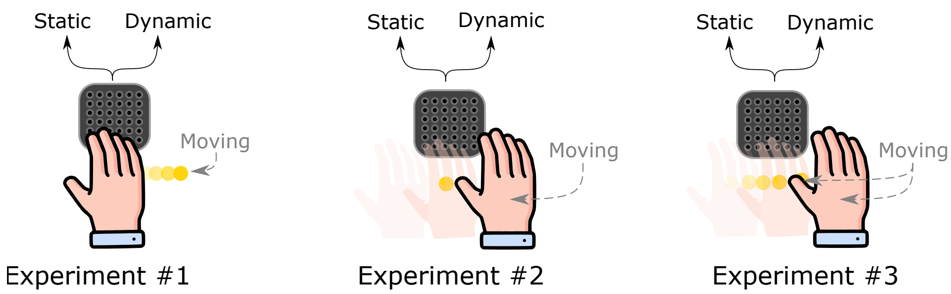


Figure 10.13: Experiment description: Experiment one evaluated a moving focal point vs. a static hand, Experiment two evaluated a moving hand vs a static focal point; and Experiment three evaluated a moving hand and moving focal point. All the experiments had static vs. dynamic arrays as conditions

However, the rendering volume is significantly bigger. If we consider a maximum ideal rendering distance of array-to-hand of approximately 45 cm [451] sensations could be rendered up to a 1.384 m radius. Moreover, if we calculate the rendering workspace using the absolute maximum rendering distance (Low-quality rendering area), the rendering space would increase to a 1.674 m radius. This maps to 3.29m^3 of rendering workspace in the nominal volume of the robot (blue volume in Figure 10.12). In this volume, the ultrasound array can always reach the hand within the ideal rendering distances. Additionally, a volume of 7.42 m^3 can be rendered with the full robot extension within the ideal ultrasound array to hand rendering distances (green volume in Figure 10.12). On the boundaries of the ideal rendering distances, an additional volume of 8.27m^3 can be rendered in a lower quality (Red volume in Figure 10.12). The total volume renderable by the proposed setup (including lower quality volumes) is 19.98 m^3. This represents an increase of 21 times over Howard et al. [452] (0.95 m^3), 13.32 times over Brice, McRoberts, and Rafferty [453] (1.5 m^3), and 9.9 times over Suzuki et al. [454] (2 m^3), making it to the best of the authors' knowledge, the biggest rendering volume achieved in mid-air haptics. See Table 10.2 for values relative to a standard static array.

10.13 Perceptual Characterization

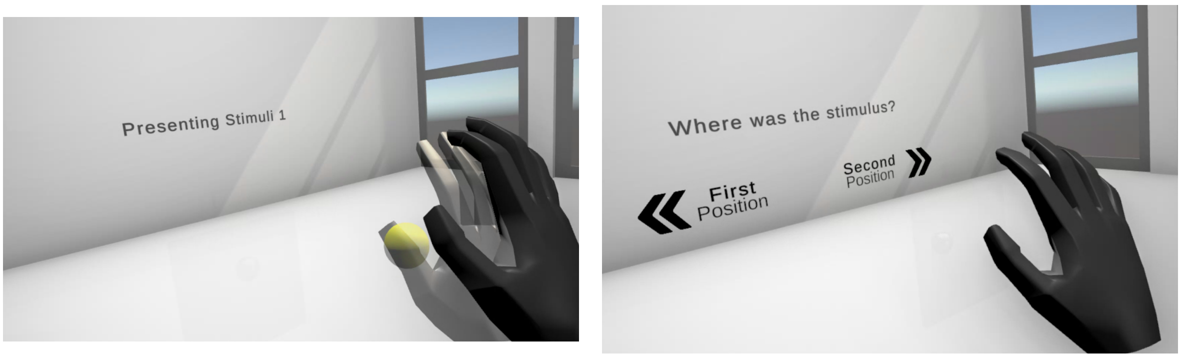
This study aimed to compare the perceptual impact of rendering ultrasonic haptic feedback using a moving array (referred to as dynamic ULTRASONIC ARRAY) versus a stationary array (referred to as static ULTRASONIC ARRAY) using a robotic manipulator. We executed this comparison to gain insights into the perceptual differences between the proposed system and the state-of-the-art in dynamic rendering setups. In detail, we conducted three experiments that cover the most relevant movement combinations of ULTRASONIC ARRAY, HAND MOVEMENT, and FOCAL POINT.

10.13.1 Hypotheses

Mid-air ultrasound haptics is rendered using stationary setups. Factors such as hand speed, hand distance from the ULTRASONIC ARRAY and the propagation of the mechanical waves in the air are typically calculated by observing the assumption that the ultrasound actuators are not being displaced from their original position. As we evaluate the same device in a dynamic setting, we expect the performance to be lower regarding the static ULTRASONIC ARRAY. In particular, we anticipate that the dynamic ULTRASONIC ARRAY will perform worse when rendering stationary elements for active touch since it has to compensate for the robot's movement. As for the moving objects, we expect the dynamic ULTRASONIC ARRAY to closely follow the performance of the static array in both dynamic and static HAND MOVEMENT. However, we expect that the decrease in quality is not significantly high and that the significant gain in the workspace can compensate for such loss.

In Summary:

- **H1:** dynamic ULTRASONIC ARRAY will perform similar to ULTRASONIC ARRAY for passive touch.



(a) Stimuli presentation screen: In the screen, guiding hand, visual focal point presentation, user's hand

(b) Forced choice screen: In the screen, forced choice questions, possible answers, and the user's hand

Figure 10.14: View of the VR environment used for the perceptual characterization.

- **H2:** dynamic Ultrasonic Array will perform worse than static Ultrasonic Array in active touch when the object is stationary in the 3D space.
- **H3:** dynamic Ultrasonic Array will achieve similar performance to Ultrasonic Array for active touch with moving objects.

To address these hypotheses, we designed three experiments: **Experiment #1** Evaluates dynamic Focal Point and static Hand Movement in terms of static and dynamic Ultrasonic Array. **Experiment #2** Evaluates static Focal Point and dynamic Hand Movement in terms of static and dynamic Ultrasonic Array. Finally **Experiment #3** Evaluates dynamic Focal Point and dynamic Hand Movement in terms of static and dynamic Ultrasonic Array, for graphic details, please refer to Figure 10.13.

We defined the aims of each experiments as follows:

- **Experiment #1**: Dynamic Focal Point, Static Hand Movement This experiment was designed to study the passive touch scenario, which is the case where the observer's hand is not moving but the element that they are touching moves. This is reflected in the current setup by moving the focal point across the bases of the proximal phalanx (Palmar digital creases). The decision to stimulate the hand across the Palmar digital creases is motivated by the density of mechanoreceptors involved in the perception of ultrasonic stimulation (Rapid Adapting Afferents) reported for the hand [455, 456]. The closer to the fingertips, the higher the innervation density; therefore, we selected the largest continuous area of the hand closer to the fingers. Volunteers were instructed to hold their hand still in the position indicated by the virtual guiding hand (more details about this virtual guiding hand in subsection 10.13.3).
- **Experiment #2**: Static Focal Point, Dynamic Hand Movement The second experiment aimed to study active touch with a static object setting. This is, the observers actively move their hands but the element they are touching is static. In the setup, this translates into a focal point static in the 3d space and the participant's hand moving from left to right.
- **Experiment #3**: Dynamic Focal Point, Dynamic Hand Movement In Experiment #3 This experiment was designed to study the effects of having several moving elements in play. In the setup, this was implemented by moving the focal point from left to right and asking the volunteers to follow the same movement with their hands.

We omitted the combinations of static Focal Point, static Hand Movement and static Ultrasonic Array since this case has been widely explored in literature and does not provide additional information regarding the set of hypotheses addressed in this investigation.

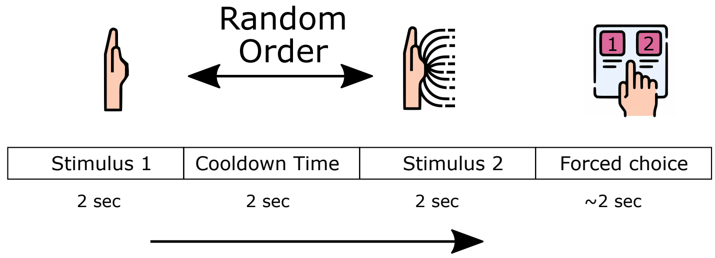


Figure 10.15: Stimuli presentation scheme: We present two stimuli sequentially, one of them contains the actual ultrasound feedback, the other is empty, the order of them is random. After the two stimuli were presented, we asked participants in a 2AFC question where the actual stimuli was presented.

10.13.2 Experimental Design

The three experiments aim to compare the perceptual quality of rendering of dynamic versus static Ultrasonic Array. Perceptual quality can be expressed as the detection sensitivity and detection threshold. Both detection sensitivity and threshold can be inferred from a psychometric function given a combination of parameters defined for **Experiment #1**, **Experiment #2**, and **Experiment #3**. We opted for an adaptive 2-down-1-up staircase method to sample the data required to calculate the psychometric functions. All the staircases started at 100% rendering intensity and adapted at a step of 20% during the first two reversals; then, the step size decreased to 5%. The order of experiments and conditions were randomized using Latin Square. Volunteers executed two conditions per experiment (static and dynamic Ultrasonic Array). The condition ended when volunteers achieved nine reversals, or 35 trials.

Task: Figure 10.15 illustrates the structure of the trials: volunteers were presented with two stimuli, one of them was an ultrasonic vibration, and the remaining one was empty. That is, it did not contain any vibration. We added a 2 second interval between stimulations to allow the robot to return to the start position before the next stimulus. The stimuli were presented in random order. In a two-alternative forced-choice decision (2AFC), volunteers had to identify which stimuli contained the vibration. The trial ended after the volunteers had contact with both stimuli. We implemented contact and movement checks to verify that the task was executed properly in every experiment. *Experiment Parameters:* The hand was placed at a distance of 25 cm from the haptic array, the speed of movement of the focal point, hand, and array was 0.3 m/s and the total distance rendered was 30 cm.

10.13.3 Setup

The experimental setup was composed of an UltraLeap for hand tracking, a Kinova Gen3 6 DoF for driving the array, A Valve Index VR headset, and a VR-Ready laptop (Acer Predator Helios 300 with a NVidia RTX 2070). We implemented a VR environment for conducting all three experiments. The ultrasonic array was positioned vertically in front of the volunteers.

Volunteers placed their hands on an armrest so they did not get tired of holding their hands in a parallel array position. Finally, we rendered a virtual (guiding) hand in VR that guides the volunteers through all the experiments. (i.e., hold the hand static or Actively explore the stimuli).

10.13.4 Participants and Procedure

Eight people participated in the study, one volunteer was female, and one participant was left-handed. Ages ranged from twenty-three to thirty-three years old, with an average age of 28.8 ($SD = 3.3$). None of the volunteers reported any skin conditions that impaired their capability to discriminate tactile sensations. One participant was left-handed. Volunteers were compensated with 10€/h for their participation in the study. *Exclusion Criteria*: We targeted young adults (less than forty years old) according to Peters et al. [457] to guarantee skin sensitivity. Additionally, people with reported skin conditions were excluded from the experiment.

Volunteers were instructed to read through the informed consent and fill up a demographics survey. Next, the experimenter explained the tasks, the elements of the study, and the stages of it. When the participant felt comfortable with the setup, we asked them to put on the headset and follow a tutorial about the experiment. The tutorial included tasks like matching the virtual and real hands, moving their hand from left to right, and selecting one stimulus in the same way they would in the experiment (2AFC). After the tutorial, the first assigned experiment started. Once the participant went through all the experiments and conditions, we debriefed them on the experiment's details. Volunteers were compensated with 10€/h. The study's average duration was one hour.

10.14 Perceptual characterization Results

We calculated the psychometric functions for every experiment and condition using bayesian estimation. Precisely, we followed the approach by Schütt et al. [458]. We modeled all the psychometric functions using normal cumulative sigmoids. The points of subjective equality (PSE) are calculated at the 75% chance of correct responses; this value also reflects the absolute detection threshold (T_A). Similarly, the middle point between the PSE and the higher chance of a correct answer is taken as the detection threshold. Finally, the just noticeable difference (JND), which represents the lower change in intensity that an observer can detect, was calculated as the difference between the PSE and detection threshold. All the data from this experiment is available in the supplementary material. For the sake of simplicity, the intensity values are presented as a percentage of the maximum rendering power of the ultrasonic array; peaks on the STRATOS platform range from $0 - 1125Pa$.

Results for passive touch (**Experiment #1**) show a slightly lower absolute threshold (T_A) for static Ultrasonic Array (T_A = 0.415, Confidence Interval (CI) = 0.630) than for dynamic

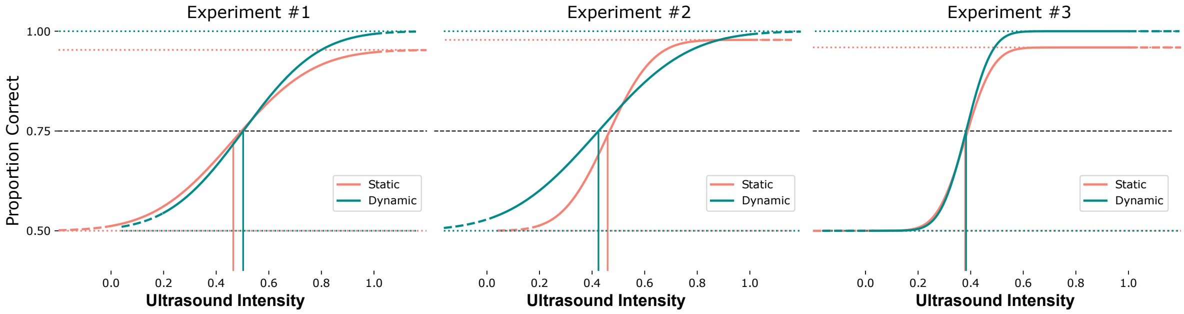
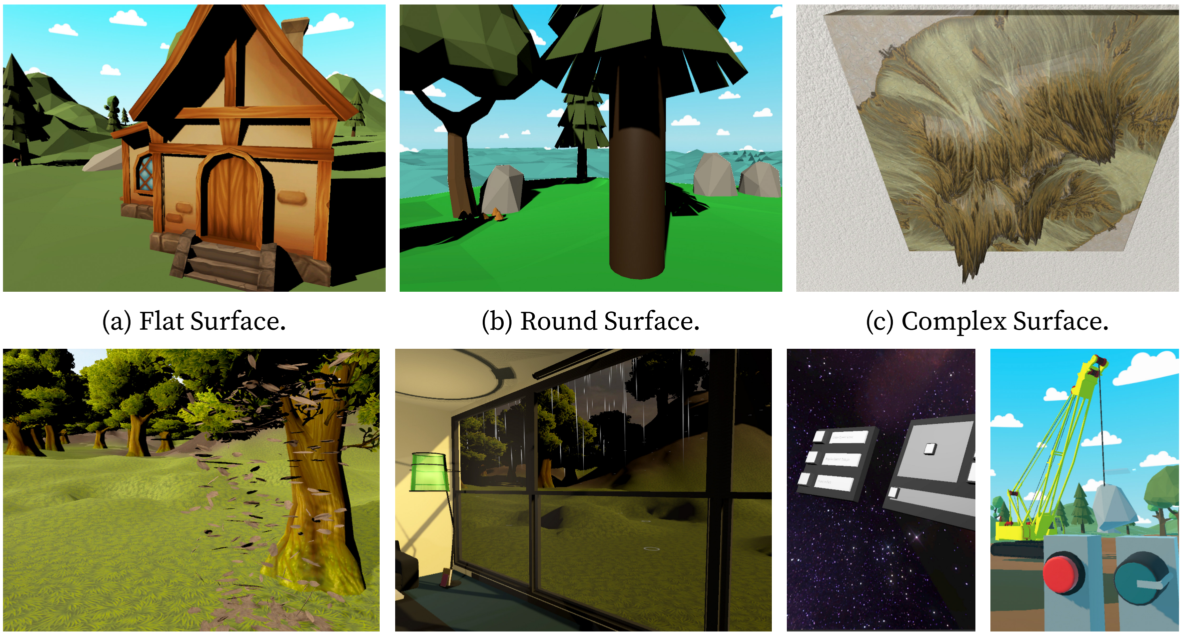


Figure 10.16: We tested the system in seven different scenarios exposing its capabilities and the challenges of the paradigm of array-in-motion mid-air rendering.

ULTRASONIC ARRAY (T_A = 0.433, CI = 0.548), however, the detection sensitivity (Sen_{Static}) was higher in dynamic ULTRASONIC ARRAY ($Sen_{\text{Dynamic}} = 0.885$) than static ULTRASONIC ARRAY ($Sen_{\text{Static}} = 0.745$). Consequently, dynamic ULTRASONIC ARRAY presents a better JND than static ULTRASONIC ARRAY ($JND_{\text{Static}} = 0.165, JND_{\text{Dynamic}} = 0.089$). In the case of active exploration of stationary objects (**Experiment #2**); The absolute threshold was, in fact, lower in the dynamic ULTRASONIC ARRAY condition (T_A = 0.454, CI = 0.747) than in the static ULTRASONIC ARRAY condition (T_A = 0.473, CI = 0.650). Nonetheless, sensitivity and JND were not better than static ULTRASONIC ARRAY ($Sen_{\text{Dynamic}} = 0.592$, $Sen_{\text{Static}} = 1.123$, $JND_{\text{Dynamic}} = 0.090$, $JND_{\text{Static}} = 0.047$). Finally, in **Experiment #3** static and dynamic ULTRASONIC ARRAY showed an identical behavior in all metrics, however, dynamic ULTRASONIC ARRAY performed marginally better sensitivity and JND ($Sen_{\text{Dynamic}} = 2.573$, $Sen_{\text{Static}} = 2.159$, $JND_{\text{Dynamic}} = 0.045$, $JND_{\text{Static}} = 0.056$) but also marginally worse in Absolute threshold; namely static ULTRASONIC ARRAY (T_A = 0.361, CI=0.420) and dynamic ULTRASONIC ARRAY (T_A = 0.366, CI = 0.404).

In all, an overarching question is whether rendering ultrasound mid-air haptics in movement is notably different from the static case, therefore, impacting the rendering quality. We assessed this question by analyzing the likelihood of the two models to be similar (Bayes Factor in favor of the null hypothesis) for the general model across all the experiments using a Bayesian Paired T-test. We found that the Bayes Factor for T_A is 3.8 ($BF_{01} = 3.864$, $Error = 0.015\%$), which means that the distribution of perceptual thresholds obtained from our characterization is 3.8 times more likely to occur under the null hypothesis than under the alternative hypothesis. This likelihood is closely followed by the JND distribution ($BF_{01} = 3.886$, $Error = 0.015\%$). In the case of the sensitivity (Sen) the distribution of data collected is 1.67 ($BF_{01} = 1.672$, $Error = 0.022\%$) times more likely to happen under the case where the systems are similar. This provides support to the fact that even if in specific scenarios, dynamic ULTRASONIC ARRAY can have mechanical advantages and disadvantages, the two systems are likely to be equivalent in terms of absolute detection threshold and just noticeable difference.



(a) Flat Surface. (b) Round Surface. (c) Complex Surface.

(d) Moving Object. (e) Environmental Haptics. (f) UI. (g) Tangibles.

Figure 10.17: We tested the system in seven different scenarios, exposing its capabilities and the challenges of the paradigm of array-in-motion mid-air rendering.

10.14.1 Improvement in Quality

The psychometric modeling of the dynamic setup evidenced similar performance in relation to a static ULTRASONIC ARRAY. In specific cases, it did not reduce the rendering quality but increased it. We discovered that the dynamic rendering mode introduces a tradeoff between rendering quality (Sensitivity & JND) and perceived intensity (Absolute threshold). For dynamic objects (**H1** and **H3**), it did not impact the rendering quality; in fact, in **Experiment #1**, sensitivity and JND values were better for dynamic ULTRASONIC ARRAY than static ULTRASONIC ARRAY. These results go in the direction of **H1** and **H2**, but also show the possibility of improving the rendering quality by moving the array at a similar speed to the moving object. However, this introduces mechanical limitations for fast-moving objects; further investigations are thus required in this regard. Finally, the dynamic setup did have an evident impact on the rendering quality for static objects. This was expected given the compensation in the position that the ULTRASONIC ARRAY has to perform in order to hold still the focal point while moving. Therefore, such behavior is encompassed with **H2**.

10.15 Show Cases

We developed seven showcases to present the systems' capabilities for rendering large-scale mid-air haptics. Thus, we designed all showcases to reflect the need for haptics in various scenarios and showcase the potential of the proposed system. We classified haptics in VR into three groups: general object haptics, environmental haptics, and haptics for user interfaces. The simplest haptic feedback is to present a *Flat Surface*, more advanced are *Round Surface*, and finally, in the group of general objects, we have *Complex Surface*. A special case of general object haptics is *Moving Objects*, which heavily relies on stable tracking. Moreover, they also deliver a feeling of energy in an environment, such as falling leaves activated by wind. Such feedback can also be ubiquitous in the case of rain, and thus, we identified the group *Environmental Haptics*. Finally, the last group we identified is haptics for user interfaces. Here, we see *User Interface* with buttons, a text field, and a virtual keyboard. On the other hand, we identified 3D user interfaces that align with *Tangibles*.

10.15.1 Flat Surface: The Door

Flat surfaces are one of the most commonly used elements that help build more complex geometric objects. However, their simplicity does not imply that they are not helpful as they are widely used in VR applications, e.g., walls, floors, doors, and tables. Thus, flat surfaces are essential for VR gaming and collaborative environments.

In detail, we used a flat 2m tall 1m wide door to showcase how the proposed system can deliver haptic feedback on a large surface, see Figure 10.17a. We had the door be part of a building situated in a low-poly flat-shaded world to support immersion. The world consisted of simple elements such as trees and mountains.

For rendering this flat surface, we used the forcefield rendering mode included in the Unity Ultraleap package that calculates the hand's intersections and renders focal points in the collision coordinates. In the flat surface case, we positioned a plane in front of the door, so every time the participant touches the door's surface, the contact points are stimulated.

In this scenario, we expected participants to use a flat open hand in active exploration of the flat surface.

10.15.2 Round Surface: The Tree

After flat surfaces, round surfaces use the third dimension, allowing for more complex combined objects. While they give designers of VR experiences a greater set to pick from, they are harder to realize from a haptic rendering perspective. Examples of such objects are door handles, advertising pillars, or trees.

We selected a tree trunk. Again, we can show the system's capabilities to render large-scale haptics, Figure 10.17b. As we only focused on scale, we used a simple large cylinder as a tree trunk with a diameter of $0.5m$. The greater scenario of this showcase is a small island with infinite see-around. We placed some other trees and some stones on the island. We again used a low-poly flat-shaded visualization with a cartoon-like skybox.

For rendering round surfaces, we used a similar approach to the first scenario. Still, this time we moved the plane tangentially based on the user's fingers' contact points and the cylinder representing the tree's trunk.

In this showcase, we expected participants to explore the surface using a more relaxed and concave hand position.

10.15.3 Complex Surface: The topology Map

All objects can be modeled as combined complex structures with flat and round surfaces, sometimes comprising thousands of small surfaces. However, this poses several new challenges; for instance, the need for haptic feedback in different directions, e.g., when grabbing smaller objects or the fast switching between haptic feedback directions while moving over a complex surface. For such a scenario, we used a map exploration scenario. Here, touching the different faces requires haptic feedback to move quickly and precisely.

We pick such a map exploration task in an education scenario for showcase, see Figure 10.17c. Therefore, they hang the map on a classroom wall as one would hang a classical 2D map. The environment was a full-scale classroom with desks, chairs, blackboards, and shelves. Large open windows reveal the skyline of the surrounding city for an immersive impression.

Using a similar approach to round surfaces, we moved the rending plane based on the contact points, generating focal points in the intersections between the fingers and the mountains' colliders.

For this showcase, we expected participants to rely mainly on their fingertips for the exploration, given the number of details presented by the topology map.

10.15.4 Moving Objects: Falling Leaves

We move away from a static showcase, with the next showcase toward moving objects. Here, we show that the system can provide static haptics at a new position and render constant haptic feedback while the user's hand follows a moving object.

We designed a tree in autumn where leaves fall to the ground to experience moving objects' constant feedback. This showcase is unique as it shows that the proposed system can provide continuous feedback in a large volume and render feedback to many small targets – the many leaves falling. To provide the users with an atmospheric autumn day, the sun is set low with long shadows, more realistic grass textures, and many trees surrounding the user's distance to deliver the impression of a glade in a forest, see Figure 10.17a.

We generated circular sensations on the palm every time the hand collided with a leaf to render the leaves. The coordinates of the circular sensations followed the leaf movement instead of the hand translation.

In this showcase, we expected participants to quickly change the configuration (from open hand to grasping) and position of their hands in order to catch the leaves.

10.15.5 Environmental Haptics: Rain

For cinematic experiences and storytelling, it is important to deliver haptic feedback for elements in the environment and render environmental haptics such as wind, rain, and snow; cf. Tatarchuk [459]. This further challenges the system as now the tracking and feedback need to provide constant haptic feedback no matter where the user is and how fast they move.

We designed a living room with a large open window to explore environmental haptics for storytelling and not overload the user with haptic feedback. Outside, the user is presented with a rainy summer night Figure 10.17e. Thus, whenever the users hold their hands, they will get heavy rain rendered onto their hands. Again, the large-scale haptic rendering system allows users to explore the entire area outside the window widely.

We render rain by generating scattered focal points in the hand every time the participant takes their hand out of the virtual window. The haptic feedback is presented as long as the user's hand is outside of the window. Furthermore, the feedback is rendered from below the user's hand and not from above. We rendered this scenario in this specific way given the limitations of the system to placing the array facing upside down at fast speeds when the users flip their hands.

We wanted to explore free-hand exploration across the rendered volume; specifically, we wanted to see the impact of participants flipping their hand in the opposite direction of the array and its impact on enjoyment and perception of the feedback.

10.15.6 User Interfaces: Digital UI

Graphical User Interfaces with buttons, sliders, and scrollbars are standard in traditional desktop applications and to navigate through virtual 3D interfaces; cf. Zhang et al. [460]. While user interface elements are typically flat surfaces, we present them in their own showcase due to their importance in interface design. To enable volunteers to experience the feedback provided by the system, we used a simple interface with buttons placed into an outer space skybox, Figure 10.17f. This enabled volunteers to experience the rapid clicking of the same buttons and the fast switching between buttons.

Digital interfaces were previously rendered using circular focal point movements on the palm. We followed the same approach for the buttons and the sliders on the interface.

In this showcase, we evaluate a popular use of mid-air haptics: interface interaction. Specifically, we wanted to check one-finger interactions. Therefore, the interface features small buttons so the participants cannot interact with them using their full hands.

10.15.7 Tangibles: Button & Knob

A special form of user interface elements is tangible UIs; here, in contrast to traditional user interface elements, we see 3D elements such as levers, buttons, and knobs; cf. Fang et al. [450].

We compiled a scene in which a user can steer the rotation of a crane using a button and a knob, see Figure 10.17g. While the knob allows manipulating the crane's current orientation, the buttons allow switching the direction of the rotation. While this is a toy example, the focus is on the tangible elements and not on the complexity of the task.

The haptic rendering in this environment aimed to simulate coherence between the visual and haptic feedback; for the push button, we simulated pressure on the palm whenever the user touched the buttons. In the case of the knob, we rendered a dial sensation on the palm, rotating at the same speed as the knob.

In this scenario, we wanted to complement the interface interactions by adding elements that must be used using the full hand and not only one-finger interactions, such as a big push button and a hand-sized knob.

10.16 Show case Evaluation

To evaluate the system and its capabilities to render haptic feedback at a large volume, we invited 12 volunteers to interact with the seven showcases above. While the proposed system can render different haptic feedback patterns, our investigation solely focused on providing feedback in a large space, allowing the user to move around an object or area of interest freely. Volunteers experienced all seven SCENES with *Haptic Feedback* and with *No Feedback*. We randomized the order of the scenes within each haptic condition and counter-balanced the independent variable HAPTIC.

While there are haptic feedback systems out there, we opted for the *No Feedback* as this is the only real possible comparison. Haptic feedback systems using vibration do require a controller in the hand of the user. The alternative to controllers is body-worn haptic systems (cf. Fang et al. [450]); however, they augment the user too. On the other hand, our system does not occupy the user's hand, nor is it body-worn. Thus, currently, there is no system that provides haptic feedback on this scale. Thus, the only fair comparison is a *No Feedback* comparison.

10.16.1 Apparatus

We used the system in a silenced $\sim 30m^2$ large room. The robotic arm was mounted on a desk situated in the center of the room. We provided visual feedback of the robot end-effector inside the virtual environment to ensure participants' safety and avoid breaks in immersion due to unexpected elements in the real world. A semi-transparent red square represents the end-effector. In all the conditions, participants had feedback about the position of their hands.

10.16.2 Participants and Procedure

In total, we invited 12 volunteers (6 female, 6 male), with an average age of 26.2 ($SD = 2.2$). Two volunteers own VR headsets, and two stated they had no prior experience with VR. In addition, all but one participant were right-handed. Finally, three volunteers indicated that they had experienced haptic devices beyond common vibration feedback (e.g., smartphones) beforehand.

After welcoming volunteers, we explained the study and answered any questions before starting the study. One experiment guided the volunteers through the survey during the study, which took around 40 minutes. Volunteers started either with or without haptic feedback. Here, volunteers were asked only to use their right hand to explore the environment. The exploration tasks did not include any dexterous manipulation, and task performance was not measured. Therefore, the handedness of the participant would not impact the execution of the task. The volunteers experienced each scene for around 1 minute, in which they were free to move and explore the target objects. On a 7-point scale, we asked volunteers to rate the following four questions: "The experience was enjoyable?" "I felt a strong feedback," "The virtual object felt real," and "I did not feel anything when my hand touched the object." The first one was inspired by the User Experience Questionnaire [461], and the other three by other haptic investigations, e.g., [450, 452, 462]. Additionally, after the volunteers finished one haptic condition, we asked them to fill out a Igroup Presence Questionnaire (IPQ) [463]. Finally, we noted down relevant comments and reactions the volunteers had during their participation. While the robot is designed for humans as an assertive robot and, thus, should not harm a human in a way, the experimenter always had an emergency button within arms reach. Moreover, to protect volunteers from the ultrasound, we asked them to wear earplugs, which we provided.

10.16.3 Results

Overall, all volunteers reported an excellent level of positive feedback on the use of the system. In particular, they were positively surprised by how well they could experience the complex surface – the mountain; for instance, P5 stated, "[the mountain] was the coolest." Other comments include P6 "I felt the leaf falling through my hand." Moreover, P12 stated that the environment feels not responsive; here, P12 described being disconnected from the virtual world.

With the designs, e.g., using low-poly flat-shaded scenes, we aimed to not focus on the feedback itself but on the system's ability to render feedback in a large space. However, volunteers would have liked to feel the tree's bark and stated that this was a missed opportunity not to render the detailed texture (P11); cf. Freeman et al. [464]. Moreover, P3 stated they felt a mismatch between the real world and VR as the feedback was mid-air and not force feedback. Finally, P1 and P2 reported tracking issues that we could trace back to the LeapMotion, as these two volunteers had particularly small hands.

10.16.4 Quantitative Feedback

As all of the questionnaire results for enjoyment, strong feedback, object felt real, and no touch sensation where not normally distributed ($W = 0.920$, $p < .001$; $W = 0.854$, $p < .001$; $W = 0.890$, $p < .001$; $W = 0.800$, $p < .001$; respectively) we aliened and ranked the data first using ARTool [465].

First, we investigated whether HAPTIC or SCENE significantly influenced **enjoyment**. Therefore, we conducted a two-way ANOVA. The analysis revealed a significant effect for HAPTIC and SCENE ($F_{1,143} = 32.196$, $p < .001$; $F_{6,143} = 3.410$, $p < .004$; respectively). However, we found no significant interaction effect ($F_{6,143} = .689$, $p = .659$). Post hoc t-tests revealed that only *Round Surface* vs. *Tangibles* is significantly different ($t(143) = -3.305$, $p = .025$), all others are ($p > .05$).

Next, we were interested if the **strong feedback** was significantly influenced by HAPTIC or SCENE. The analysis revealed a significant effect for HAPTIC and SCENE, and a significant interaction effect ($F_{1,143} = 267.386$, $p < .001$; $F_{6,143} = 2.547$, $p = .023$; respectively). Due to the significant interaction effects, we only looked at the six important post hoc comparisons

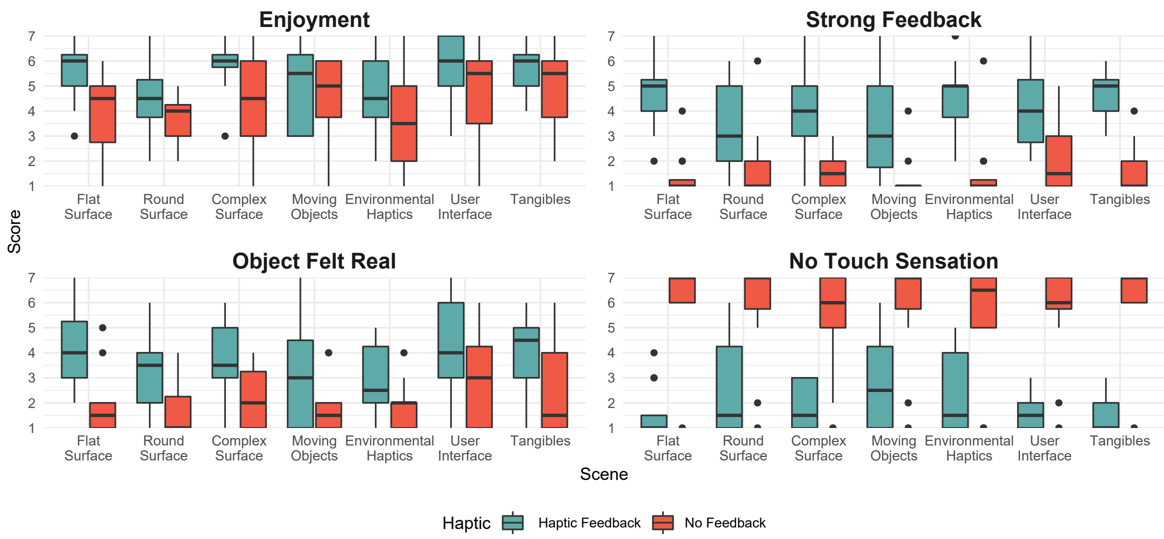


Figure 10.18: The score of the four questions independent for HAPTIC × SCENE.

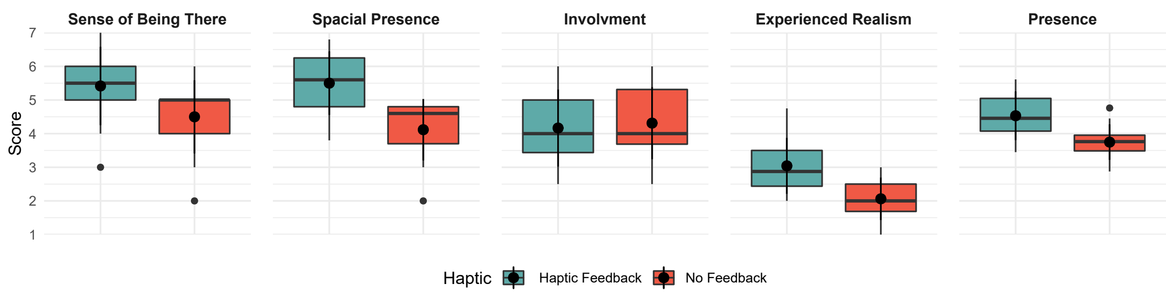


Figure 10.19: The rating of the four sub-scales of the IPQ questionnaire [463] and the combined presents the score for the two conditions *Haptic Feedback* and *No Feedback*.

comparing only within each SCENE. T-test reveal that all differences are $p < .001$, see Figure 10.18.

Next, we investigated whether HAPTIC or SCENE significantly influenced if the **object felt real**. We again conducted a two-way ANOVA. The analysis revealed a significant effect for HAPTIC and SCENE ($F_{1,143} = 67.197, p < .001$; $F_{6,143} = 2.684, p < .017$; respectively). However, we found no significant interaction effect ($F_{6,143} = .757, p = .605$). Post hoc t-tests reviled that only *Environmental Haptics* vs. *Tangibles* is significantly different ($t(143) = -3.181, p = .038$), all others are ($p > .05$).

Finally, we were interested if the feeling of **no touch sensation** was significantly influenced by HAPTIC or SCENE. The analysis revealed a significant effect for HAPTIC and a significant interaction effect ($F_{1,143} = 248.679, p < .001$; $F_{6,143} = 2.231, p = .044$; respectively). However, no significant effect for SCENE ($F_{6,143} = 1.611, p = .148$). Due to the significant interaction effects, we again looked at the six meaningful post hoc comparisons, comparing only within each SCENE. The T-test reveals that all differences are $p < .001$, see Figure 10.18.

10.16.5 Presence

First, we analyzed the IPQ questionnaire [463] (scale: from 1 to 7) to determine whether HAPTIC significantly influenced the presence in VR, see Figure 10.19. As the sub-scale "Sense of Being There" was not normally distributed ($W = 0.908, p = 0.032$), we decided to run the Wilcoxon signed-rank test, which does not assume the normality of the data.

The tests revealed that the sub-scales "Sense of Being There," "Spacial Presence," and "Experienced Realism" are all significantly different between the condition *Haptic Feedback* vs. *No Feedback* ($V = 2, p = 0048$; $V = 1, p = .005$; $V = 0, p = .004$; respectively) where *Haptic Feedback* outperformed *No Feedback*. However, there was no significant difference for the subscale "Involvement" ($V = 27, p = .635$). Finally, the overall feeling of presence scored by the IPQ was higher in the condition with haptic feedback ($V = 5, p = .005$).

10.17 Discussion

We encompassed our efforts towards outperforming the current state-of-the-art in workspace size and render quality for mid-air ultrasound haptics for Virtual Reality. With the proposed system, we achieved a rendering volume almost ten times bigger than the highest rendering volume reported in the literature (Suzuki et al. [454]). The synergy between an ultrasound emitter and a serial robot increases the haptic array's workspace by optimizing the array rotation and location within the nominal actuation range of the robot. However, physically moving parts introduce an additional set of safety considerations to the design process. We tackled this by maintaining the robot's safety features (such as collision detection and maximum speed) and implementing an over-damped PD control that ensures soft movements of the haptic array. On the other hand, observing such safety measures reduces the system's versatility in speed or acceleration. It is technically possible to increase the robot's reaction speed to match the hand's movement speed (for example, with a full PID controller setup and overriding the speed constraints). However, additional safety measures should be introduced to protect the user from unwanted collisions. We executed a set of tests to model the end-effector behavior in terms of rotations and translations. Additionally, the hand guidance strategy used to enable the robot to follow the hand favors configurations where the haptic array is facing outwards from the base of the robot; While this strategy can exploit the potential of the increased workspace, a more elaborated strategy could allow the end-effector to quickly relocate facing down to render sensations even if the palm of the user is facing up.

10.17.1 Challenges of extending Ultrasound Haptics Workspace using Moving Arrays

Besides the technical challenges that can be faced when using a setup like the one presented in this manuscript, a primary design concern is the impact of motion on rendering quality. This is how the movement of the emitter alters the perceived quality of the ultrasound stimuli. We explored this impact using psychometric characterization as this is a central topic in the context of this manuscript. This puts our perceptual characterization as the first study reporting the impact of the emitter's movement on the perceived rendering quality of mid-air ultrasound haptics.

With this exploration, we identified that while a moving array closely matching the speed of the virtual moving object could lead to better perceptual features and, in the opposite case, a static object with a moving array could tend to perform less efficiently. We discovered that the overall performance of both systems is likely to be equivalent (Bayesian T-test). This means that even with the gain on rendering volume, both arrays can perform similarly; with an added gain, the rendering distance of the dynamic array to the hand is smaller, so the hand stays within the optimal rendering distance.

10.17.2 Contributions of an extended Haptic Rendering Workspace to Presence

Only one IPQ scale did not perform better than the baseline. However, this is not surprising as involvement with the scene and the environment's interaction did not change. Moreover, we found that volunteers enjoyed exploring the different showcases. We also found that we could significantly boost their enjoyment when providing haptic mid-air feedback. Combined with the technical evaluation and its result that the workspace is more than 19 m^3 large, we believe that with the proposed setup, we make a step towards large-scale mid-air haptic feedback, which many systems can benefit from.

In the STRONG FEEDBACK questionnaire, a small number of participants rated the feedback as strong (above 1) in the NO FEEDBACK condition; the reason for this phenomenon could be the presence of phantom vibrations after being exposed to the vibrations produced by mid-air haptics. However, more evidence is necessary since the mid-air haptics community has not explored this phenomenon so far.

10.17.3 Extending Ultrasound Haptics Workspace; What is the best Approach?

While several authors have reported a significant increase in the workspace of ultrasound-based mid-air haptics, these approaches have been highly heterogeneous, and yet ours significantly differs from those already reported in the literature:

Howard et al. [452] pan-tilt device tackles the problem by mounting the haptic array on a two-degree-of-freedom rotational platform. Such a setup considerably increases the workspace (around 17 times). However, the fundamental problem of the rendering quality remains untouched, given that the array is still spatially anchored to a fixed point. Instead, the proposed system drives the array closer to the user's hand, which impacts the rendering intensity and quality. A significant disadvantage of the proposed system is the control complexity and cost difference (Howard et al. [452] reported that their system cost only 150 Euro). However, the rendering volume and quality difference is considerable; the proposed system has a workspace 21 times larger than Howard et al. [452].

Brice, McRoberts, and Rafferty [453] approach instead was to relocate the haptic array in a set of pre-defined locations using a serial manipulator. Their approach increased the volume-rendered around 27 times with five pre-defined locations. While this setup can drive the array from one position to the other, it does not dynamically adapt to the user's hand position. Furthermore, it does not consider rotations of the haptic array and does not render while the array is relocating. Such an approach is inherently simpler from the control perspective and safer in many conditions, given that the number of situations where the robot is moved is considerably lower. Nevertheless, the implementation carries many of the static array limitations given that the approach is fundamentally the same but with added relocation. The proposed system instead continuously optimizes the position and rotation of the array to be close to the user's hand across a bigger workspace.

Finally, perhaps the most divergent approach in this ecosystem is the one proposed by Suzuki et al. [454]; Enlarging the ultrasound array by increasing the number of transductors. The approach by Suzuki et al. [454] advantage is the lack of moving parts in the setup, making it inherently safe concerning collisions within the interaction space. Additionally, the focal point relocation is potentially faster in extreme positions, given that the emitters do not have to switch positions or locations. Suzuki et al. [454] proposed setup enlarged the rendering volume around 36 times compared with the original Ultraleap workspace. However, the setup is technically more complex and energy inefficient, considering the overheating of the haptic arrays. Furthermore, the rendered volume is still considerably smaller than the one achieved with the proposed setup.

A major advantage of the proposed setup over the cited alternatives is the system's scalability. Cobots are increasingly becoming more popular and affordable. They can be used in a broad number of contexts, including haptics. This favors the proposed modular design that allows us to mount and unmount the haptic array from the robot's end-effector to switch from one proposed to the other easily. Moreover, a potential extension of the current work could be the integration of encountered-type haptics and extended ultrasound mid-air haptics to render kinaesthetic and tactile-only sensations with a single setup. Such an advantage is unfortunately not possible with the compared alternatives given the highlighted limitations on their approaches.

Encountered-type Thermal Haptics in Virtual Reality

Figure 10.20: We introduce an Encountered-type Haptic Display that incorporates thermal feedback to enrich VR experiences, our study shows significant improvements in user immersion and haptic realism.

In this section, we designed an (Encountered-type Thermal Haptic Display) ETHD end-effector that enables VR thermal rendering. Unlike previous approaches, we do not use pneumatics or hydraulics to vary the temperature of the tactile surface; instead, we strategically placed an array of Peltier elements in the end-effector surface to generate the desired thermal stimuli while keeping the hands of participant unoccluded for interaction. This section describes

Figure 10.21: In the VR experience, the participant interacts with a microwave and toaster, experiencing thermal feedback. Subsequent images show real-world views of the scenarios.

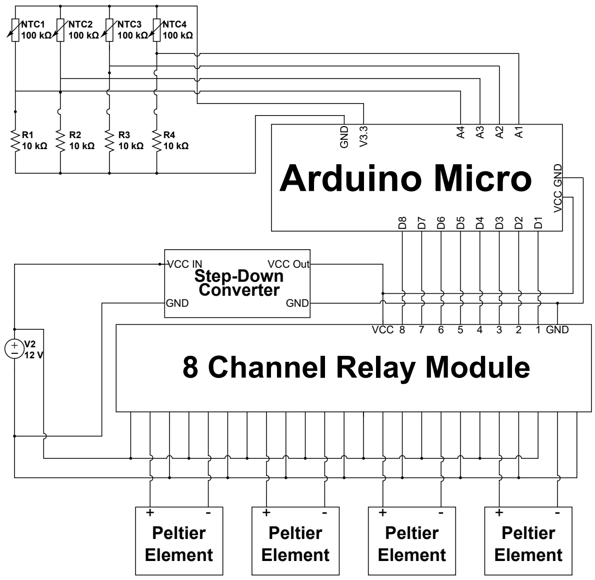


Figure 10.22: System electronic schematic: The system is powered by an Arduino micro microcontroller and an 8 channel relay module for power control of the Peltier modules

our thermal encountered-type haptic display system and design strategies. In detail, we (1) describe the hardware required to build the interface, (2) report the user intent prediction strategy to locate the end effector in the virtual environment timely, and (3) propose two different thermal rendering strategies. All the materials used in this section are available at `https://github.com/mimuc/RoboThermalHaptics` for repplicability.

10.18 Hardware

We used a Kinova Gen3 Cobot as our base platform. The cobot was controlled directly from the graphic engine Unity3D using a control package provided by Villa and Mayer [Exp7]. We used an Acer Predator Laptop with a Nvidia 2070 graphic card and an HTC Vive Pro Headset to run the VR application. We also used a LeapMotion for hand tracking and a Vive Tracker for detecting the hand position outside the LeapMotion tracking as done previously by [Exp5].

10.19 End-effector Design

Our end-effector design is based on a 3D-printed model containing all the system components. The model is designed to be attached to the robot arm using a default Kinova connection. It comprises four parts: the main body, 225mmX225mm cooper plate, securing frame, and a circuitry box. The main body contains the core components of the system, the copper plate serves as a thermal transfer medium and provides a flat surface for the thermal stimuli, the securing frame is used to keep the copper plate in place and pressured to the peltier elements, and, finally, the circuitry box host the electronics of the end-effector.

The cover is attached to the main body using four M3 screws. The main body is attached to the robot arm using the Kinova default connector. The end-effector is designed to be printed on a consumer 3D printer (Bambu Lab p1s). The end-effector weighs 2.3 kg, which is well within the payload capacity of the robot arm (4kg).

The main body of the end effector has four spaces of 40mmx40mm to fit the Peltier modules and four smaller 10mm radius holes for the temperature sensors in the frontal side. On the back side, it has four screw spaces around the Peltier holes to fit heat sinks and cooling fans. The interface between the Peltier modules and the copper plates was filled with small copper inserts and thermal pads for optimal heat transfer, while the interface between the Peltier modules and the heat sinks was filled with thermal paste for dissipation. The holes of the temperature sensors were filled with thermal paste to ensure heat transfer between the copper plate and the body of the sensors. We used 4 Peltier elements rated 57 Watt; for the heat sink and cooling fans, we used the AMD Wraith Stealth Socket AM4. For the temperature sensors, we used NTC thermistors. We selected copper as the conducting material given its high thermal transfer compared to materials such as aluminum or iron, and the plate had a thickness of 1mm to optimize thermal transfer speed.

10.20 Circuit Design

The electronics driving the end-effector functionality is composed of an Arduino micro microcontroller which controls an 8-channel relay module that drives the power from the power supply to the Peltier modules. The power supply is set to deliver 12V and a maximum of 10A to the system. We used a step-down circuit to reduce the voltage from 12V to 5V supported by the Arduino board. For temperature measurements, we use four NTC thermistors $100k\Omega$. We used an HCS 3602 USB power supply which can deliver a maximum of 32V at 30A in the current prototype. For details on the circuit configuration, see Figure 10.22

10.21 Firmware Design

The microcontroller featured a simple binary delayed setpoint strategy for temperature control; As thermodynamic phenomena are typically slow in nature, we implemented a binary temperature control strategy with a dead zone and a switch delay. Temperature control is implemented for each peltier element individually. When the temperature measured by the closer temperature sensor is below the setpoint plus the temperature tolerance, the Peltier element will be set to heat; when the measured temperature by the closer temperature sensor is above the setpoint minus the temperature tolerance, the Peltier element will be set to cool-down. The system communicates with the virtual environment using serial communication via USB and allowing individual control of the Peltier elements and set of target temperatures for each element.

10.22 Intent Prediction Strategy

The intent prediction algorithm incorporates gaze tracking and hand velocity assessment, with the interaction phase dictated by the latter's speed. The system persistently monitors the hand's position and velocity, triggering a gaze-based target prediction via raycasting from the head's position until the hand speed crosses a 3 cm/s threshold.

Upon reaching or surpassing this velocity threshold, the algorithm assumes a straight-line reaching trajectory and employs a direction vector from the index finger's tip for raycasting to identify potential interaction points within the environment. If the projected interaction target lies beyond the system's reachable workspace—defined as a radius of 90 cm from the robot base—a reach redirection strategy is implemented. Here, the redirection origin is marked at the point where the hand velocity first exceeds the threshold.

This approach integrates the REACH+ [15, 466] algorithm for refining the hand position offset and applies a smoothstep interpolation for a natural interaction flow. User disengagement is recognized when the hand's average speed toward an intended target falls below 0.5 cm/s for a duration exceeding 100 ms. For experimental purposes, the system architecture was augmented to incorporate a reactive relocation feature, enabling the robot to adjust its position towards a pre-established location of the interactable object as dictated by the experimental task parameters. This enhancement retains the foundational intent prediction mechanisms of hand speed thresholding and gaze tracking yet introduces a dynamic spatial adjustment to facilitate user interaction.

10.23 Thermal Rendering Strategies

Thermal phenomena are slow in nature, making it challenging to achieve drastic changes in temperature in a short time, especially when the area to heat up or cool down is big. To address these challenges, we propose two thermal rendering strategies that can be leveraged depending on the rendering requirements:

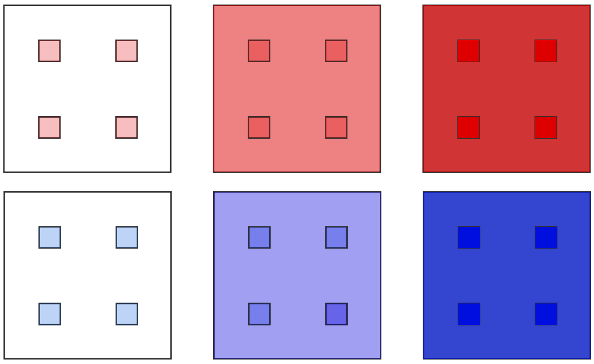


Figure 10.23: Render Strategy 1: All the Peltier modules act together to achieve a whole-plate temperature level

10.23.1 Whole plate rendering method

With this method, it is possible to render a target temperature across the whole plate, enabling a bigger area of touch. In this method, the target temperature is set to the desired value, and the plate will reach the desired temperature within a given time. Then, using the end-effector intent prediction, the end effector is located in the encountered locations during the interaction. As a disadvantage, the time to reach the target temperature can be high, especially in temperatures far from room temperature, given extreme temperatures require a higher energy consumption (see Figure 10.23).

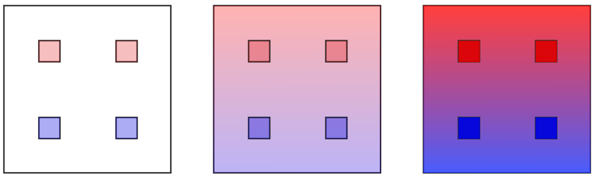


Figure 10.24: Render Strategy 2: A thermal gradient is rendered by setting two elements to heat up and two to cool down, temperature selection is done by setting the encountered location at the desired temperature

10.23.2 Gradient rendering method

With this method, it is possible to render a spectrum of temperatures across the whole plate, quickly enabling access to a wide range of temperatures. In this method, half of the Peltier elements are set to a high temperature. In contrast, the other half is set to a low temperature and held in this configuration during interaction. Then, using the end-effector intent prediction, the end effector is located in the encountered locations during the interaction, and the target temperature is modulated using the REACH+ algorithm to redirect the users to touch the desired temperature point in the gradient (see Figure 10.24). As a disadvantage, the temperature interaction area can significantly be reduced, which can be especially noticeable when the user uses the whole hand to interact or slides their finger through the surface.

10.24 Static Thermal Test

In order to characterize the end-effector's thermal capabilities, we conducted a thermal test following three scenarios: (1) Starting from room temperature and heating up the plate until 85 degrees celsius. (2) Starting from room temperature and cooling down until 10 degrees celsius. Finally, (3) starting from room temperature and generating a heat gradient in the plate. All temperatures were measured using an Optris 28-0023 thermal camera. Figure 10.25 and Figure 10.26 depict the transient temperature progression in each scenario, while Figure 10.27 shows the measured temperature plots in each element. The test for scenario (1) yielded that the end effector achieved a temperature of 70 degrees Celsius in the first minute and a temperature of 85 degrees Celsius for almost four minutes, an absolute difference of 56 degrees Celsius. In contrast, in scenario (2), After four minutes, the system reached an average of 14 degrees Celsius, an absolute temperature change of 14 degrees Celsius. Achieving a stable gradient took a total of four minutes; the measured temperature between the sides of the gradient was 20 degrees Celsius.

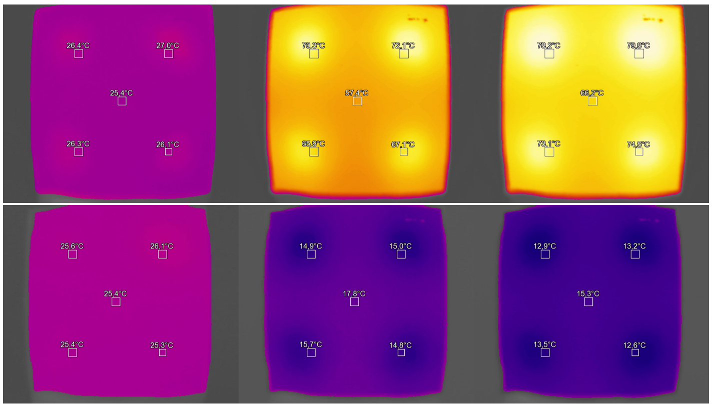


Figure 10.25: Thermal test for whole-plate temperature control. Top: Heating up, Bottom: Cooling down, both starting from room temperature

10.25 User Experience Assessment

ETHDs have been shown to improve immersion [467], realism [468], and overall experience in VR [469]; therefore, the baseline condition of this study was chosen accordingly: we evaluate

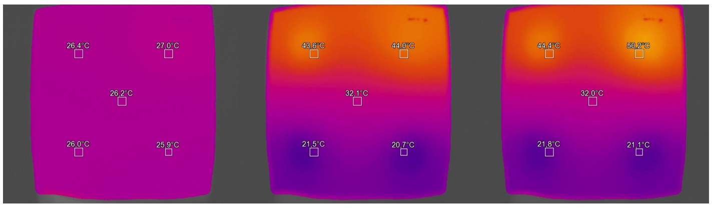


Figure 10.26: Thermal test for gradient temperature control

the encountered type of haptic feedback with and without additional thermal feedback on top. This reduces confounding variables, such as the kinaesthetic feedback being the main contributor to the haptic experience or immersion. Therefore, the results of this study have to be considered in addition to the benefits of ETHDs that have already been reported in the literature [470]. To assess the capabilities of the proposed system to enhance the VR experience, we conducted a within-subject, in-lab study with 26 participants using the gradient rendering method of the system with and without thermal feedback. In this section, we describe and report on such a study.

10.25.1 Task

Participants were required to engage in a VR exploration game involving interaction with various objects within the virtual environment: a cutting board, cereal box, microwave, toaster, sandwiches in Fridge, and a cake container. They were instructed to navigate around these objects to explore them, receiving directional cues for subsequent interactions. Each object was designed to convey thermal properties reflective of its type and condition. The task was finished once the participant interacted with all the objects.

10.25.2 Participants

26 participants took part in the experiment, from which 2 were removed from the analysis given due to irregular setup behavior, leading to a total of 24 participants; participants were primarily University Students with an average age of 23 years old (M=23.04, SD= 2.20); 6 participants self-reported to be female, 16 to be male and 1 preferred not to disclose. One participant reported high familiarity with VR, 11 reported using VR often, and 12 reported low familiarity with VR. Participants were recruited using the university communication channels. Each participant was compensated with 10 Euros/Hour. The study had an average duration of 50 minutes (M=50.60, SD=11.8). The recruitment and study procedures were conducted in accordance with the LMU Munich IRB guidelines to ensure the ethical treatment of all participants.

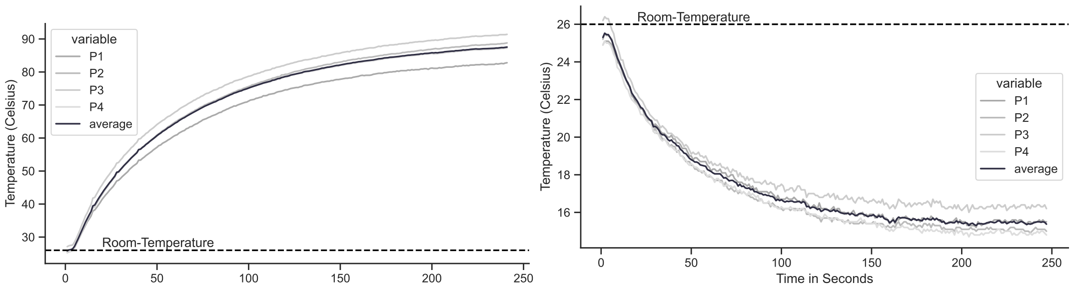


Figure 10.27: Thermal response. Top: Heating up, Bottom: Cooling down, starting from room temperature. PN refers to Peltier elements

Table 10.3: HX Questionnaire Items: Please notice that we used the original item labeling proposed by the authors of the questionnaire

Item	Question	Factor
R1	The haptic feedback was realistic	**Realism**
R2	The haptic feedback was believable	**Realism**
R3	The haptic feedback was convincing	**Realism**
H3	The haptic feedback felt disconnected from the rest of the experience	**Harmony**
H5	The haptic feedback felt out of place	**Harmony**
I1	The haptic feedback distracted me from the task	**Harmony**
H2	I like having the haptic feedback as part of the experience	**Involvement**
I2	I felt engaged with the system due to the haptic feedback	**Involvement**
E4	The haptic feedback changes depending on how things change in the system	**Expressivity**
E5	The haptic feedback reflects varying inputs and events	**Expressivity**
E1	The haptic feedback all felt the same	**Expressivity**

10.25.3 Measures

We collected VR experience data with two questionnaires; the Igroup Presence Questionnaire questionnaire measured presence in three subscales: Spatial Presence, Involvement, and Realism [471]. And the HX model questions proposed by Anwar et al. [472] and originally derived from Sathiyamurthy et al. [473] to measure haptic experience across multiple haptic modalities. This set of items includes four factors: Realism, Harmony, Involvement, and Expressivity. While the realism factor measured by the IPQ focuses on the contrast between virtual and real feedback, assessing how closely virtual experiences mimic real life, the HX realism evaluates the plausibility of the haptic feedback itself. Finally, we included 3 custom questions: *How easy was it to identify objects through physical interaction, like touching an object or bumping into an object? (Q1).* , *How easily did you adjust to the control devices used to interact with the virtual environment? (Q2).* , and, *Was the information provided through different senses in the virtual environment (e.g., vision, hear- ing, touch) consistent? (Q3).*

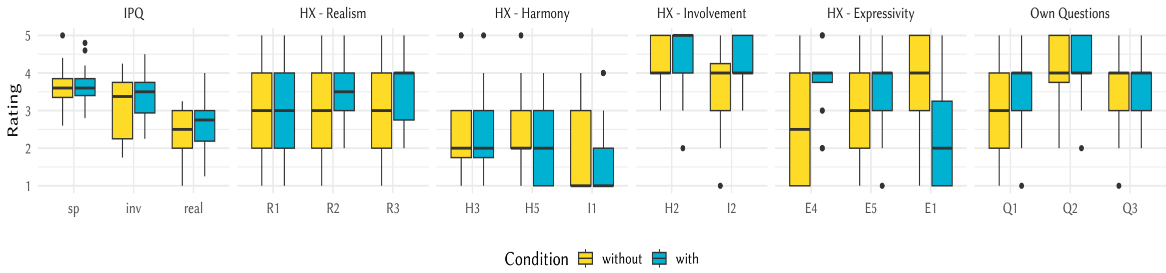


Figure 10.28: Boxplots of the dependent variables meassured in the study.

10.25.4 Procedure

On arrival, the experimenter introduced the participant to the study goals and procedure and requested informed consent for participation. Afterward, the participant was asked to fill out a pre-study questionnaire, asking about their previous experience with VR, handedness, and demographic information. The experimenter then explained the study's procedure and the experimental tasks.

The participant was then asked to put on the VR headset, and the experimenter proceeded to calibrate the VR setup for the participant. Afterward, the participants were asked to interact with the objects to familiarize themselves with the VR setup and the interaction mechanics. The task included interaction with the whole system, where the participant had to touch several points in the virtual objects with a virtual representation of the robot so that the participants knew that the physical robot was moving.

After the training tasks, the participant was asked to perform the experimental tasks without feedback from the real robot location so as not to break immersion. The participant was asked to perform the tasks in a counterbalanced order and complete the questionnaires after each condition. After completing all tasks, the participants participated in a semi-structured interview, were debriefed about the experiment, and were finally compensated.

10.25.5 Results

In the following section, we report the results of our user study in which we compared the utility of temperature feedback (`with temperature`) against a system without temperature feedback (`without temperature`). For aggregated (grouped and averaged) values like the IPQ questionnaire, we report the mean (M) and standard deviation (SD). For individual questions, we report the median and the Median Absolute Deviation (MAD) as a measure of variability. Figure 10.28 depicts all dependent variables grouped by our two conditions.

10.25.6 IGroup Presence questionnaire

We analyzed the impact of our independent variable on the participants' perceived presence using the IGroup Presence questionnaire (IPQ). Similarly to the approach of the IPQ's authors[10], we calculated the group means and used parametric tests to analyze the data. For this, we first checked the data for normality using Shapiro–Wilk's test, where we found no violation of the assumption of normality (sp: $W = 0.93, p = 0.11$, inv: $W = 0.94, p = 0.18$, real: $W = 0.92, p = 0.07$). Therefore, we proceeded with paired-samples t-tests. The analysis indicated a significant ($p < .05$) effect on *realism (real)* (`without temperature`: $M = 2.44$, $SD = 0.6$, `with temperature`: $M = 2.65$, $SD = 0.65$). Besides that, the analysis did not reveal significant effects for the other two subscales *spatial presence (sp)* (`without temperature`: $M = 3.64$, $SD = 0.55$, `with temperature`: $M = 3.69$, $SD = 0.47$) and *involvement (inv)* (`without temperature`: $M = 3.08$, $SD = 0.83$, `with temperature`: $M = 3.39$, $SD = 0.59$).

Table 10.4: Significance test for the dependent variables.

	DV	Test	Statistic	p	sig
IPQ	sp	t-test	t(23) = -0.45	.659	
	inv	t-test	t(23) = -1.8	.086	
	real	t-test	t(23) = -2.29	.032	*
HX - Realism	R1	wilcox	W = 23	.239	
	R2	wilcox	W = 16.5	.092	
	R3	wilcox	W = 56	.651	
HX - Harmony	H3	wilcox	W = 62	.916	
	H5	wilcox	W = 68	.668	
	I1	wilcox	W = 30	1.000	
HX - Involvement	H2	wilcox	W = 12	.562	
	I2	wilcox	W = 9	.009	**
HX - Expressivity	E4	wilcox	W = 26	.007	**
	E5	wilcox	W = 10.5	.013	*
	E1	wilcox	W = 167.5	.002	**
Own Questions	Q1	wilcox	W = 3.5	.027	*
	Q2	wilcox	W = 8	.453	
	Q3	wilcox	W = 24	.505	

10.25.7 Haptic Experience

Further, we assessed the haptic experience of our participants using the Haptic Experience questionnaire as proposed by Anwar et al. [472] Similar to [474], we decided to test the items individually. Because of the non-parametric nature of the data, we used Wilcoxon signed-rank

[10] `https://www.igroup.org/pq/ipq/data.php`, Retrieved July 22, 2024

test to test for significant differences between the paired samples. In the following, we only report the significant results. The remaining questions as well as all significance tests can be found in Table 10.4.

The analysis indicated a significant ($p < .01$) effect on the item *I felt engaged with the system due to the haptic feedback (I2)* with a higher agreement for `with temperature` ($M = 4$, $MAD = 0$) compared to `without temperature` ($M = 4$, $MAD = 1.48$).

Also, we found a significant ($p < .01$) effect on the item *The haptic feedback changes depending on how things change in the system (E4)* with, again, higher agreement for `with temperature` ($M = 4$, $MAD = 0$) compared to `without temperature` ($M = 2.5$, $MAD = 2.22$). Additionally, the analysis showed a significant ($p < .05$) effect on the item *The haptic feedback reflects varying inputs and events (E5)*. As before, `with temperature` ($M = 4$, $MAD = 1.48$) resulted in higher agreement compared to `without temperature` ($M = 3$, $MAD = 1.48$). Finally, we found a significant ($p < .01$) effect on item *The haptic feedback all felt the same (E1)* with higher agreement for `without temperature` ($M = 4$, $MAD = 1.48$) compared to `with temperature` ($M = 2$, $MAD = 1.48$).

10.25.8 Custom Questions

Besides the standardized questionnaires, we employed three custom questions to gain further insights into the appropriateness of temperature feedback for encountered-type haptic feedback. First, we asked participants *How easy was it to identify objects through physical interaction, like touching an object or bumping into an object? (Q1)*. The analysis indicated a significant ($p < .01$) effect with higher ratings for `with temperature` ($M = 4$, $MAD = 1.48$) compared to `without temperature` ($M = 3$, $MAD = 1.48$). Second, we asked our participants *How easily did you adjust to the control devices used to interact with the virtual environment? (Q2)*. The analysis did not indicate a significant difference between `with temperature` ($M = 4$, $MAD = 1.48$) and `without temperature` ($M = 4$, $MAD = 1.48$). Finally, we asked our participants *Was the information provided through different senses in the virtual environment (e.g., vision, hearing, touch) consistent? (Q3)*. Again, the analysis did not show a significant difference between `with temperature` ($M = 4$, $MAD = 0.74$) and `without temperature` ($M = 4$, $MAD = 1.48$).

10.26 Discussion

Through our two evaluations, we have shown the technical feasibility and the suitability of the approach to deliver more realistic VR experiences. In the following, we discuss the implications of our findings.

10.26.1 The Implementation of a temperature-enabled ETHD

We found from our technical tests that the system takes around 2.5 seconds to achieve a target point 40cm away from the initial point at the maximum recommended speeds for cobot interaction. Aligned with literature [466, 470], this highlights the relevance of implementing intent prediction strategies. In this work, we used a combination of gaze-based prediction and hand trajectory tracking to relocate the cobot end-effector preemptively.

With regard to thermal rendering, we propose two rendering strategies. First, setting all the Peltier elements simultaneously to the target temperature allows for a bigger rendering area, which is especially useful for multi-finger or full-hand surface palpation and also allows a more consistent temperature rendering across the surface. Yet, it requires more time to be ready for temperature rendering as it has to change the temperature of the whole thermal plate for each target temperature. This is especially critical at higher temperature differences and is further impaired when trying to reach low temperatures, which are generally more energy-intensive when using Peltier elements. The second rendering strategy exploits the multiple Peltier elements in the end effector to create a thermal gradient containing the lowest and highest temperatures. This rendering method provides the advantage of rapidly switching temperatures, given that the temperature is location-based within the end-effector plate, and the contact point can be altered using hand redirection. However, temperatures are not uniform through the plate, which means that sliding through the plate in the direction of the gradient will let the user know that the temperature is not uniform. We show how such a rendering strategy would work in section 10.23.

On the other hand, the non-uniform temperature rendered using gradient rendering can be beneficial for generating thermal affordances such as temperature-based sliders, where, on one side, the slider's value is cold, and, on the other, it is warm.

10.26.2 The integration of temperature feedback enhances the realism of haptic VR Experiences

From the user study, we found that the added thermal feedback influenced the participant's presence as defined by the IPQ questionnaire Realism factor but not in the Spatial Presence or Involvement factors. The ratings for these two factors were already positive compared to those for traditional ETHD. At the same time, realism had a more mixed rating from participants, suggesting that adding thermal feedback can support these ratings where ETHD does not perform well. On a perceptual level, this can be explained by sensory immersion; object properties include stiffness, temperature, and texture. While typical ETHD provides kinaesthetic and tactile sensory information, the proposed system adds additional stimulation that increases haptic immersion.

This contrasts the realism factor of the Haptic Experience questionnaire, which addresses the plausibility of the haptic feedback (**HX - Realism**) rather than the contrast of the virtual and real feedback (**IPQ - Real**). In this sense, thermal ETHD and typical ETHD had similar ratings in the **HX - Realism** with only indistinguishable higher ratings (higher is better) favoring thermal ETHD.

Regarding the Haptic Experience questionnaire, we found that the thermal ETHD significantly enhanced the haptic experience's Expressivity component (**HX - Expressivity**) in all the subscale items (notice that E1 is a reversed polarity item). This suggests the thermal ETHD was perceived as more dynamic by the experiment participants and better integrated with the events occurring in the virtual environment.

Regarding the Involvement factor (**HX - Involvement**), we found a significant impact of the thermal ETHD over the typical ETHD, with higher ratings in the item *I felt engaged with the system due to the haptic feedback (I2)* (higher is better), but not for the other item (H2) which may be due to the already positive ratings for both versions of the system. This might be partially a limitation of the questionnaire item, given that this item apparently saturates at some level of haptic immersion (we can call this a ceiling effect).

The final factor from the Haptic Experience questionnaire is the Harmony factor (**HX - Harmony**), which presented no significant differences nor ceiling effects, suggesting no improvement of thermal ETHDs over the overall haptic experience.

Regarding the custom questions, we found significant effects only in Q1: *How easy was it to identify objects through physical interaction, like touching an object or bumping into an object?* suggesting that thermal ETHDs do improve the identification of material properties given the higher haptic immersion, coherently with **IPQ-Real**.

Although typical ETHD has been shown to improve the overall VR and haptic experience, we found that adding thermal feedback can substantially improve ETHD, making it reasonable to consider it as a component to be integrated into future haptic interfaces.

11 TOOLS AND TECHNOLOGIES FOR AUGMENTATION SIMULATION

This chapter is based on the following publications:

Cobity: A Plug-And-Play Toolbox to Deliver Haptics in Virtual Reality - Steeven Villa and Sven Mayer - *In: Proceedings of Mensch und Computer Conference 2022 (MuC '22). Association for Computing Machinery.*

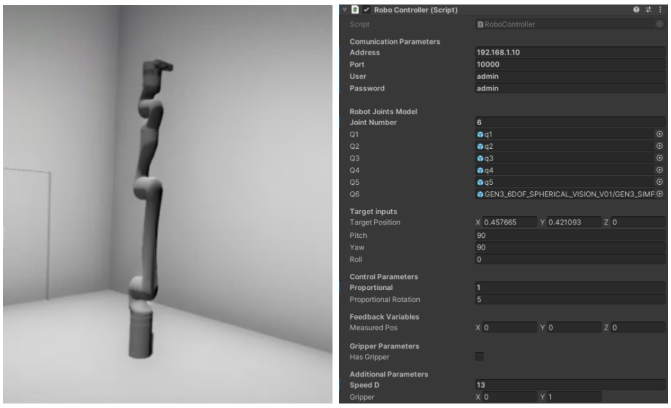
Towards a Haptic Taxonomy of Emotions: Exploring Vibrotactile Stimulation in the Dorsal Region - Steeven Villa, Thuy Duong Nguyen, Benjamin Tag, Tonja-Katrin Machulla, Albrecht Schmidt, Jasmin Niess - *In: Proceedings of the 2023 ACM International Symposium on Wearable Computers (ISWC '23). Association for Computing Machinery.*

11.1 Towards Plug-And-Play Robotics

Developing tools and prototypes for human augmentation research is fundamental, in this chapter I list two tools a unity plugin for controlling robots in VR and a vibrotactile haptic vest. We developed the Cobity plugin as a bridge to bring robotics closer to HCI researchers, therefore prioritizing easiness to use and setup. The plugin runs in the same operating system (Windows) and does not require network communication with auxiliary platforms. It does not need to launch additional software to establish the connection and communicate with the robot. In the following, we present the details of our implementation. On the other hand, we developed the haptic vest in order to enable sensory substitution in the dorsal area.

11.1.1 Architecture

To reduce computational costs in the host computer and the complexity of the communication setup, we aimed to provide direct communication from the graphic engine to the robot joints. However, such implementation raises a series of conflicts, such as the control frequency required to make the robot movement smooth and stable. A direct control loop over the robot joints involves a minimum of 1Khz to avoid instabilities and oscillations in the motor; however, a graphic engine such as unity typically runs at $\sim$30 to 120 fps. While this is frequently acceptable for graphics, it's not suitable for robot control. Therefore, we developed a dynamic library that manages this communication by running a velocity control over the 1KHz loop of



(a) Robot 3D Model (b) Plugin Interface

Figure 11.1: The two visual components of our plugin.

the robot, see Figure 11.2. This introduces a middle loop that sets targets to the High-frequency loop whenever it reads them from the physics loop in the graphic engine. The central loop reads positions in coordinates centered in the robot's base. The cartesian coordinates are converted into velocity vectors (Rotational and Translational), communicated to the kinematics loop, and finally applied to the joint motors. The feedback from the encoders is read from the encoders in the robot and then sent to the graphic engine to animate the cobot's model.

11.1.2 Control

We implemented a PD control [475] in the translation and rotation axis. Proportional-Derivative (PD) controls are widely used for high-level control in robotics [476, 477, 478, 479]. Similarly, we used a PD controller to create a velocity control based on cartesian coordinates given by Unity. The Kortex Library[1] internally computes inertia, Gravity and Coriolis.

Equation 11.1 and 11.2 describe the PD control based on Dorf and Bishop:

$$P_{x,y,z} = P_{x,y,z,bias} + K_p * e(t) + K_d * (e(t) - e_{-1}(t)) \tag{11.1}$$

$$R_{\theta,\psi,\phi} = R_{\theta,\psi,\phi,bias} + K_{pR} * e(t) + K_{dR} * (e(t) - e_{-1}(t)) \tag{11.2}$$

Where $P_{x,y,z}$ and $R_{\theta,\psi,\phi}$ represent the pose of the end effector (Cartesian position and rotations), $P_{x,y,z,bias}$ and $R_{\theta,\psi,\phi,bias}$ are the standard offsets in position and rotation, K_p, K_d and K_{pR}, K_{dR} are the controller components (PD) for position and rotation respectively and finally $e(t)$ represent the error between target position/rotation and current position/rotation.

[1] `https://github.com/Kinovarobotics/kortex`

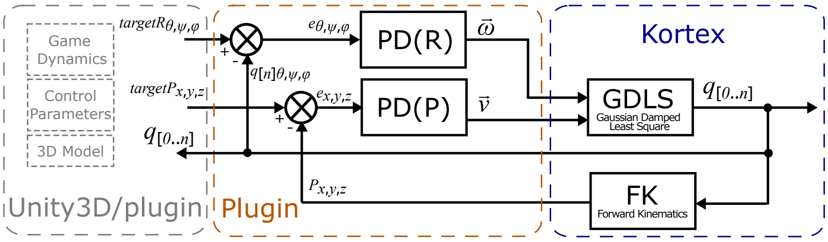


Figure 11.2: Plugin communication scheme, a C++ dynamic library manages the information exchange between the graphic engine and the control loop of the robot: The graphic engine manages the positions in the virtual environment, the target positions and measured feedback are sent and retrieved from the plugin. The plugin then communicates with the robot loops using the Kortex API to manipulate the robot's end effector towards the target position received from the virtual environment. Also, feedback from the joints and estimated end-effector position are sent back to the virtual environment.

We used them to interface the target position and rotation vectors and the robot end-effector position/rotations (joint $q_{[n]}$). Rotations and translations are handled separately, given the mechanical and interaction implications of each of them. While an overshoot in the robot's translation can easily lead to a collision with the user, an overshoot in rotation (in-place) of the end-effector will preserve the distance between the end-effector to the user's hand. The velocities (Cartesian and angular) are introduced in the Gaussian Damped Least Square inverse kinematics solver. This solver is based on Jacobian inversion and adds a gaussian damping factor to handle the behavior of the jacobian matrix near singularity configurations. The reader can find further information about this algorithm in Phuoc et al. [480]. Figure 11.2 illustrates this control loop in more detail.

The plugin sends cartesian ($\vec{v}$) and angular velocities ($\vec{\omega}$) and reads the robot's joint rotation angles to animate the 3D model in the scene.

11.1.3 Interface

Our plugin visual interface is divided into two main components: (1) the editor interface that works as GUI input to the robot and (2) the 3D representation of the robot that provides feedback about the robot's current pose.

Editor interface

The interface of our plugin (see Figure 11.1b) enables the user to set *communication parameters* such as IP address and login information (required to access the manipulator), as well as *Target inputs:* which are the goal coordinates and angles that the user wants the end effector to adopt, the pose is communicated in real-time to the robot. Access to these variables is also possible from external scripts by making a call to the script instance. *Control Parameters* are the inputs fields for K_p, K_d, K_{pR}, K_{dR}, these values can be used for online tuning of the robot behavior or for damping the speeds. The field *Measured Position* provides feedback of the

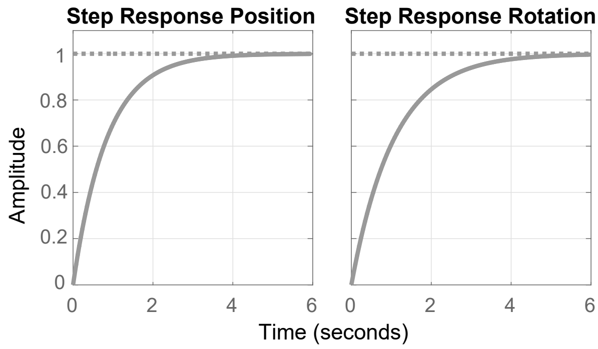


Figure 11.3: Step Responses to Rotational and translational degrees of freedom (Normalized response)

end-effector position in the robot base coordinates frame and is drawn using a gizmo in the Graphic engine 3d space.

Robot 3D representation

We render the robot's 3D model in the virtual environment using the values obtained from the cobot's encoders. Herefore, we use the values of the forward kinematics that are implicitly calculated by setting those rotations to the 3d subcomponents of the cobot (*gameobjects*). We also render the forward kinematics obtained from the Kortex solver as a gizmo in the debug window of the 3D engine. Our library assumes that the robot is in the center of the coordinate system, as this is not always the case; the control script of the robot harmonizes the mismatch in coordinate systems using the virtual robot 3D position. Therefore, the end-effector can be set to follow an object in the scene without requiring to provide relative coordinates of the robot, allowing a more straightforward game logic.

11.1.4 Technical Evaluation

To test the functionality and stability of our plugin, we ran a set of technical tests and showcased the capabilities of the system by developing two common use-cases in haptics for VR; encountered-type haptics, and object manipulation.

Our system has two primary design criteria: (1) Safety of use in shared environments with humans, and (2) speed performance to timely meet the user requirements. To facilitate the understanding of the dynamics of the system, we characterized the response of the cobot using Cobity, detailed values about the transfer function can be found in the plugin repository.

We executed an automated test of the robot speed in every translational axis. The robot was programmed to move from position A to B ($A - B = 40cm$). Then, we recorded the time it needed to reach the final position. Similarly, we evaluated the system's rotation axis; We rotated the end effector from an angle α_a to α_b ($\alpha_a - \alpha_b = 90°$). Figure 11.3 shows the step

(a) Encoutered Type Haptics (b) Encoutered Type Haptics Virtual Environment (c) Extended Mid-Air Haptics (d) Extended Mid-Air Haptics Virtual Environment

Figure 11.4: Four additional example use cases which we can envision to deliver haptic feedback in VR.

responses for the system without any PD control tuning. The standard response of the system is constrained in speed in order to meet the regulations for cobots (EN ISO 10218-2:2011).

11.1.5 Showcases

In the following, we present four showcases that highlight the potential for rapid prototyping of our plug-and-play solution. For each application, we connect the virtual environment with the cobot to deliver haptic feedback.

Encountered-type Haptics When using a robotic manipulator, encountered-type haptics (ETH) is probably one of the first use cases to imagine. We implemented a virtual environment to replicate a simple ETH scenario; the participant's hand can be tracked using VIVE Trackers or Leap motion (or the headset built-in hand tracking). A straightforward approach to enable ETH is to move the robot in the plane of the surface that is required to be rendered, constrained to the hand's movement. Using our plugin, it is necessary only to trace a line from the hand position to the surface to be rendered and move the robot according to the hand movement. Further improvements can be added, for example making a predictive control of the end effector to anticipate the hand's future position. Figure 11.4a and 11.4b depicts a user touching a flat surface that corresponds to a virtual door. A more realistic rendering of flat surfaces can be obtained using a round rotating surface as described in [467]. The latter approach also enables the rendering of different surface textures.

Mid-air Extended Haptics Ultrasound-based mid-air haptics uses sound waves generated by an array of transducers to render tactile sensations at the palm of the hand. A well-known constraint of mid-air haptics is the rendering workspace [452, 481, 482, 483], the usage of a cobot as a driver of the haptic array can help to overcome such limitations. State of the art approaches proposes to increase the number of arrays [482], attach the array to a rotary joint [452] or switch the positions of the array as required [483].

However, an online driving of the array in the 3D space using a manipulator emerges as a more robust and beneficial approach. We attached a haptic array to the end effector of our robot and guided the robot's movement using the position of the palm, given by a Leap Motion. Using the transform of the end effector (simulated thanks to the rotation of the joints read by the plugin), we transformed the coordinates of the hand from the leap motion coordinates to the world coordinates. Figure 11.4c and 11.4d illustrate a user interacting with a dynamic haptic array; the array is kept at a distance of ∼30cm of the palm to preserve a high rendering quality.

Social Touch Social VR has been increasing its presence in VR stores; apps like VR chat, RecRoom or PokerStars VR are becoming more popular. Social VR allows multiple users to join in a shared Virtual Environment and let the participants interact in a more natural way than 2D interactions. However, touch is still a missing component in this context. Social touch has been demonstrated to increase the perceived human likeliness in virtual agents [484]. However, this sensation depends highly on the kinaesthetic feedback provided by the human hand; therefore, although versatile, vibrotactile actuation does not create such perception. Alternatively, robotic manipulators with human hand alike end effectors can automate this task and enhance social VR environments.

Figure 11.5: The Social Touch example; we used Cobity to deliver human like touch sensations in virtual reality.

Figure 11.5 shows our implementation of social touch using a silicon human hand that features a heat-able foil. The prototype is driven by a cobot using the Cobity plugin. The participant interacts in a VR environment. Whenever the virtual avatar touches the user's shoulder, the robot moves the hand to their real shoulder. This setup could be further improved by using rigged hands as end effectors, for example, the Shadow Dexterous Hand[2].

[2] https://www.shadowrobot.com/dexterous-hand-series/

On Demand Tangible objects Tangible props are a common approach to introducing haptics in a VR scene, yet, the usage of tangibles requires a previous preparation of the physical scene to match the haptic-enabled VR objects, reducing the flexibility of this method. Recent advances have demonstrated that it is possible to alter the perception of Stiffness and friction of tangible prop [485, 486]. Furthermore, De Tynguy [487] explored the extent to which the virtual representation of tangible props can be altered without perceiving such mismatch. In addition to those approaches, the ability to switch the tangibles presented on the scene would significantly enhance the versatility of such a method. Mercado et al. presented a remarkable proof of concept addressing this use case.

11.1.6 Discussion

Cobot control remains a highly technical task, keeping apart HCI designers with non-engineering backgrounds. Current solutions require deploying extensive middle-ware and, in some cases, involve more than one operative system for simple prototyping. To facilitate introducing cobots in VR applications and reduce the time of experience prototyping, we introduced Cobity, a solution to use a virtual environment to control the cobot directly from Unity. We developed a real-time C++ library acting as a bridge between the cobot and the graphic engine. In this way, we transfer the 3D positions of the VR application onto real cobot motions.

We presented a dynamic library that enables the usage of Unity as rapid cobot experience prototyping. We run a technical evaluation using our new plugin. Finally, we presented a range of showcases that evidence the flexibility of cobot-based haptics, from kinaesthetic to ethereal sensations, including social touch. The goal of Cobity is to facilitate rapid prototyping for cobot usage in VR instead of replacing the standard architecture ROS. The purpose of such a plugin is to speed up the prototyping of applications within the HCI domain. In the bigger picture, this plugin will facilitate the mediation with different types of cobots used for HCI research.

Takeaway: Cobity is available at `https://github.com/xteeven/Cobity` and maintained by the Media Informatics Group at LMU Munich. New features and development will be added to this repository.

As of today, our system only considers the serial robot Kinova Gen3 (6DoF + 7Dof) and Gen3 Lite; we envision a compatibility enlargement to include other widely-used models as the Universal Robots line (UR3, UR5, and UR10), as this line of cobots is more common in HCI environments. The following stages of the plugin require implementing a simulation system to help designers have development speed even higher. Moreover, individual joint control is required for more complex scenarios, allowing a more comprehensive range of applications for our plugin, especially those that demand path planning or obstacle avoidance.

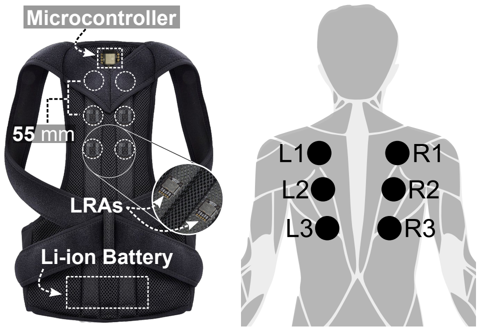


Figure 11.6: (Left) The vibrotactile vest uses a posture corrector as the base, an ESP32 FireBeetle micro-controller, six Linear Resonant Actuators (LRAs) and one Li-Ion Battery with a charging circuit. (Right) The LRAs' locations on the back.

11.2 A Haptic Vest for Sensory Substitution in the Dorsal Region

To investigate a haptic taxonomy of emotions, we developed a prototype consisting of a vibrotactile vest and a pattern creation interface, which was connected to the vest via Bluetooth Low Energy (BLE). In this section, we describe the specifics of this prototype.

The haptic vest was constructed by integrating a commercial posture corrector to ensure ergonomic contact with the user's back. Six Pimoroni DRV2605L breakouts were strategically positioned in the dorsal region inside the vest, maintaining an approximate spacing of 55 millimeters between each actuator as illustrated in Figure 11.6. Actuators were placed in pairs along the left and right scapular lines of the back: L1 and R1 in the suprascapular, L2 and R2 in the intrascapular, and L3 and R3 in the infrascapular region.

Each breakout integrated a Texas Instruments DRV2605L driver, which offered built-in auto-resonance control. Each breakout also featured an ELV1411A Linear Resonant Actuators with a resonant frequency of approximately 150 Hz and amplitude of $14.7m/s^2(1.5g@100g)$. Notably, the upper back region exhibits lower sensitivity to tactile stimulation compared to other areas [488, 489], requiring meticulous calibration of resonant frequencies, intensities, and spatial resolution for effective tactile rendering. Considering this low tactile sensitivity, we spaced the actuators slightly above the just noticeable difference (JND) thresholds for tactile stimulation on the back (at approximately 55 millimeters, see Figure 11.6), estimated to be between 45 and 50 millimeters [490]. This spacing enables both optimal discrimination of individual stimulations and the ability to interpolate stimulation between adjacent actuators.

The pattern creation interface was developed using Python and the Kivy UI framework[3]. This interface incorporated controls to enable the customization of tactile patterns. Specifically, it featured two sliders responsible for adjusting the continuous variables of AMPLITUDE and FREQUENCY, i.e., to increase or decrease their intensity. To govern the activation of the actuators, six buttons were arranged in a 2x3 layout, emulating the SPATIAL LOCATIONS of

[3] https://kivy.org/

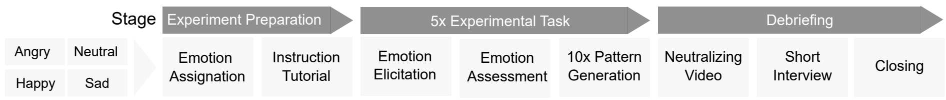


Figure 11.7: A between-subject experiment was conducted; participants were randomly assigned to one of four conditions. Participants classified the induced emotion and subsequently generated ten vibrotactile patterns.

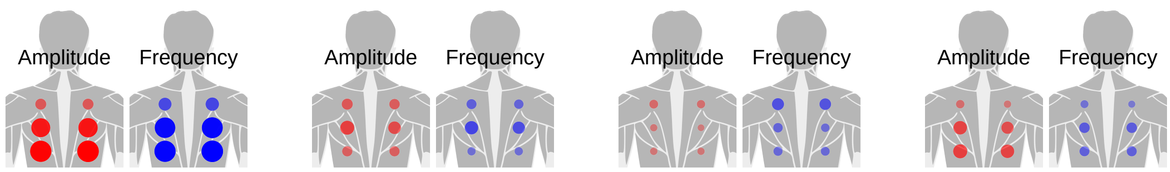
the haptic vest. Two extra buttons were implemented; one for rendering the selected pattern and the second for submitting the pattern and progressing to the next phase. The submit button remained inactive until the render button was pressed, ensuring participants were aware of the stimulation before submitting. Upon submitting the pattern, all input values were automatically reset to their default settings. Additionally, the interface served as a proxy for the experimenter by providing instruction screens, playing the emotion elicitation videos, and displaying the emotion assessment questions.

11.2.1 User Study

We conducted a between-subject study with one independent variable Assigned Emotion with four levels: ANGRY, HAPPY, NEUTRAL, and SAD, manipulated using video emotion elicitation (Figure 11.2.1). Three dependent variables that constitute the building blocks of the vibrotactile patterns were measured: AMPLITUDE, FREQUENCY, and SPATIAL LOCATION, while maintaining a constant duration of stimulation of 2 seconds. Participants were randomly assigned to one of the levels of Assigned Emotion and asked to develop ten vibrotactile patterns describing the elicited emotional state.

Measures First, we analysed *vibrotactile patterns* in detail. As a standard haptic technique, vibrotactile rendering allows a wide number of dimensions that can influence the perceptual attributes of haptic patterns [491, 492]. Yet, in this study, we focused on three key aspects [493]: AMPLITUDE, FREQUENCY, and SPATIAL LOCATION of the stimuli.

AMPLITUDE is the intensity of the stimulation applied to the user, ranging from $0m/s^2$ to $14.7m/s^2$. FREQUENCY, in this study, refers to the envelope frequency of the stimulation, ranging from $1Hz$ to $60Hz$, while the underlying frequency is set to the resonant frequency of the actuator ($150Hz$). Lastly, SPATIAL LOCATION refers to the regions on the dorsal region



(a) Angry (b) Happy (c) Neutral (d) Sad

Figure 11.8: Predominant SPATIAL LOCATION per Assigned Emotion in terms of AMPLITUDE and FREQUENCY. A bigger circle size represents a higher predominance in an specific location

where stimuli are delivered (see Figure 11.6 for visualization of the vest and the stimulated regions).

The Big Five questionnaire encompasses the personality traits Conscientiousness, Agreeableness, Neuroticism, Openness to Experience, and Extraversion. These traits can be linked to the experience of positive and negative emotions. We used this measure to perform a group allocation check. This screening ensured that the distribution of participants across conditions was balanced for their personality traits, thus reinforcing the validity of the findings.

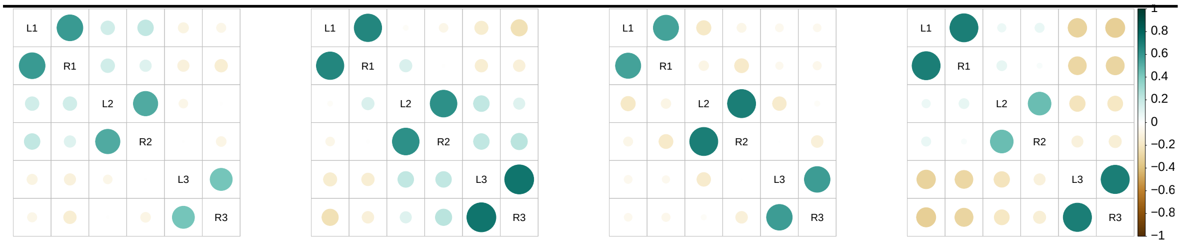
To measure *emotion classification*, we quantified the effectiveness of the emotion elicitation process by comparing the Classified Emotion and the Assigned Emotion. After every emotion elicitation video, we asked participants:"How did this video make you feel?", with four possible answers: Angry, Happy, Neutral, and, Sad.

The Self-Assessment Manikin (SAM) scale [494] measures Arousal, Valence, and Dominance. In the current study we collected SAM data, however it was not analyzed.

Emotion Elicitation Empirical evidence suggests that emotions play a role in people's decision-making processes [76]. In fact, some researchers argue that decision-making would be impossible or suboptimal without emotional involvement [76]. In line with this perspective, Damasio's Somatic Marker Hypothesis (SMH), suggests that emotions implicitly bias human behavior. Based on this theoretical framework, we aimed to induce an emotionally congruent state in participants prior to each pattern creation stage, i.e., aligning their current emotional experience with the intended category of patterns to be created. In other words, we wanted to increase the likelihood of generating patterns that reflect the target emotional qualities. We used the video emotion elicitation method over alternatives, as it has been shown to have a comparatively superior performance [495, 496, 497]. We collected videos from different datasets and selected five per emotion [498, 499]. The videos reported in these datasets are mostly scenes from movies (see Jurásová and Spajdel, Gilman et al. for details).

Participants We recruited 41 participants using the university's mailing list. One was excluded due to a technical problem. From the 40 remaining participants, 24 self-identified as male, 15 as female, and 1 as non-binary. The mean age of our participants was 23 years old (M = 23.77, SD = 2.91). They were compensated with 10 Euros/hour for their participation.

Task and Procedure Participants were informed of the experiment's details, risks, and benefits, and asked for consent. Emotions were randomly assigned without the participants' knowledge. The prototype and interface were explained and a tutorial round of pattern generation was conducted. Then the experimenter left the room and communication was limited to text to avoid influencing emotions. Participants performed the experimental task five times. They watched an emotion elicitation video (Figure 11.2.1) and assessed the induced emotion using the SAM scale and emotion classification. They created 10 vibrotactile patterns using combinations of: (1) Amplitude, (2) Frequency, and (3) Spatial Location and asked to render each pattern before submitting it. Once a set of 10 patterns were submitted,



(a) Angry (b) Happy (c) Neutral (d) Sad

Figure 11.9: Correlation matrices among SPATIAL LOCATIONS: Symmetrical patterns observed across all measured emotions. A value closer to 1 represents possitive correlations, while a value closer to -1 represents a negative correlation

the next elicitation video would play, participants were allowed to submit the same pattern multiple times. This process was repeated 5 times using distinct videos eliciting the same emotion, resulting in a total of 50 patterns per participant. At the end, a neutralizing video was shown to counter emotional carry-over effects. The experimenter returned, conducted a brief interview, and concluded the study (refer to Figure 11.7 for a graphical overview).

11.2.2 Findings and Discussion

This section presents an analysis and discussion of the key features exhibited in the generated vibrotactile patterns. An examination of the pattern profiles AMPLITUDE and FREQUENCY profiles is followed by an investigation of the influence of SPATIAL LOCATION. Lastly, we explore the significance of each feature in effectively describing the Assigned Emotion.

First, we assess the data's validity by conducting a group allocation and emotion elicitation check.

To address potential group allocation imbalances, we conducted an analysis of variance (ANOVA) on the factor Assigned Emotion for each subscale of the Big Five questionnaire. The results indicated no significant differences.

Further, participants were asked to classify the emotion induced by the elicitation video, in the following sections, we call this classification made by the participants "Classified Emotion". We compared this Classified Emotion to the Assigned Emotion. The results revealed that participants accurately classified the emotion in 68.23% of cases (chance level 25%). Further

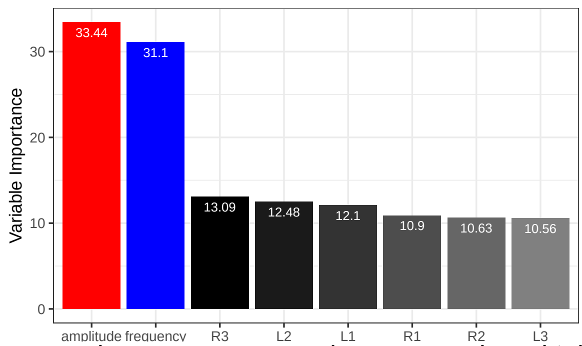


Figure 11.10: Feature relevance chart: AMPLITUDE and FREQUENCY have higher contribution to pattern classification.

computations were exclusively performed using data from instances where the emotions were correctly classified.

Amplitude and Frequency To study the impact of the Assigned Emotion on AMPLITUDE and FREQUENCY, we computed a Multivariate analysis of variance (MANOVA) using Assigned Emotion as a predictor. The MANOVA yielded a significant overall effect of Assigned Emotion on AMPLITUDE and FREQUENCY ($F(6, 2312) = 47.704, p < 0.005$). We then performed one Tukey Honest Significant Differences (HSD) per variable.

For AMPLITUDE we found significant differences for all pairs of Assigned Emotion. Specifically, when compared to the ANGRY emotion, the HAPPY emotion showed a lower amplitude (diff = $-2.955 m/s^2$, $p < 0.001$), this was also true for the NEUTRAL (diff = $-5.202 m/s^2$, $p < 0.001$), and the SAD emotion (diff = $-2.002 m/s^2$, $p < 0.001$). Likewise, the NEUTRAL emotion showed a lower amplitude compared to the HAPPY emotion (diff = $-2.246 m/s^2$, $p < 0.001$), and the SAD emotion showed a higher amplitude compared to the HAPPY emotion (diff = $0.952 m/s^2$, $p = 0.004$). Finally, the SAD emotion showed a higher amplitude compared to the NEUTRAL emotion (diff = $3.199 m/s^2$, $p < 0.001$).

For FREQUENCY we found significant differences for the HAPPY, NEUTRAL, and SAD emotions compared to the ANGRY emotion; In detail, the HAPPY (diff = $-9.869 Hz, p < 0.001$), NEUTRAL (diff = $-8.672 Hz, p < 0.001$), and SAD emotions showed a lower frequency (diff = $-8.370 Hz, p < 0.001$) compared to the ANGRY emotion. However, no significant differences were observed between the NEUTRAL and HAPPY emotions (diff = $1.197 Hz, p = 0.722$), the SAD and HAPPY emotions (diff = $1.499 Hz, p = 0.505$), or the SAD and NEUTRAL emotions (diff = $0.302 Hz, p = 0.993$).

The findings indicate that AMPLITUDE is the primary property with greater descriptive power for emotion communication. FREQUENCY, in constrast, exhibits descriptive capabilities for a variety of emotions, although some emotions exhibit frequencies that are not distinct enough for effective classification. ***Consequently, we recommend that designers prioritize maintaining consistent intensity (Amplitude) between the desired emotion to be communicated and the suggested ranges outlined in this section.***

Locations To evaluate the impact of Assigned Emotion on SPATIAL LOCATION, we computed a weighed average calculation across all locations. The results are displayed in Figure 11.8 where it can be observed that Anger had a higher intensity than the other emotions, mostly in the intra and infrascapular region. Similar locations were relevant for sadness, but with a lower intensity. Happiness was mostly assigned to the interscapular region with lower intensities in the supra and infrascapular regions, and Neutral to the suprascapular region for amplitude, with a lower frequency intensity in the same regions. To further understand the correlation among SPATIAL LOCATION and Assigned Emotion We computed correlation matrices for each level of Assigned Emotion. The result show a correspondence between scapular regions. This is between L1 & R1 (suprascapular), L2 & R2 (intrascapular), and, L3

& R3 (infrascapular). Suggesting that participants preferred symmetrical over unbalanced patterns.

The findings indicate that specific locations can also contribute to the descriptive nature of the intended communicated emotion, predominantly in terms of vertical rather than horizontal location. This has implications for the quantity and placement of actuators in prototypes. ***Based on these results, we recommend that designers exploit the descriptive capacity of the back's location, particularly emphasizing the infra and suprascapular regions, which exhibit the highest potential to discern between emotions among all the locations.***

Feature Importance To determine the relevance of each feature, we perform a recursive elimination of features for variable importance [502]. Results are displayed in Figure 11.10. AMPLITUDE and FREQUENCY were identified as the primary discriminating properties for distinguishing between emotions based on the generated patterns, followed by the SPATIAL LOCATION. It is noteworthy that the less descriptive features are associated with complementary locations that are categorized as more descriptive, indicating redundancy due to high correlations between actuators within the same region.

Therefore, we recommend that designers prioritize Amplitude and Frequency as the primary factors and utilize the available locations whenever feasible to enhance the clarity of the communicated emotion.

Reflections on Part IV

This part explored advancements in haptic augmentation, motor learning, and virtual reality (VR) interfaces, providing empirical evidence on how emerging technologies influence skill acquisition, perception, and user interaction. Across multiple studies, we examined different feedback mechanisms, their impact on learning and perception, and their implications for HCI design.

We first investigated Electrical Muscle Stimulation (EMS) for motor learning, comparing it with electrotactile feedback and no-feedback conditions. Our findings confirmed that EMS not only enhances immediate motor performance but also facilitates long-term motor learning, contradicting traditional assumptions that learning requires conscious correction. However, electrotactile feedback demonstrated a higher learning rate, suggesting that different feedback mechanisms may be more suitable depending on the learning goal. We outlined best practices for using EMS in skill acquisition and emphasized the role of agency in EMS-driven learning, proposing future research directions in physiological mechanisms and optimal intervention strategies.

We investigated virtual augmentation in VR, highlighting its potential for immersive experiences while emphasizing the need for natural, goal-driven interactions beyond novelty. Motor-focused augmentations face technical limitations, suggesting a shift toward enhancing perception rather than action. Additionally, we explored ultrasound haptics (UMH) and temperature-enabled haptics (ETHD), demonstrating their role in improving multisensory realism. UMH enhances object congruency and temperature perception, particularly for liquid and gaseous simulations, while ETHD increases realism by integrating thermal cues. We also introduced a moving-array UMH approach to extend haptic feedback across larger VR environments, balancing rendering quality with safety. These findings underscore the importance of multi-modal haptics in enhancing immersion beyond traditional force feedback.

Finally, we developed Cobity, a real-time Unity-based framework for controlling collaborative robots (cobots) in HCI research. By simplifying cobot programming and enabling rapid prototyping within VR environments, Cobity lowers the barrier for HCI designers without extensive robotics expertise. This contributes to making cobots more accessible for haptic research, immersive VR experiences, and interaction design applications.

V

Ends and Means

12 DISCUSSION

> *"Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end."*
>
> **– Immanuel Kant**

Human augmentation changes how we think, act, and interact. And, as these augmentations move from labs into everyday life, their influence extends beyond usability, shaping autonomy, decision-making, and social norms. This thesis examined human augmentation through two critical lenses: autonomy—how these technologies alter individual agency and behavior—and heteronomy—the external factors that shape the experiences of augmented individuals.

Throughout this thesis, I investigated three research questions. First, how do augmentation technologies influence individual behavior and decision-making? I approached this question from a behavioral and psychological perspective, focusing on decision-making in the context of risk-taking, autonomy, and agency. Second, what are societal perceptions of augmented humans, and how do these perceptions affect social acceptance and integration? I examined this through a sociological lens, using a cross-cultural study to develop a standardized measurement of attitudes toward human augmentation. Third, how can humans experience and interact with augmentation technologies to explore their potential and limitations? I addressed this question by comparing augmentation with learning processes and proposing VR technologies to enable the investigation of Augmented Human interaction.

In this section, I revisit these three main research questions: first, I address **RQ1** on autonomy, elaborating on how augmentation technologies shape individual agency and influence decision-making and agency within the context explored in Part I. Then, I discuss **RQ2** on heteronomy, discussing how societal perceptions and external structures affect augmented individuals. Next, I analyze **RQ3** on experience, exploring multiple facets of being augmented. I then examine the tensions that emerge across these discussions before expanding on autonomy and heteronomy as lenses for understanding human augmentation, identifying factors that influence both or shift between them. Next, I elaborate on the implications of this work, using the implications taxonomy proposed by Van Berkel and Hornbæk [503], with a focus on methodology and HCI, design, and society. Finally, I discuss the limitations of this work and directions for future research.

12.1 How Do Augmentation Technologies Shape Human Behavior and Decision-Making?

The mere presence of an augmentation—regardless of its actual functionality—affects decision-making. In chapter 3, when we used a sham augmentation, participants believed the system was supporting them, even though it provided no actual functionality. This belief influenced risk-taking behavior, particularly among those with elevated expectations of the system's capabilities. Participants who assumed the augmentation was functional tended to take greater risks compared to those with lower expectations. This behavioral shift was also reflected in augmentation technology at a physiological level, where the P300 component—typically associated with decision-making, specifically self-relevance, showed a reduction when participants anticipated augmentation support. **Thus, technologies, as well as expectations, alter human behavior; in this work, we found this to be valid for risk-taking and that this is measurable by the intensity of the ERP P300 response in the central electrodes.**

Beyond risk-taking, the presence of augmentation also impacted the sense of agency. We analyzed the alpha and beta bands in the EEG data, which are commonly linked to agency in motor augmentation contexts. The results showed that agency-related neural correlates were lower when participants were informed that the augmentation was active. Given this, users could be attributing part of their actions to the system itself, potentially diminishing their perceived control over the task. **Therefore, the agency could also be compromised when participants are informed that AI is augmenting them, and this is measurable with correlates in the Alpha and Beta bands of the EEG.**

Beyond non-functional augmentations, functional cognitive augmentation systems, such as AI-enabled smart glasses for contextual suggestions, also impact the sense of agency. Our findings indicate that some individuals perceive themselves as mere instruments of the augmentation system, mainly when the system provides direct guidance or suggestions. This effect was explored in a conversational setting, where participants engaged in dialogues while receiving AI-generated suggestions. **Participants showed greater reliance on the augmentation in situations where they had less prior knowledge about the topic. While this improved engagement and fluency, it also led to a loss of agency, as individuals attributed their conversational output more to the system than to themselves.**

At a behavioral level, the modality of information presentation (visual vs. auditory) influenced how participants structured their speech and the degree of independence from AI-suggested content. For example, individuals with visual AI assistance presented longer discourses but also followed the textual recommendation more closely, while auditory AI suggestions led to a lower response similarity and length. **Thus, these findings show that AI support modality in AR-augmented conversation shapes an individual's discourse, especially when the topic of the conversation is not familiar to the individual.**

Additionally, the way suggestions were triggered played a role in perceived agency. Participants exhibited higher agency when they had to explicitly trigger the augmentation system through deliberate mechanical action, such as pressing a button, compared to implicit activations, where the system provided suggestions based on, for example, gaze. This aligns with motor action-consequence associations [253], where explicit control reinforces a sense of agency over decisions.

Yet, across the autonomy Part, our findings showed that expectations of functionality play a critical role in shaping human behavior and decision-making when using augmentation technologies. **Beliefs about an augmentation's capabilities influence risk-taking, reliance, and perceived agency, even before actual interaction.** In the iExpect scale, we studied this phenomenon further and found that pre-use expectations align with the core dimensions of the Technology Acceptance Model (TAM), especially *Usefulness* and *Ease of Use—but* also introduce an affective dimension: *anticipated enjoyment*. **This, aligned with modern HCI models, implies that users do not merely assess augmentation technologies based on functionality but also on the emotional and experiential value they expect to derive from them. These expectations, in turn, shape how users engage with the technology, how much they rely on it, and the extent to which they attribute agency to themselves versus the system.**

12.2 How Are Augmented Humans Perceived, and What Affects Their Social Acceptance?

The way augmented humans are perceived extends beyond their *actual* autonomy to how society *interprets* that autonomy. As we found across studies, these perceptions shape how augmented individuals are judged and integrated into society. Yet, there are more nuances to this; in Part II, **we identified six key factors that influence societal attitudes toward augmented humans: Peril, Privacy, Access, Motivation, Ownership, and Achievement.** Additionally, we observed a recurring personal dimension related to an individual's own preference for augmentation. We presume that these factors are not static but are shaped by broader societal contexts, with differences emerging based on cultural norms, technological familiarity, and ethical perspectives.

One of the most significant concerns shaping attitudes toward augmented humans is the potential risk they pose. This risk is not always clearly attributed, **some view the individual as the source of danger, while others see the augmentation technology itself as the risk factor.** This ambiguity complicates societal perceptions, especially in cases where augmentations enhance physical or cognitive abilities beyond natural human limits. Augmented humans are sometimes compared to enhanced weapons, raising fears about their potential to cause harm in ways that non-augmented individuals cannot.

Privacy concerns shape how augmented humans are perceived. **Many fear that augmentation could be used for surveillance, manipulation, or heteronomous control, introducing ethical considerations about who governs an individual's actions when augmented.** This extends to both the augmented individual and those interacting with them—whether people with augmentations have unfair advantages or whether external entities (governments, corporations, or bad actors) could exploit them. However, some acknowledge the benefits of augmentation, particularly in improving social understanding, such as emotion recognition or enhanced communication abilities.

Beyond privacy and risk, social acceptance is also influenced by why someone seeks augmentation. People who seek augmentation for medical or assistive purposes tend to be perceived more favorably than those who seek it for performance enhancement, status, or competitive advantage. This distinction influences trust; augmentations seen as necessary for restoring function are generally accepted, while those perceived as optional enhancements generate skepticism and, in some cases, social resistance.

Yet, there is also an element of envy in perceptions of augmented humans. Some non-augmented individuals, even while voicing concerns, expressed jealousy toward those with access to enhancements, revealing an underlying social tension between fairness and personal aspiration. **Augmentation is not only a technological issue, it is a status-defining factor that can reinforce social hierarchies and disparities.**

Social acceptance also depends on whether an augmented individual is seen as being in control of their enhancement or if the technology itself dictates their abilities. The degree of augmentation, its visibility, and its level of automation play a role in shaping whether people see augmented humans as autonomous individuals or as beings partially controlled by external forces. **When agency over one's augmentation is perceived as diminished, it raises concerns about authenticity, responsibility, and fairness** in interactions with others.

I speculate that this could lead to stigmatization, which can vary across cultural contexts. In some societies, augmented humans may face exclusion or discrimination, particularly in settings like healthcare, employment, and social relationships. This stigma may push some individuals to hide their augmentations, yet one may argue that transparency is necessary for ethical and social reasons too. **The debate around visibility versus hiding reveals another core factor in social acceptance: whether augmentation is seen as a personal choice, a social obligation, or an unfair advantage.**

Another challenge to social acceptance is the legitimacy of success among augmented individuals. If an augmented person excels in athletics, academics, or professional fields, is their achievement recognized on the same terms as non-augmented individuals? This question mirrors historical debates on performance-enhancing substances in sports and cognitive enhancers in academia. **Augmented success raises concerns about equity and merit, influencing whether augmented individuals are seen as rightfully accomplished or unfairly advantaged.** An interesting case to make a parallel about this factor is the movement *Enhanced Games*[1], which advocates for the validity of using human enhancement in athletics and other disciplines to showcase the maximum performance humanly possible in

[1]https://www.enhanced.com/science-is-real

these areas.

An interplay of concerns about safety, privacy, fairness, and agency shapes the acceptance of augmented humans. These perceptions are not uniform but vary across cultural, economic, and technological landscapes, highlighting the need for context-sensitive approaches to developing human-centered augmentation technologies.

As augmentation technologies develop, their social acceptance will depend on how designers address these concerns in their designs and how much autonomy the user can (1) actually have and (2) be perceived to have. Addressing these concerns requires designers to balance fairness, and public trust, ensuring that augmentation aligns with social values and ethical principles.

12.3 How Can Humans Engage with Augmentation Technologies and Assess Their Potential?

Human interaction with augmentation technologies is shaped by how these systems extend capabilities, how users adapt to them, and whether they enable temporary or lasting changes. These interactions are not solely defined by individual intent but are also influenced by the design of the augmentation, the expectations it creates, and the constraints it imposes. While some technologies provide a seamless extension of human abilities, others introduce dependencies or reshape how users engage with their own skills and decision-making. To explore these dynamics, this thesis focuses on two critical aspects: **sensory experiences in VR** and the distinction between **augmentation** and **learning** for human augmentation.

12.3.1 Augmentation vs. Learning: When Does Augmentation Become an Acquired Skill?

An emergent challenge in designing augmentation experiences is understanding when an augmentation merely enhances performance temporarily and when it leads to long-term knowledge consolidation. In motor learning, for example, studies claim to support learning but primarily measure short-term improvements within a single session [404]. These effects often indicate augmentation rather than actual learning, as performance gains disappear once the augmentation is removed [23]. A similar pattern emerges in cognitive and sensory augmentation, where users show rapid adaptation while assisted by a device but struggle to maintain the same performance without it.

Recognizing this distinction is important for developing augmentation systems that support skill acquisition rather than dependence. **Temporary augmentation is helpful in contexts where situational enhancement is needed for functional or hedonic reasons, but if the goal is for users to retain new capabilities, systems may be designed to encourage and facilitate learning beyond the period of augmentation use.**

To explore how augmentation can support both temporary enhancement and lasting skill acquisition, I investigated EMS as an augmentation tool in motor learning. Initially, I hypothesized that EMS might interfere with learning by altering the sensorimotor representation of a task, making it harder for participants to internalize movement patterns [Exp1]. However, our multi-session study revealed that EMS not only improved performance while active but also contributed to skill retention after the system was removed.

This finding suggests that **augmentation and learning can coexist—certain systems can both augment performance and facilitate long-term skill development.** Understanding how augmentation interacts with sensorimotor and cognitive adaptation is required for designing systems that empower users rather than create reliance.

For augmentation to be meaningful, it must be designed with an understanding of whether it serves as a temporary extension of human ability or a pathway to skill acquisition. Some augmentations will always be situational, enhancing abilities only while in use, others can be de-skilling while others may serve as bridges to learning.

12.3.2 Virtual Reality as a Platform for Prototyping Human Augmentation

Understanding the potential of augmentation technologies requires direct interaction, experimentation, and iteration, yet technical complexity, material limitations, and prototyping costs often constrain the development of physical augmentations [432, 436]. To overcome these barriers, VR has emerged as a platform for exploring augmentation concepts, allowing users to experience, evaluate, and refine augmentation interactions before physical implementation. In the past five years, research has increasingly investigated the validity of VR for prototyping human augmentation, offering a controlled environment where users can engage with augmentations that may not yet be feasible in reality.

This approach was further explored in a semester-long course on human augmentation development, reported in Part IV of this thesis. One of the key challenges that emerged from this work was the absence of haptic feedback in VR [Exp3], a limitation that affects how users perceive and engage with virtual augmentations. While this challenge applies to VR broadly, it is particularly relevant for human augmentation research, where physical interaction is central to understanding an augmentation's impact, for example, as tactile feedback is often used in sensory substitution to map different senses, or concepts [Exp8].

Rather than addressing this limitation by adding wearable haptic devices or handheld controllers, we explored an alternative approach, leveraging interaction space where the VR experience takes place to provide tactile cues. Minimizing additional devices that users must wear. Throughout multiple projects, I investigated how VR-based augmentation prototypes could effectively simulate augmentation experiences despite these constraints and what design principles could enhance their realism and applicability. Through the works presented in Part IV I proposed multiple methods to provide (1) tactile [Exp4], (2) thermal, and (2) kinaesthetic haptic feedback [Exp6] in VR in a large space [Exp5] without requiring the user to wear any new device rather than the typical VR equipment, enable to stimulate and prototype tactile sensations for human augmentation and VR in general.

12.3.3 Creating Tools for Human Augmentation

As I will discuss further at the end of this section, advancing Human Augmentation as a discipline requires robust tools and methods for prototyping, evaluation, and research. So far, I have introduced two methodological tools developed in this thesis [Het2, Aut4]. However, beyond methodologies, there is also a need for dedicated design and prototyping tools (see [436] for ideation tools).

In the final part of Part IV, I presented two tools for prototyping human augmentation. The first is a vest designed to enable both sensory substitution and sensory extension [Exp8]. The second, introduced in [Exp7], includes the primary tools I used in [Exp5, Exp6] to implement robot-based encountered-type haptics for prototyping sensory-rich virtual reality applications.

With these methodological and technical contributions, I aim to provide a foundation that facilitates faster and more effective research in Human Augmentation. Cobity [Exp7] bridges the gap between robotics and VR design by allowing interaction between kinova cobots and Unity, making haptic prototyping more accessible to researchers and designers. Meanwhile, the vibrotactile vest [Exp8] enables studies on how different vibration patterns can convey emotions, providing insights into non-verbal communication through haptic augmentation. These tools contribute both methodologically and technically, supporting the development of augmentation technologies.

12.4 Integration and Control as Foundations of Autonomy in Human Augmentation

As previously defined in this thesis, autonomy refers to an augmented individual's ability to achieve self-imposed goals. However, within the context of human augmentation, autonomy also implies having full control over the augmentation in use. This includes understanding its capabilities and limitations, ensuring it responds to the user's intentions, and maintaining a strong sense of agency over it.

This connects to Don Norman's concepts of the *Gulf of Evaluation* and the *Gulf of Execution* [504], which are especially relevant in human augmentation. The way individuals perceive, interact with, and modify their environment depends on how well their mental model of an augmentation aligns with its actual functionality. A significant mismatch can be problematic, particularly given the potential for increased risk-taking or altered behavior. If an augmentation fails to perform as expected in a critical moment, the consequences can be severe.

To mitigate this, designers and manufacturers must consider shaping user expectations to match real-world functionality, ensuring clear documentation, and, in high-risk cases, even requiring training or certification. This awareness of an augmentation's functionality is essential for preserving autonomy, aligning with Kant's premise that true autonomy requires rational understanding. Consequently, factors such as placebo effects or expectation mismatches can undermine autonomy just as much as technical failures.

A potential ideal augmented human would be one that has interiorized the augmentation capabilities and limitations, the one that can translate their own intentions to actions through the system without explicit instructions, *becoming one with the augmentation* as one would ideally say.

Autonomy can also be understood as the ability to define one's own identity, including the intentional integration of technology into that identity. An individual may set the goal of incorporating an external body part, a technological artifact, or an augmentation as part of their sense of self. Similarly, an augmented human may voluntarily seek integration into society, embracing heteronomy to fit in, adopting technology for social acceptance or aesthetic appeal, or even delegating tasks to technology as a matter of convenience. In some cases, this delegation extends to bodily control, whether for specific functions or purely for play—an idea extensively explored by Patibanda.

From a heteronomous perspective, I have framed this term through a Kantian lens, which has been supported by various authors within and beyond the HCI field as a robust alternative to other interpretations of autonomy and heteronomy. In this thesis, heteronomy refers to external factors that influence an individual's autonomy, with society identified as the primary heteronomous force affecting augmented humans. However, heteronomy is not limited to societal influence; it can also stem from the augmentation itself, particularly when it is not fully integrated or when designers attribute agency to the system. This is evident in frameworks such as human-computer symbiosis and human-computer integration.

12.4.1 Conceptualizing Human-Machine Relationships Through Agency and Heteronomy

When an augmentation is assigned agency—or even autonomy—it can shape human decision-making and self-imposed goals. Looking toward cyborgs, the question emerges: which framework best describes this relationship? Should we conceptualize it through the lens of *human augmentation*, *integration*, or *symbiosis*? This remains an open question for future research.

At the risk of oversimplification, I propose an "agency-heteronomy continuum" to illustrate the distinctions among these frameworks (see Figure 12.1). On one end, *human augmentation* prioritizes individual autonomy—humans lead, define goals, and use technology as an extension of their agency. At the opposite end, *full automation* represents systems that operate independently, without human intervention. Between these extremes, *human-computer integration* and *symbiosis* occupy an intermediary space, reflecting varying degrees of shared agency.

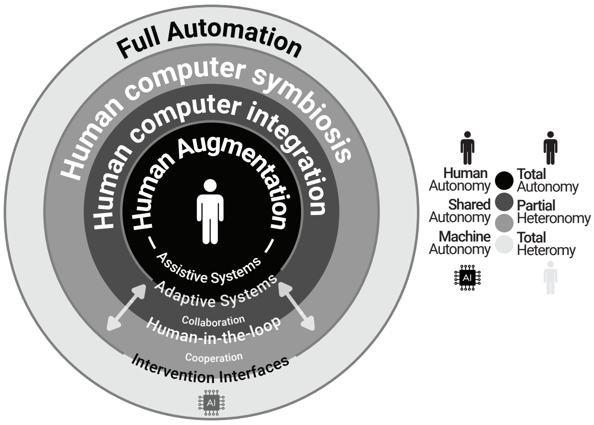


Figure 12.1: Layers of autonomy and their connection with human-computer relationships in HCI: The inner circles represent frameworks where the human has (or aims) Total autonomy; the more external the circle, the less human autonomy, until the human has no autonomy and all the autonomy is in the computer

Several HCI concepts, such as *intervention interfaces* [505], *human-in-the-loop systems*, and *human-computer collaboration* and *cooperation*, also define human-computer relationships. However, they differ fundamentally from *augmentation*, *symbiosis*, and *integration*: rather than focusing on the continuous interplay between human and machine as a joint entity, they primarily structure task-oriented interactions. Nevertheless, for completeness, I also situate them within this continuum.

At the farthest end, full automation operates without human oversight. However, the concept of intervention interfaces, as coined by Schmidt and Herrmann [505], introduces systems that retain high autonomy but allow selective human intervention when necessary. This marks the threshold where the human starts becoming part of the process, though still within a domain where machine autonomy dominates. At this same boundary, *human-in-the-loop* frameworks begin to apply [506], spanning the continuum until the point where the system no longer holds primary control.

From this boundary, Licklider [45]'s vision of symbiosis takes shape—a stage where humans and machines function as complementary entities, each relying on the other's strengths. *Human-computer collaboration* aligns with this layer, as both contribute independently to a shared process. Progressing inward, cooperation emerges as a more tightly integrated interaction, where humans and machines operate in closer coordination with more seamless communication. At this point, *symbiosis* transitions into *integration*, a concept developed further by Mueller et al. [46], where the system becomes deeply interwoven with human cognition and action.

As autonomy shifts toward the human, the continuum reaches adaptive systems—technologies that respond dynamically to human states and physiological signals based on self-defined goals. In the innermost ring, two concepts take precedence: *human augmentation*, where technology amplifies or extends human capabilities, and *assistive systems*, which provide human autonomy by compensating for impairments through ability substitution.

An interesting parallel exists with [507], which also operationalizes human-computer relationships in terms of autonomy and heteronomy. However, its perspective is inverted, centering on the machine and measuring autonomy as a function of decreasing human intervention, rather than considering the interplay of agency between human and system.

12.5 Emerging Tensions

The integration of augmentation technologies into human life introduces fundamental tensions. *Disclosure vs. concealment* questions whether augmentation should be visible or remain unnoticed. *Transparency and merit* affect how achievements are recognized when technology is involved. *Autonomy boundaries* must be reassessed as augmentation shifts control between individuals and systems. *Rational and non-rational autonomy* challenges design, requiring consideration of both deliberate choices and human impulses. Finally, *instrumentalization* raises the risk of treating augmented individuals as tools rather than autonomous agents. These tensions call for clear design decisions and ethical accountability.

We identified an emerging tension between disclosing and concealing augmentation—whether to make it visible to others or keep it unnoticed. Some segments of society view disclosure as important, creating friction with the concept of ubiquitous computing, where technology is seamlessly integrated into everyday life without explicit visibility. However, when it comes to the human body, this principle of "computers everywhere, yet unseen" does not seem to hold. Perhaps Mark Weiser did not anticipate computers being embedded within human bodies or underestimated the complexities of social acceptance when technology becomes part of the self.

This issue extends beyond transparency and location; it also involves transparency and merit. Human achievements are typically attributed to individuals, but when technology plays a role, society may assess these accomplishments differently, as explored in Part III. If there is no clear distinction between human effort and technological assistance, achievements may be more readily attributed to the individual. Conversely, visibly using augmentation can shift attribution toward the technology. Yet, the actual contributions of humans and machines to a given outcome remain inherently intertwined, regardless of visibility. This misattribution effect, therefore, represents a potential emerging bias in the use of human augmentation.

Another emerging tension for augmented humans revisits a long-standing philosophical debate: where does one individual's autonomy begin, and where does another's end? While *social contracts* mediate this balance among non-augmented individuals, the introduction of technology disrupts these boundaries. Designers are confronted with a fundamental dilemma: should they prioritize the autonomy of the augmented user or that of others? Any choice, or even the refusal to choose, imposes constraints, either limiting the user's control over their augmentation or restricting the autonomy of those around them. This creates an unavoidable contradiction: any attempt to safeguard autonomy for one party risks undermining it for another.

I agree with the interpretation that autonomy is grounded in rationality and that only rationality can enable true autonomy. However, I also recognize the irrational aspects of human behavior and the need to account for them in the design of human augmentation technologies. In this sense, the pursuit of autonomy is aspirational rather than absolute.

Kant argued that autonomy is dictated by the rational self, that individuals act according to rational goals. Dostoevsky, in *Notes from Underground*, critiques this view through the Underground Man—an individual who believes the world can be navigated through pure rationality yet struggles to grasp the complexities of human nature and its non-rational impulses. This thesis does not adhere to a strictly rationalist view of autonomy in goal-setting but instead acknowledges that autonomy is expressed through an individual's intent, whether rational or not.

Thus, at one level, the design of augmented human augmentation technologies should optimize for the autonomy of intentions—ensuring individuals can pursue their chosen goals. Yet, aspirationally, design should also support rational autonomy, enabling individuals to make informed, deliberate choices while recognizing the inherent unpredictability of human nature.

One of the central conclusions of this thesis, as a reader familiar with Kant's philosophy may have anticipated, is the imperative to treat both humans and augmented humans as *ends*, not merely as *means*. Kant firmly argued that rational beings should never be treated solely as instruments for external purposes.

We must ensure that the design of future human augmentations does not reduce augmented humans, who are moral agents, to mere tools. They must not be instrumentalized. Traditional tools, such as a hammer, exist purely as instruments, and even a personal computer, though more complex, remains external to the user, functioning as a separate entity. In contrast, an augmentation such as a neural interface or a prosthetic limb becomes part of the individual's lived experience. When technology becomes integrated into the self, it is no longer just an instrument, it holds intrinsic value. The augmented human, therefore, retains and reinforces their intrinsic worth as an *end in themselves*.

12.5.1 Are Autonomy and Heteronomy discrete categories?

In [Het2], we analyzed concerns about enhancement and other HCI technologies, conducting both Exploratory and Confirmatory Factor Analyses. The results identified two fundamental factors shaping attitudes toward augmented humans: Perceived Autonomy and Perceived

Control over Agency. Initially, we referred to the latter as social threat, but further reflection suggests a stronger alignment with the concept of heteronomy.

These factors are not discrete categories. There is no clear boundary where autonomy ends, and heteronomy begins. Moreover, while these emerged as the most relevant factors, they are not necessarily the only ones influencing the integration of augmented humans into society.

Autonomy must also be understood in a social context. It does not exist in isolation but must be balanced with the autonomy of others—a fundamental concern of ethics. Participation in any social contract inherently limits full autonomy, as individuals must navigate interdependencies within a society. Further, there are factors that originate in one side of the autonomy heteronomy spectrum and influence the other side, in the following, I reflect on those:

Originating from Autonomy Autonomy and heteronomy are interdependent, with no evident boundary between them. External influences constantly shape individual autonomy, and in turn, an individual's autonomous choices can contribute to heteronomous structures. At the individual level, decision-making, agency, and self-determination influence how autonomy develops over time. However, autonomy is not isolated. For example, an augmented human in a leadership position may directly or indirectly influence others' decisions to adopt augmentation. On a smaller scale, an individual's augmentation can alter the behaviors and attitudes of their immediate social environment, constraining the autonomy of others.

Originating from Heteronomy Heteronomy affects both individual autonomy and societal structures. External narratives, such as media hype around AI and emerging technologies, shape individual expectations and decisions. As shown in this thesis, exaggerated expectations can influence behavior, just as social stigma can discourage adoption or disclosure of augmentation technologies. Beyond individual impact, heteronomous processes shape societal norms and policies. Media representations reinforce biases, cultural expectations emerge around augmentation use, and regulatory frameworks develop in response to technological adoption. These forces define the conditions under which augmented humans integrate into society.

12.6 Implications

In this section, I elaborate on the overarching implications of this thesis, however, for the reader that is looking for more specific, actionable implications, please refer to the specific chapters. In the following I elaborate on the implications of this work using the taxonomy proposed by Van Berkel and Hornbæk [503]. To relate these implications to existing HCI frameworks, I use Laaksoharju [508] for theory and design and Tan et al. [509]'s recently published framework for the implications for the Augmented Humans community.

12.6.1 Implications for Theory

The HCI community has made multiple calls to produce more theories and models, yet theory-building remains an uncommon practice [510]. This work aims to help establish a foundation for developing theories in HCI. In particular, I revisit Laaksoharju [508]'s Theory of Use in HCI, which addresses two key questions: First, *What should be designed?*, —or, as Laaksoharju [508] puts it, *"how designers of computerized artifacts come to understand what they should design."*, Second *How can we determine whether a computerized artifact is appropriate for a given purpose?*

To answer the first question, Laaksoharju [508] advocates for a deductive theorizing approach grounded in *falsification*, drawing from the philosophy of science and psychological research on problem-solving and decision-making. I align with this view[2], recognizing it as a fundamental driver of HCI as a proto-science in pursuit of discovery.

Laaksoharju [508] operationalizes this approach by requiring an explicit articulation of the designers' understanding in the form of a *Theory of Use*. This theory must account for the assumptions designers make about users and usage contexts, as well as their assumptions of what constitutes a "good solution." These assumptions should be formulated as falsifiable statements, allowing designers to test and refine their understanding based on user feedback systematically. Additionally, the *Theory of Use* serves as a set of requirements against which design hypotheses can be evaluated.

To guide this evaluation, Laaksoharju [508] emphasizes the autonomy of the artifact's users, which, in this case, is the autonomy of the augmented human. He argues that autonomy is essential for human well-being, yet current design practices often overlook this principle.

This thesis contributes to the Theory of Use by addressing the second question, specifically in the context of augmented humans, identifying factors that span the spectrum from autonomy to heteronomy.

Finally, Laaksoharju [508]'s Theory of Use not only provides designers with the autonomy necessary for developing expertise but also enhances communication with users. It ensures that designers remain focused on understanding what they are designing for, while giving users control over the technology that shapes their lives.

12.6.2 Implications for methodology

The findings from [Aut1] reinforce Kosch et al. [184]'s position that placebo research should recognize a new subcategory of placebos introduced by digital artifacts.

[2]The call to align HCI to the falsification approach from Popper [511] has not only helped other disciplines such as psychology to achieve a status of science [512] but also has been previously recommended as an approach for HCI to follow in order to build and test theories [510].

Placebo effects in research are typically explained by two concurrent processes. Expectancy-oriented theories attribute placebo effects to increased treatment efficacy beliefs, while conditioned response explanations link placebo effects to previously established stimulus-response associations (e.g., taking a drug and feeling better). Given that augmentation technologies are novel, such stimulus-response linkages could not have formed. Our results, therefore, support expectancy-based mechanisms as the primary driver of placebo effects in this context. One might argue that higher-order associations between novel technology and subjective improvement exist, but this does not account for the physiological effects we observed, particularly those related to loss-information integration. Other placebo induction mechanisms, such as social learning [513], likely play a role—individuals may experience benefits simply by observing others using augmentation technologies. These alternative pathways warrant further exploration.

Our replication of placebo effects in augmentation technologies highlights the need for placebo control in this research domain, much like in psychological and medical intervention studies. However, unlike these fields, technology user studies often involve participants who are acutely aware of the novelty of the technology, making group assignments more transparent. As [184] suggests, placebo control methods must be adapted to the specific constraints of technology research. We propose five approaches for controlling placebo effects in augmentation technologies, acknowledging that these are neither exhaustive nor universally applicable. While some extend to other areas of technology evaluation, each study must be carefully designed to isolate genuine effects beyond placebo influences.

Five ways of addressing the placebo effect in the evaluation of augmentation technologies:

1. Present a placebo condition with a non-functional augmentation technology and compare it to the functional system – *placebo-control*
2. Control for contextual aspects [180] that are known to increase placebo effects – *placebo-reduction*
3. Poll expectations before and after use – *placebo-indicator*
4. Consider indirect measures (e.g., physiological measures) when probing the augmentation technology – *placebo-indicator*
5. Assess users' qualitative statements in an interview can highlight a mismatch between expectation – *placebo-indicator*

One could argue that this research paves the way for a variety of follow-up investigations for each new technology. However, medical placebo trials can provide a framework for defining the limits of such follow-up research.

First, only studies that can identify the conditions and mechanisms under which placebo effects occur and the potential consequences of placebo effects are relevant to technology evaluation. Here, there is substantial knowledge in the medical literature to start and replicate effects that generalize across technologies.

Second, AI in human-centered AI, or augmentation technologies, are examples of technology that create high expectations in their users. Therefore, another constraint is that only technologies that raise high expectations may need placebo control.

Third, the placebo effects we found are small, and thus, false-positive inferences due to placebo effects may only be relevant for user studies that found small effects in statistical comparisons. Overall, while placebo research must be considered in the evaluation of technology, we have to understand the constraints and mechanisms of placebo effects in the evaluation of technology before invalidating large amounts of prior research.

12.6.3 Implications for the Augmented Humans community

This work addresses how to harmonize human augmentation with the self (augmented human) and society. To expand on this, I will draw on the Assistive Augmentation framework recently published in *Interactions* by Tan et al. [509]. The authors structure assistive augmentation around two key pillars: *ability* and *integration*. The *ability* pillar contains *perceptual*, *physical*, and *cognitive* domains, aligning with established *Sensory*, *Motor*, and *Cognitive* augmentation frameworks, as the one used here, originally proposed by Raisamo et al. [266]. It also considers the augmentation method, categorized as *amplify*, *substitute*, or *extend*.

The second pillar, *integration*, is particularly relevant to this thesis as it embraces multiple dimensions that have been addressed in this thesis. In detail, it examines four dimensions: (1) *body integration*, which concerns how naturally the augmentation becomes part of the user's physical experience; (2) *temporal integration*, which considers the consistency of augmentation use in daily life; (3) *identity integration*, which reflects how well the augmentation aligns with the user's sense of self; and (4) *sociocultural integration*, which addresses how augmentation is perceived within societal and cultural contexts.

For the first pillar, based on the findings of [Het1], one can argue that *amplifying* and *substituting* are more socially acceptable as they can take place in order to compensate for an impairment, which in turn can serve as the "entry point" of augmentations into society. Extending, as seen before, if introduced, one must consider having a clear narrative on why it is necessary to extend a given skill, otherwise, users could face some level of rejection or segregation given the uncertainty of the purpose of having an augmentation, naturally, this applies also for substituting and amplifying in the case they are not using in the context of impairments.

For the second pillar, I would argue that (1) and (3) are tightly codependent and highly influenced by the factors studied in Part II, especially the sense of agency, as an external body part that is not fully internalized, that the user fails to attribute their functioning to their own intentions, in a physical or cognitive level (if we also consider the potential scenario of brain implants or brain stimulation). Thus, designing also for the agency as a subfactor of autonomy is fundamental, especially considering that agency can easily be lost in the case of a technical failure, high execution delay, or even extremely low latency [406].

As discussed in [Aut3], personal discourse can closely align with AI-generated discourse when AI provides contextually relevant suggestions. In such cases, individuals often adopt the readily available AI-generated options rather than following their own initial thoughts. Given this, it is important to consider how augmentation might shape an individual's personality and self-expression.

Beyond the findings of this thesis, additional evidence suggests that individuals are already adapting their speech patterns to incorporate expressions commonly associated with large language models [514]. While any technology deeply embedded in daily life (2) inevitably influences its users, designers of augmentation systems should strive to minimize their impact on personality and self-determination. In the end, user autonomy should remain a core design principle in the long term.

12.6.4 Implications for design

To elaborate on the implications for design, I frame the main implications for the design of this work using seven key questions proposed by Laaksoharju [508] in *Designing for Autonomy*. Based on my findings, I address the following:

Is it important that users do not make mistakes? Users should have the ability to make and learn from mistakes rather than relying entirely on automation. Augmentation should provide mechanisms for error recovery without restricting autonomy. Systems that overly prevent errors can limit user engagement and adaptability. Instead, designs should allow controlled exposure to errors, ensuring that users develop an understanding of the augmentation and its use until they ultimately implicitly interact with it.

Is it important that users follow regulations? For augmentation systems that impact safety or decision-making, regulatory compliance might be necessary. In such scenarios, it might be complex for designers to both, comply with potential safeguard mechanisms required by regulators and allow the augmented human total control over the augmentation, in order to avoid social rejection by bystanders, one might argue that designers and manufacturers should not externally control augmented humans, even in the case of regulations, and that such, should me mediated directly between the individual and the regulating entity.

Is it important that users can learn to use the system easily? Augmentation should integrate into daily routines with minimal friction. Interfaces must be clear and *transparent to use,* reducing the learning curve without sacrificing functionality. Systems should be designed so that users can understand and use them effectively without extensive prior training, making augmentation more accessible across different expertise levels.

Is it important that users have full control over what they do? Augmentation should assist rather than dictate. There must be no actions that the user needs to override; all controls should be in the human; not guaranteeing so would have psychological and societal implications for the users of human augmentations [Het1, Het2, Aut1, Aut2, Aut3].

Is it important that users can appropriate the tool to better fit their needs? Looking at the future, when augmentations can actually have a permanent nature, one would expect designs to be individual-specific; even so, the broad functionality would be similar. This resembles how some assistive technologies are designed and adapted to the individuals using them; I think in this regard, we, as augmented human designers, still have a lot to learn from assistive technologies designers, and more, we should consider ways to scale this up to a general public.

Is it important that users develop expertise? Some augmentation systems should support skill-building rather than de-skilling. This is why leaving autonomy to the human is so important, in the moment that the system starts having autonomy and doing task for the human, the human loses the need to preserve a given skill, this problem has already been experience in automation and human-in-the-loop systems [515].

Is it important that users can take responsibility for their decisions? When a system influences decision-making, users must remain accountable for their choices. Systems should present information clearly, ensuring that users understand how system-generated suggestions are derived. Designs should avoid passive reliance on AI-driven outputs and instead encourage critical engagement, particularly in areas where augmentation shapes communication, judgment, or behavior [Aut3].

12.6.5 Implications for practice

With AI-integrated smart glasses becoming more common, these findings point to important design choices. Devices like Ray-Ban Meta rely on auditory feedback, while others, such as Vuzix Blade and Xreal Air, integrate both visual and auditory modalities. How these systems provide assistance, when they intervene, and the level of user control will shape how people experience agency.

Addressing agency loss requires rethinking interaction strategies. Instead of optimizing efficiency alone, systems should reinforce user autonomy. One approach is adaptive AI, which adjusts assistance based on user confidence, ensuring augmentation remains a support rather than a directive. Another is designing interfaces that prompt users to evaluate AI suggestions rather than accept them passively. These directions need further study, particularly in high-stakes environments where misaligned reliance on augmentation could have serious consequences.

12.6.6 Implications for policy

Marketing and regulation of augmentation technologies must account for the behavioral shifts they induce. Heightened expectations, often reinforced through marketing, shape how users interact with these systems. Users may overestimate augmentation capabilities, leading to misplaced trust or overreliance. **A key regulatory step would be independent verification of augmentation claims before market release.**

Another challenge is that even when users are informed of an augmentation's limitations, they often disregard this information. My findings, consistent with prior work, suggest that disclaimers alone are ineffective. Interventions should make these limitations tangible—through controlled failure scenarios during onboarding or interactive risk assessments that expose users to edge cases where augmentation does not function as expected.

This also raises questions about accountability. If an augmentation contributes to a failure, who is responsible? The user, the manufacturer, or both?.

12.6.7 Implications for society

As augmentation technologies become more common, they may introduce new social expectations. Augmented individuals could face additional ethical and legal constraints on how they use their enhancements. Whether these rules emerge informally or through policy, they will shape how augmentation fits into society.

A largely unexamined issue is how augmentation interacts with aging. Biological bodies deteriorate, but augmented parts do not necessarily follow the same pattern. If augmentation preserves function beyond natural limits, does it redefine aging? Could an individual reach an age where their biological components decline while their augmented capabilities remain stable?

This has implications for inequality. If augmentation becomes essential for maintaining function in old age, access to upgrades may determine who remains capable and who experiences decline. A future where some individuals can continually update their augmentations while others cannot could create new divides, not just in wealth but in ability.

These are not distant concerns. Augmentation is already reshaping human experience, and the questions it raises—about autonomy, responsibility, and access—need answers now.

13

APPENDIXES:

A APPENDIX

A.1 Clarification on Contributions

This section reports the specific contributions to each paper using the Contributor Roles Taxonomy (CRediT). The CRediT framework provides a taxonomy to transparently identify and report the roles played by each contributor in a paper. In 2022, CRediT was formally adopted as an ANSI/NISO standard (Z39.104-2022), therefore, I adhere to this standard. Table A.1, A.2, and A.3 provide an overview of the author's and collaborators' contributions to the core publications included in this thesis.

The taxonomy defines the following 14 roles: [1]

Conceptualization: Ideas; formulation or evolution of overarching research goals and aims.

Methodology: Development or design of methodology; creation of models.

Software: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.

Validation: Verification, whether as a part of the activity or separate, of the overall replication and reproducibility of results/experiments and other research outputs.

Formal Analysis: Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.

Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.

Resources: Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.

Data Curation: Management activities to annotate (produce metadata), scrub data, and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse.

Writing – Original Draft: Creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).

Writing – Review & Editing: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary, or revision – including pre- or post-publication stages.

[1]Textually extracted from: `https://credit.niso.org/contributor-roles/resources/`

Visualization: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.

Supervision: Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.

Project Administration: Management and coordination responsibility for the research activity planning and execution.

Funding Acquisition: Acquisition of the financial support for the project leading to this publication.

Table A.1: Author contributions for the chapter on **Autonomy**. The contributions of the thesis author are highlighted.

	Conceptualization	Data curation	Formal Analysis	Funding acquisition	Investigation	Methodology	Project administration	Resources	Software	Supervision	Validation	Visualization	Writing – original draft	Writing – review & editing
[Aut1] In: *Computers in Human Behavior* 146 [2023]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Thomas Kosch	✓	✓	✓			✓							✓	✓
Felix Grelka		✓			✓				✓		✓			
Albrecht Schmidt	✓			✓			✓			✓				✓
Robin Welsch	✓		✓			✓	✓				✓	✓	✓	✓
[Aut2] In: *Computers in Human Behavior: Artificial Humans* [2025]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Lisa Barth		✓	✓		✓					✓	✓	✓		
Francesco Chiossi						✓								✓
Robin Welsch	✓												✓	✓
Thomas Kosch	✓												✓	✓
[Aut3] In: *Proc. of the 26th Intl. Conference on Multimodal Interaction.* [2024]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Yannick Weiss										✓				
Mei Yi Lu		✓	✓		✓				✓			✓		
Moritz Ziarko					✓									
Albrecht Schmidt	✓			✓			✓			✓				✓
Jasmin Niess	✓		✓			✓							✓	✓

Table A.1: (Continued) Author contributions for the chapter on **Autonomy**. The contributions of the thesis author are highlighted.

[Aut4] In: ***Computers in Human Behavior Reports*** **[2025]**	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Thomas Kosch	✓										✓		✓	✓
Agnes Mercedes Kloft		✓			✓									✓
Jasper Quinn		✓			✓									
Sari Kujala														✓
Robin Welsch	✓		✓		✓	✓	✓			✓	✓		✓	✓

Table A.2: Author contributions for the chapter on **Heteronomy**. The contributions of the thesis author are highlighted.

	Conceptualization	Data curation	Formal Analysis	Funding acquisition	Investigation	Methodology	Project administration	Resources	Software	Supervision	Validation	Visualization	Writing – original draft	Writing – review & editing
[Het1] In: ***Proc. of the 2023 CHI Conference on Human Factors in Computing Systems.*** **[2023]**														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓		✓	✓	✓	✓
Jasmin Niess			✓				✓			✓	✓		✓	✓
Takuro Nakao					✓									
Jonathan Lazar		✓			✓	✓					✓		✓	
Albrecht Schmidt	✓			✓		✓	✓			✓			✓	✓
Tonja Machulla	✓	✓	✓		✓	✓	✓			✓	✓		✓	✓
[Het2] In: ***Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 7.3*** **[2023]**														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓		✓	✓	✓	✓
Jasmin Niess	✓					✓				✓	✓		✓	✓
Albrecht Schmidt	✓			✓			✓			✓				✓
Robin Welsch	✓		✓			✓	✓			✓	✓	✓	✓	✓

Table A.3: Author contributions for the chapter on **Experiencing Augmentation**. The contributions of the thesis author are highlighted.

Contributor	Conceptualization CON	Data curation DCU	Formal Analysis FAN	Funding acquisition FUN	Investigation INV	Methodology MET	Project administration PAD	Resources RES	Software SOF	Supervision SUP	Validation VAL	Visualization VIS	Writing – original draft WOR	Writing – review & editing WRE
[Exp1] In: *Proc. of the 2025 CHI Conference on Human Factors in Computing Systems.* [2025]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Finn Krammer		✓			✓	✓			✓		✓			
Yannick Weiss							✓			✓				
Robin Welsch			✓									✓	✓	✓
Thomas Kosch	✓					✓				✓			✓	✓
[Exp2] In: *Proc. of the 2021 ACM Intl. Symposium on Wearable Computers.* [2021]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Jasmin Niess						✓					✓		✓	✓
Bettina Eska			✓		✓				✓		✓			✓
Albrecht Schmidt	✓			✓			✓			✓				✓
Tonja Machulla	✓		✓			✓				✓				✓
[Exp3] In: *Proc. of the Intl. Conference on Mobile and Ubiquitous Multimedia.* [2024]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Robin Neuhaus		✓	✓		✓						✓		✓	✓
Yannick Weiss	✓		✓			✓				✓	✓		✓	✓
Marc Hassenzahl														✓
[Exp4] In: *Proc. ACM Hum.-Comput. Interact. 8.MHCI* [2024]														
	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Yannick Weiss	✓					✓	✓							✓
Niklas Hirsch		✓	✓		✓				✓		✓	✓		
Alexander Wiethoff	✓									✓				✓

Table A.3: (Continued) Author contributions for the chapter on **Experiencing Augmentation**. The contributions of the thesis author are highlighted.

	CON	DCU	FAN	FUN	INV	MET	PAD	RES	SOF	SUP	VAL	VIS	WOR	WRE
[Exp5] In: ***Proc. of the ACM on Human-Computer Interaction 6.ISS*** **[2022]**														
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Sven Mayer	✓	✓	✓			✓			✓	✓		✓	✓	✓
Jess Hartcher-O'Brien	✓					✓	✓				✓			
Albrecht Schmidt	✓			✓			✓			✓				✓
Tonja Machulla	✓		✓			✓	✓			✓	✓	✓	✓	✓
[Exp6] In: ***2024 IEEE Intl. Symposium on Mixed and Augmented Reality (ISMAR)*** **[2024]**														
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Kenji Ishihara		✓			✓				✓		✓	✓		
Moritz Ziarko					✓				✓					✓
Sebastian Günther	✓												✓	✓
Florian Müller	✓		✓			✓	✓				✓		✓	✓
[Exp7] In: ***Proc. of Mensch und Computer 2022.*** **[2022]**														
Steeven Villa	✓	✓			✓	✓	✓		✓		✓	✓	✓	✓
Sven Mayer	✓	✓				✓			✓	✓	✓		✓	✓
[Exp8] In: ***Proc. of the 2023 ACM Intl. Symposium on Wearable Computers.*** **[2023]**														
Steeven Villa	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓
Thuy Duong Nguyen		✓			✓				✓				✓	
Benjamin Tag						✓							✓	✓
Tonja Machulla	✓						✓							
Albrecht Schmidt	✓			✓			✓			✓				✓
Jasmin Niess	✓					✓				✓	✓		✓	✓

Bibliography

Autonomy

[Aut1] **Villa, Steeven**, Kosch, Thomas, Grelka, Felix, Schmidt, Albrecht, and Welsch, Robin. 'The placebo effect of human augmentation: Anticipating cognitive augmentation increases risk-taking behavior.' In: *Computers in Human Behavior* 146 (2023), p. 107787. DOI: 10.1016/j.chb.2023.107787.

[Aut2] **Villa, Steeven**, Barth, Lisa, Chiossi, Francesco, Welsch, Robin, and Kosch, Thomas. 'Whose Mind is it Anyway? A Systematic Review and Exploration on Agency in Cognitive Augmentation.' In: *Computers in Human Behavior: Artificial Humans* (2025).

[Aut3] **Villa, Steeven**, Weiss, Yannick, Lu, Mei Yi, Ziarko, Moritz, Schmidt, Albrecht, and Niess, Jasmin. 'Envisioning Futures: How the Modality of AI Recommendations Impacts Conversation Flow in AR-enhanced Dialogue.' In: *Proceedings of the 26th International Conference on Multimodal Interaction*. ICMI '24. Association for Computing Machinery, 2024, 182–193. DOI: 10.1145/3678957.3685731.

[Aut4] **Villa, Steeven**, Kosch, Thomas, Kloft, Agnes Mercedes, Quinn, Jasper, Kujala, Sari, and Welsch, Robin. 'Inventory of User Expectations for Technology (iExpect).' In: *Computers in Human Behavior Reports* (2025).

Heteronomy

[Het1] **Villa, Steeven**, Niess, Jasmin, Nakao, Takuro, Lazar, Jonathan, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Understanding Perception of Human Augmentation: A Mixed-Method Study.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 2023, pp. 1–16. DOI: 10.1145/3544548.3581485.

[Het2] **Villa, Steeven**, Niess, Jasmin, Schmidt, Albrecht, and Welsch, Robin. 'Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale.' In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7.3 (2023). Publisher: ACM, pp. 1–23. DOI: 10.1145/3610915.

Experiencing Augmentation

[Exp1] **Villa, Steeven**, Krammer, Finn Jacob Eliyah, Weiss, Yannick, Welsch, Robin, and Kosch, Thomas. 'Understanding the Influence of Electrical Muscle Stimulation on Motor Learning: Enhancing Motor Learning or Disrupting Natural Progression?' In:

Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan (2025).

[Exp2] **Villa, Steeven**, Niess, Jasmin, Eska, Bettina, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Assisting Motor Skill Transfer for Dance StudentsUsing Wearable Feedback.' In: *Proceedings of the 2021 ACM International Symposium on Wearable Computers*. ISWC '21. Association for Computing Machinery, 2021, pp. 38–42. DOI: 10.1145/3460421.3478817.

[Exp3] **Villa, Steeven**, Neuhaus, Robin, Weiss, Yannick, and Hassenzahl, Marc. 'Exploring Virtual Reality as a Platform for Early-Stage Design for Human Augmentation Technologies.' In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. MUM '24. Association for Computing Machinery, 2024, 464–466. DOI: 10.1145/3701571.3703386.

[Exp4] **Villa, Steeven**, Weiss, Yannick, Hirsch, Niklas, and Wiethoff, Alexander. 'An Examination of Ultrasound Mid-air Haptics for Enhanced Material and Temperature Perception in Virtual Environments.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI (2024). DOI: 10.1145/3676488.

[Exp5] **Villa, Steeven**, Mayer, Sven, Hartcher-O'Brien, Jess, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'Extended Mid-air Ultrasound Haptics for Virtual Reality.' In: *Proceedings of the ACM on Human-Computer Interaction* 6.ISS (2022), 578:500–578:524. DOI: 10.1145/3567731.

[Exp6] **Villa, Steeven**, Ishihara, Kenji, Ziarko, Moritz, Günther, Sebastian, and Müller, Florian. 'Touch It Like It's Hot: A Thermal Feedback Enabled Encountered-type Haptic Display for Virtual Reality.' In: *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2024, pp. 700–709. DOI: 10.1109/ISMAR62088.2024.00085.

[Exp7] **Villa, Steeven** and Mayer, Sven. 'Cobity: A Plug-And-Play Toolbox to Deliver Haptics in Virtual Reality.' In: *Proceedings of Mensch und Computer 2022*. MuC '22. ACM, 2022, pp. 78–84. DOI: 10.1145/3543758.3543775.

[Exp8] **Villa, Steeven**, Nguyen, Thuy Duong, Tag, Benjamin, Machulla, Tonja-Katrin, Schmidt, Albrecht, and Niess, Jasmin. 'Towards a Haptic Taxonomy of Emotions: Exploring Vibrotactile Stimulation in the Dorsal Region.' In: *Proceedings of the 2023 ACM International Symposium on Wearable Computers*. ISWC '23. Association for Computing Machinery, 2023, pp. 50–54. DOI: 10.1145/3594738.3611363.

References

[1] Chiossi, Francesco, Kosch, Thomas, Menghini, Luca, **Villa, Steeven**, and Mayer, Sven. 'SensCon: Embedding Physiological Sensing into Virtual Reality Controllers.' In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI (2023), 223:1–223:32. DOI: 10.1145/3604270.

[2] Chiossi, Francesco, Welsch, Robin, **Villa, Steeven**, Chuang, Lewis, and Mayer, Sven. 'Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience.' In: *Big Data and Cognitive Computing* 6.2 (2022). Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, p. 55. DOI: 10.3390/bdcc6020055.

[3] Schött, Svenja Y., Capsi-Morales, Patricia, **Villa, Steeven**, Butz, Andreas, and Piazza, Cristina. 'Would You Hold My Hand? Exploring External Observers' Perception of Artificial Hands.' In: *Multimodal Technologies and Interaction* 7.7 (2023). Number: 7 Publisher: Multidisciplinary Digital Publishing Institute, p. 71. DOI: 10.3390/mti7070071.

[4] Weiss, Yannick, **Villa, Steeven**, Grootjen, Jesse W, Hoppe, Matthias, Kale, Yasin, and Müller, Florian. 'Exploring Redirection and Shifting Techniques to Mask Hand Movements from Shoulder-Surfing Attacks during PIN Authentication in Virtual Reality.' In: *Proc. ACM Hum.-Comput. Interact.* 8.MHCI (2024), 257:1–257:24. DOI: 10.1145/3676502.

[5] Shen, Chenxinran, Karpus, Jurgis, Kosch, Thomas, Fernandes, Daniela, Mello, Beatriz, Welsch, Robin, and **Villa, Steeven**. 'The Impact of Asymmetric AI Assistance on Decision-Making in Social Dilemmas: A Study on Human Augmentation in Economic Games.' In: *Augmented Humans 2025*. AHs 2025. Association for Computing Machinery, 2025.

[6] Leusmann, Jan, **Villa, Steeven**, Berberoglu, Burak, Wang, Chao, and Mayer, Sven. 'Developing and Validating the Perceived System Curiosity Scale (PSC): Measuring Users' Perceived Curiosity of Systems.' In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan* (2025).

[7] Leusmann, Jan, **Villa, Steeven**, Liang, Thomas, Wang, Chao, Schmidt, Albrecht, and Mayer, Sven. 'An Approach to Elicit Human-Understandable Robot Expressions to Support Human-Robot Interaction.' In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25), Yokohama, Japan* (2025).

[8] Chiossi, Francesco, **Villa, Steeven**, Hauser, Melanie, Welsch, Robin, and Chuang, Lewis. 'Design of On-body Tactile Displays to Enhance Situation Awareness in Automated Vehicles.' In: *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*. ISSN: 2377-9322. 2022, pp. 1–6. DOI: 10.1109/CIVEMSA53371.2022.9853690.

[9] Chiossi, Francesco, Welsch, Robin, **Villa, Steeven**, Chuang, Lewis L., and Mayer, Sven. ‘Designing a Physiological Loop for the Adaptation of Virtual Human Characters in a Social VR Scenario.’ In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2022, pp. 578–579. doi: 10.1109/VRW55335.2022.00140.

[10] Dietz, Dennis, **Villa, Steeven**, Ziarko, Moritz, Bonfert, Michael, Müller, Florian, and Butz, Andreas. ‘Embracer: A Wearable Encountered-Type Haptic Controller for 3 DoF Input and Feedback.’ In: *Proceedings of the 2024 ACM International Symposium on Wearable Computers*. ISWC ’24. Association for Computing Machinery, 2024, pp. 140–143. doi: 10.1145/3675095.3676626.

[11] Faltaous, Sarah, Hubert, Anna, Karolus, Jakob, **Villa, Steeven**, Kosch, Thomas, and Wozniak, Pawel W. ‘EMStriker: Potentials of Enhancing the Training Process of Racket-based Sports via Electrical Muscle Stimulation.’ In: *Proceedings of the Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction*. TEI ’22. Association for Computing Machinery, 2022, pp. 1–6. doi: 10.1145/3490149.3505578.

[12] Kloft, Agnes Mercedes, Welsch, Robin, Kosch, Thomas, and **Villa, Steeven**. ‘"AI enhances our performance, I have no doubt this one will do the same": The Placebo effect is robust to negative descriptions of AI.’ In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI ’24. Association for Computing Machinery, 2024, pp. 1–24. doi: 10.1145/3613904.3642633.

[13] Pangratz, Elisabeth, Chiossi, Francesco, **Villa, Steeven**, Gramann, Klaus, and Gehrke, Lukas. ‘Towards an Implicit Metric of Sensory-Motor Accuracy: Brain Responses to Auditory Prediction Errors in Pianists.’ In: *Proceedings of the 15th Conference on Creativity and Cognition*. C&C ’23. Association for Computing Machinery, 2023, pp. 129–138. doi: 10.1145/3591196.3593340.

[14] Tannert, Isabel, Schulleri, Katrin H., Michel, Youssef, **Villa, Steeven**, Johannsen, Leif, Hermsdörfer, Joachim, and Lee, Dongheui. ‘Immediate Effects of Vibrotactile Biofeedback Instructions on Human Postural Control.’ In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. ISSN: 2694-0604. 2021, pp. 7426–7432. doi: 10.1109/EMBC46164.2021.9630992.

[15] Weiss, Yannick, **Villa, Steeven**, Schmidt, Albrecht, Mayer, Sven, and Müller, Florian. ‘Using Pseudo-Stiffness to Enrich the Haptic Experience in Virtual Reality.’ In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI ’23. Association for Computing Machinery, 2023, pp. 1–15. doi: 10.1145/3544548.3581223.

[16] Fernandes, Daniela, **Villa, Steeven**, Nicholls, Salla, Haavisto, Otso, Buschek, Daniel, Schmidt, Albrecht, Kosch, Thomas, Shen, Chenxinran, and Welsch, Robin. AI Makes

You Smarter, But None The Wiser: The Disconnect Between Performance and Metacognition. arXiv:2409.16708 [cs]. 2024. DOI: 10.48550/arXiv.2409.16708.

[17] schraefel, mc, Andrés, Josh, Tabor, Aaron, Bateman, Scott, Wanyu Liu, Abby, Jones, Mike, Kunze, Kai, Murnane, Elizabeth, and **Villa, Steeven**. 'Body As Starting Point 4: Inbodied Interaction Design for Health Ownership.' In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Association for Computing Machinery, 2021, pp. 1–5. DOI: 10.1145/3411763.3441335.

[18] **Villa, Steeven**, Welsch, Robin, Denisova, Alena, and Kosch, Thomas. 'Evaluating Interactive AI: Understanding and Controlling Placebo Effects in Human-AI Interaction.' In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. CHI EA '24. Association for Computing Machinery, 2024, pp. 1–4. DOI: 10.1145/3613905.3636304.

[19] Koelle, Marion, Ananthanarayan, Swamy, and Boll, Susanne. 'Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2020, 1–19.

[20] Koelle, Marion, El Ali, Abdallah, Cobus, Vanessa, Heuten, Wilko, and Boll, Susanne CJ. 'All about Acceptability? Identifying Factors for the Adoption of Data Glasses.' In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Chi '17. Association for Computing Machinery, 2017, 295–300. DOI: 10.1145/3025453.3025749.

[21] Chignell, Mark, Hancock, Peter A., and Takeshita, Harumi. 'Chapter 11 - Human—Computer Interaction: The Psychology of Augmented Human Behavior.' In: *Human Performance and Ergonomics*. Ed. by P. A. Hancock. Handbook of Perception and Cognition (Second Edition). Academic Press, 1999, pp. 291–328. DOI: 10.1016/B978-012322735-5/50012-9.

[22] Abdelrahman, Yomna, Sahami Shirazi, Alireza, Henze, Niels, and Schmidt, Albrecht. 'Investigation of material properties for thermal imaging-based interaction.' In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 15–18.

[23] Kasahara, Shunichi, Nishida, Jun, and Lopes, Pedro. 'Preemptive Action: Accelerating Human Reaction using Electrical Muscle Stimulation Without Compromising Agency.' In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–15.

[24] Huber, Jochen, Shilkrot, Roy, Maes, Pattie, and Nanayakkara, Suranga, eds. Assistive Augmentation. Cognitive Science and Technology. Springer, 2018. DOI: 10.1007/978-981-10-6404-3.

[25] Farooq, Umer and Grudin, Jonathan. 'Human-computer integration.' In: *Interactions* 23.6 (2016), pp. 26–32. DOI: 10.1145/3001896.

[26] Pinker, Steven. Enlightenment Now: The Case for Reason, Science, Humanism, and Progress. Google-Books-ID: xn83DwAAQBAJ. Penguin UK, 2018.

[27] Engelbart, Douglas C. AUGMENTING HUMAN INTELLECT: A Conceptual Framework. Tech. rep. 1962.

[28] Engelbart, Douglas C. and English, William K. 'A research center for augmenting human intellect.' In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)*. ACM Press, 1968, p. 395. DOI: 10.1145/1476589.1476645.

[29] Wobbrock, Jacob O. and Kientz, Julie A. 'Research contributions in human-computer interaction.' In: *interactions* 23.3 (2016), pp. 38–44. DOI: 10.1145/2907069.

[30] Blandford, Ann, Furniss, Dominic, and Makri, Stephann. Qualitative HCI research: Going behind the scenes. Vol. 9. 1. Morgan & Claypool Publishers, 2016, pp. 1–115.

[31] Pfeiffer, Max, Duente, Tim, and Rohs, Michael. 'Let Your Body Move: A Prototyping Toolkit for Wearable Force Feedback with Electrical Muscle Stimulation.' In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '16. event-place: Florence, Italy. Association for Computing Machinery, 2016, pp. 418–427. DOI: 10.1145/2935334.2935348.

[32] Ven, Katinka van de and Mulrooney, Kyle JD. 'An introduction to human enhancement drugs.' In: *Human Enhancement Drugs*. Routledge, 2020.

[33] Socci, Valentina, Tempesta, Daniela, Desideri, Giovambattista, De Gennaro, Luigi, and Ferrara, Michele. 'Enhancing human cognition with cocoa flavonoids.' In: *Frontiers in Nutrition* 4 (2017), p. 19.

[34] Soderberg, William. 'Human Genetic Modifications and Parental Perspectives.' In: *Proceedings of the XXIII World Congress of Philosophy*. Vol. 3. 2018, pp. 115–120.

[35] Raisamo, Roope, Rakkolainen, Ismo, Majaranta, Päivi, Salminen, Katri, Rantala, Jussi, and Farooq, Ahmed. 'Human augmentation: Past, present and future.' In: *International Journal of Human-Computer Studies* 131 (2019), pp. 131–143.

[36] Guerrero, Graciela, Silva, Fernando José Mateus da, Fernández-Caballero, Antonio, and Pereira, António. 'Augmented Humanity: A Systematic Mapping Review.' In: *Sensors* 22.2 (2022), p. 514. DOI: 10.3390/s22020514.

[37] Proulx, Michael J. 'Synthetic synaesthesia and sensory substitution.' In: *Consciousness and Cognition* 19 (2010), pp. 501–503.

[38] Danilov, Yuri P., Tyler, Mitchell E., and Kaczmarek, Kurt A. 'Vestibular sensory substitution using tongue electrotactile display.' In: Birkhäuser Basel, 2008.

[39] Kumar, Vikash, Hote, Yogesh Vijay, and Jain, Shivam. 'Review of Exoskeleton: History, Design and Control.' In: *2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE)* (2019), pp. 677–682.

[40] Boeck, De. 'STRUCTURING HUMAN AUGMENTATION WITHIN PRODUCT DESIGN.' In: (2021), pp. 16–20. DOI: 10.1017/pds.2021.534.

[41] Schmidt, Albrecht. 'Augmenting Human Intellect and Amplifying Perception and Cognition.' In: *IEEE Pervasive Computing* 16.1 (2017). Conference Name: IEEE Pervasive Computing, pp. 6–10. DOI: 10.1109/MPRV.2017.8.

[42] Nabeshima, Junichi, Saraiji, MHD Yamen, and Minamizawa, Kouta. 'Arque Artificial Biomimicry-Inspired Tail for Extending Innate Body Functions.' In: *ACM SIGGRAPH 2019 Posters* (2019). DOI: 10.1145/3306214.

[43] Saraiji, M. H.D. Yamen, Sasaki, Tomoya, Kunze, Kai, Minamizawa, Kouta, and Inami, Masahiko. 'MetaArmS: Body remapping using feet-controlled artificial arms.' In: *UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* 18 (2018), pp. 65–74. DOI: 10.1145/3242587.3242665.

[44] Cornelio, Patricia, Haggard, Patrick, Hornbaek, Kasper, Georgiou, Orestis, Bergström, Joanna, Subramanian, Sriram, and Obrist, Marianna. 'The sense of agency in emerging technologies for human–computer integration: A review.' In: *Frontiers in Neuroscience* 16 (2022).

[45] Licklider, J. C. R. 'Man-Computer Symbiosis.' In: *IRE Transactions on Human Factors in Electronics* HFE-1.1 (1960). Conference Name: IRE Transactions on Human Factors in Electronics, pp. 4–11. DOI: 10.1109/THFE2.1960.4503259.

[46] Mueller, Florian Floyd, Lopes, Pedro, Strohmeier, Paul, Ju, Wendy, Seim, Caitlyn, Weigel, Martin, Nanayakkara, Suranga, Obrist, Marianna, Li, Zhuying, Delfa, Joseph, et al. 'Next Steps for Human-Computer Integration.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–15.

[47] Inami, Masahiko, Uriu, Daisuke, Kashino, Zendai, Yoshida, Shigeo, Saito, Hiroto, Maekawa, Azumi, and Kitazaki, Michiteru. 'Cyborgs, Human Augmentation, Cybernetics, and JIZAI Body.' In: *Augmented Humans 2022*. AHs 2022. Association for Computing Machinery, 2022, 230–242. DOI: 10.1145/3519391.3519401.

[48] Kelly, Michelle E., Duff, Hollie, Kelly, Sara, McHugh Power, Joanna E., Brennan, Sabina, Lawlor, Brian A., and Loughrey, David G. 'The impact of social activities, social networks, social support and social relationships on the cognitive functioning

of healthy older adults: a systematic review.' In: *Systematic Reviews* 6.1 (2017), p. 259. DOI: 10.1186/s13643-017-0632-2.

[49] Feng, Chunliang, Eickhoff, Simon B., Li, Ting, Wang, Li, Becker, Benjamin, Camilleri, Julia A., Hétu, Sébastien, and Luo, Yi. 'Common brain networks underlying human social interactions: Evidence from large-scale neuroimaging meta-analysis.' In: *Neuroscience & Biobehavioral Reviews* 126 (2021), pp. 289–303. DOI: 10.1016/j.neubiorev.2021.03.025.

[50] Boeck, M. De and Vaes, K. 'Human augmentation and its new design perspectives.' In: *International Journal of Design Creativity and Innovation* 12 (2024), pp. 61–80. DOI: 10.1080/21650349.2023.2288125.

[51] Cao, Heng, Ling, Zhengyang, Zhu, Jun, Wang, Yu, and Wang, Wei. 'Design frame of a leg exoskeleton for load-carrying augmentation.' In: *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. Ieee. 2009.

[52] Marcheschi, Simone, Salsedo, Fabio, Fontana, Marco, and Bergamasco, Massimo. 'Body Extender: Whole body exoskeleton for human power augmentation.' In: *2011 IEEE International Conference on Robotics and Automation*. Ieee. 2011.

[53] Brown, M., Tsagarakis, N., and Caldwell, D.G. 'Exoskeletons for human force augmentation.' In: *Industrial Robot: An International Journal* 30.6 (2003), pp. 592–602.

[54] Takahashi, Akifumi, Brooks, Jas, Kajimoto, Hiroyuki, and Lopes, Pedro. 'Increasing Electrical Muscle Stimulation's Dexterity by means of Back of the Hand Actuation.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 2021, pp. 1–12. DOI: 10.1145/3411764.3445761.

[55] Garcia, Gabriela, Arauz, P., Alvarez, Isabel, Encalada, Nicolas, Vega, Shirley, and Martin, B. 'Impact of a passive upper-body exoskeleton on muscle activity, heart rate and discomfort during a carrying task.' In: *PLOS ONE* 18 (2023). DOI: 10.1371/journal.pone.0287588.

[56] Nassour, John, Zhao, Guoping, and Grimmer, M. 'Soft pneumatic elbow exoskeleton reduces the muscle activity, metabolic cost and fatigue during holding and carrying of loads.' In: *Scientific Reports* 11 (2021). DOI: 10.1038/s41598-021-91702-5.

[57] Malvezzi, M., Iqbal, Z., Valigi, M. C., Pozzi, Maria, Prattichizzo, D., and Salvietti, G. 'Design of Multiple Wearable Robotic Extra Fingers for Human Hand Augmentation.' In: *Robotics* 8 (2019). DOI: 10.3390/robotics8040102.

[58] Xiloyannis, M., Chiaradia, D., Frisoli, A., and Masia, L. 'Physiological and kinematic effects of a soft exosuit on arm movements.' In: *Journal of NeuroEngineering and Rehabilitation* 16 (2019). DOI: 10.1186/s12984-019-0495-y.

[59] Luger, T., Seibt, R., Cobb, Timothy J., Rieger, M., and Steinhilber, B. 'Influence of a passive lower-limb exoskeleton during simulated industrial work tasks on physical load, upper body posture, postural control and discomfort.' In: *Applied ergonomics* 80 (2019), pp. 152–160. DOI: 10.1016/J.APERGO.2019.05.018.

[60] Demidas, Aneta. 'Electrical stimulation: then and now. Applications and limitations.' In: *Physiotherapy Quarterly* 29.4 (2021), pp. 81–86. DOI: 10.5114/pq.2021.110987.

[61] Geddes, L.A. 'The first stimulators-reviewing the history of electrical stimulation and the devices crucial to its development.' In: *IEEE Engineering in Medicine and Biology Magazine* 13.4 (1994), pp. 532–542. DOI: 10.1109/51.310995.

[62] Knibbe, J., Alsmith, A., and Hornbæk, K. 'Experiencing Electrical Muscle Stimulation.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2.3 (2018). DOI: 10.1145/3264928.

[63] Moreno-Aranda, J. and Seireg, A. 'Electrical parameters for over-the-skin muscle stimulation.' In: *Journal of Biomechanics* 14.9 (1981), pp. 579–585. DOI: https://doi.org/10.1016/0021-9290(81)90083-X.

[64] Parry, Selina M, Berney, Sue, Granger, Catherine L, Koopman, Renè, El-Ansary, Doa, and Denehy, Linda. 'Electrical muscle stimulation in the intensive care setting: a systematic review.' In: *Critical care medicine* 41.10 (2013), pp. 2406–2418. DOI: 10.1097/CCM.0b013e3182923642.

[65] Toshiaki Miyamoto Hiroyuki Kamada, Akira Tamaki and Moritani, Toshio. 'Low-intensity electrical muscle stimulation induces significant increases in muscle strength and cardiorespiratory fitness.' In: *European Journal of Sport Science* 16.8 (2016), pp. 1104–1110. DOI: 10.1080/17461391.2016.1151944.

[66] Porcari, John P, McLean, Karen Palmer, Foster, Carl, Kernozek, Thomas, Crenshaw, Ben, and Swenson, Chad. 'Effects of electrical muscle stimulation on body composition, muscle strength, and physical appearance.' In: *The Journal of Strength & Conditioning Research* 16.2 (2002), pp. 165–172.

[67] Abdelrahman, Yomna, Knierim, Pascal, Wozniak, Pawel W, Henze, Niels, and Schmidt, Albrecht. 'See through the fire: evaluating the augmentation of visual perception of firefighters using depth and thermal cameras.' In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 2017, pp. 693–696.

[68] Schmidt, Albrecht. 'Augmenting human intellect and amplifying perception and cognition.' In: *IEEE Pervasive Computing* 16.1 (2017), pp. 6–10.

[69] Schoeller, F., Haar, A J H, Jain, Aarthi, and Maes, Pattie. 'Enhancing human emotions with interoceptive technologies.' In: *Physics of life reviews* (2019).

[70] Cernea, Daniel and Kerren, Andreas. 'A survey of technologies on the rise for emotion-enhanced interaction.' In: *J. Vis. Lang. Comput.* 31 (2015), pp. 70–86.

[71] Tsetserukou, Dzmitry and Neviarouskaya, Alena. 'FeelIM: Augmenting Emotions during Online Communication.' In: *IEEE Computer Graphics and Applications* 30.5 (2010), pp. 72–80. DOI: `10.1109/MCG.2010.88`.

[72] Derks, Daantje, Fischer, Agneta H., and Bos, Arjan E. R. 'The role of emotion in computer-mediated communication: A review.' In: *Comput. Hum. Behav.* 24 (2008), pp. 766–785.

[73] Foo, Esther W., Dunne, Lucy E., and Holschuh, Brad. 'User Expectations and Mental Models for Communicating Emotions through Compressive and Warm Affective Garment Actuation.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5.1 (2021). DOI: `10.1145/3448097`.

[74] Charles, Fred, Pizzi, David, Cavazza, Marc, Vogt, Thurid, and André, Elisabeth. 'EmoEmma: emotional speech input for interactive storytelling.' In: *Adaptive Agents and Multi-Agent Systems*. 2009.

[75] Berenguer, Carmen, Baixauli, Inmaculada, Gómez, Soledad Andrés, Andres, M. D., and Stasio, Simona De. 'Exploring the Impact of Augmented Reality in Children and Adolescents with Autism Spectrum Disorder: A Systematic Review.' In: *International Journal of Environmental Research and Public Health* 17 (2020).

[76] Pfister, Hans-Rüdiger and Böhm, Gisela. 'The multiplicity of emotions: A framework of emotional functions in decision making.' In: *Judgment and Decision Making* 3.1 (2008), 5–17. DOI: `10.1017/s1930297500000127`.

[77] Barrett, Lisa Feldman. 'Are Emotions Natural Kinds?' In: *Perspectives on Psychological Science* 1 (2006), pp. 28–58.

[78] Arafsha, Faisal, Alam, Kazi Masudul, and El Saddik, Abdulmotaleb. 'Design and development of a user centric affective haptic jacket.' In: *Multimedia Tools and Applications* 74 (2013), pp. 1–18. DOI: `10.1007/s11042-013-1767-3`.

[79] Mazzoni, Antonella and Bryan-Kinns, Nick. 'How does it feel like? An exploratory study of a prototype system to convey emotion through haptic wearable devices.' In: *2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN)* (2015), pp. 64–68.

[80] Ko, Chang Yong, Chang, Yunhee, Kim, Sol Bi, Kim, Shin-Ki, Kim, Gyoosuk, Ryu, Jeicheong, and Mun, Mu-Seong. 'Evaluation of physical and emotional responses to

vibrotactile stimulation of the forearm in young adults, the elderly, and transradial amputees.' In: *Physiology & Behavior* 138 (2015), pp. 87–93.

[81] Huang, Hsinfu and Hsieh, Ming-Hsuan. 'Tactile Emotional coding: The Perceptual Linking of Vibrotactile Stimuli with Basic Emotions.' In: *2019 IEEE 2nd International Conference on Knowledge Innovation and Invention (ICKII)* (2019), pp. 134–137.

[82] Réhman, Shafiq ur and Liu, Li. 'iFeeling: Vibrotactile Rendering of Human Emotions on Mobile Phones.' In: *WMMP*. 2008.

[83] Lemmens, Paul M. C., Crompvoets, Floris, Brokken, Dirk, Eerenbeemd, Jack van den, and Vries, Gert-Jan de. 'A body-conforming tactile jacket to enrich movie viewing.' In: *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (2009), pp. 7–12.

[84] Zupan, Barbra and Eskritt, Michelle. 'Eliciting emotion ratings for a set of film clips: A preliminary archive for research in emotion.' In: *The Journal of Social Psychology* 160.6 (2020), pp. 768–789.

[85] Chandra, Yurike, Tag, Benjamin, Peiris, Roshan Lalintha, and Minamizawa, K. 'Preliminary Investigation of Across-Body Vibrotactile Pattern for the Design of Affective Furniture.' In: *2020 IEEE Haptics Symposium (HAPTICS)* (2020), pp. 671–676.

[86] Yoo, Yongjae, Yoo, Taekbeom, Kong, Jihyun, and Choi, Seungmoon. 'Emotional responses of tactile icons: Effects of amplitude, frequency, duration, and envelope.' In: *2015 IEEE World Haptics Conference (WHC)* (2015), pp. 235–240.

[87] Rantala, Jussi, Salminen, Katri, Raisamo, R., and Surakka, Veikko. 'Touch gestures in communicating emotional intention via vibrotactile stimulation.' In: *Int. J. Hum. Comput. Stud.* 71 (2013), pp. 679–690.

[88] Arafsha, Faisal, Alam, Kazi Masudul, and Saddik, Abdulmotaleb El. 'Design and development of a user centric affective haptic jacket.' In: *Multimedia Tools and Applications* 74 (2015), pp. 3035–3052.

[89] Wikström, Valtteri, Makkonen, Tommi, and Saarikivi, Katri. 'Exploring universal expression of emotion with vibrotactile patterns composed using dynamic features.' In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 2017, pp. 77–78. DOI: 10.1109/ACIIW.2017.8272590.

[90] Cinel, Caterina, Valeriani, Davide, and Poli, Riccardo. 'Neurotechnologies for Human Cognitive Augmentation: Current State of the Art and Future Prospects.' In: *Frontiers in Human Neuroscience* 13 (2019), p. 13. DOI: 10.3389/fnhum.2019.00013.

[91] Cai, Runze, Janaka, Nuwan Nanayakkarawasam Peru Kandage, Zhao, Shengdong, and Sun, Minghui. 'ParaGlassMenu: Towards Social-Friendly Subtle Interactions in Conversations.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 2023, pp. 1–21. DOI: 10.1145/3544548.3581065.

[92] Hodges, Steve, Williams, Lyndsay, Berry, Emma, Izadi, Shahram, Srinivasan, James, Butler, Alex, Smyth, Gavin, Kapur, Narinder, and Wood, Ken. 'SenseCam: A Retrospective Memory Aid.' In: *UbiComp 2006: Ubiquitous Computing*. Ed. by Paul Dourish and Adrian Friday. Lecture Notes in Computer Science. Springer, 2006, pp. 177–193. DOI: 10.1007/11853565_11.

[93] Kapur, Arnav, Kapur, Shreyas, and Maes, Pattie. 'AlterEgo: A Personalized Wearable Silent Speech Interface.' In: *23rd International Conference on Intelligent User Interfaces*. ACM, 2018, pp. 43–53. DOI: 10.1145/3172944.3172977.

[94] Ksibi, Amel, Alluhaidan, Ala Saleh D., Salhi, Amina, and El-Rahman, Sahar A. 'Overview of Lifelogging: Current Challenges and Advances.' In: vol. 9. 2021, pp. 62630–62641.

[95] Brich, Irina R., Bause, Inga M., Hesse, Friedrich W., and Wesslein, Ann-Katrin. 'Working memory affine technological support functions improve decision performance.' In: *Computers in Human Behavior* 92 (2019), pp. 238–249. DOI: https://doi.org/10.1016/j.chb.2018.11.014.

[96] Dobrowolski, Pawel, Hanusz, Krzysztof, Sobczyk, Bartosz, Skorko, Maciek, and Wiatrow, Andrzej. 'Cognitive enhancement in video game players: The role of video game genre.' In: *Computers in Human Behavior* 44 (2015), pp. 59–63. DOI: https://doi.org/10.1016/j.chb.2014.11.051.

[97] Cristina, Olarte-Pascual, Jorge, Pelegrín-Borondo, Eva, Reinares-Lara, and Mario, Arias-Oliva. 'From wearable to insideable: Is ethical judgment key to the acceptance of human capacity-enhancing intelligent technologies?' In: *Computers in Human Behavior* 114 (2021), p. 106559. DOI: https://doi.org/10.1016/j.chb.2020.106559.

[98] Zisk, Alyssa Hillary and Dalton, Elizabeth. 'Augmentative and Alternative Communication for Speaking Autistic Adults: Overview and Recommendations.' In: *Autism in Adulthood* 1.2 (2019), pp. 93–100. DOI: 10.1089/aut.2018.0007.

[99] Kane, Shaun K. and Morris, Meredith Ringel. 'Let's Talk About X: Combining Image Recognition and Eye Gaze to Support Conversation for People with ALS.' In: *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM, 2017, pp. 129–134. DOI: 10.1145/3064663.3064762.

[100] Vargas, Mauricio Fontana de, Dai, Jiamin, and Moffatt, Karyn. 'AAC with Automated Vocabulary from Photographs: Insights from School and Speech-Language Ther-

apy Settings.' In: *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '22. Association for Computing Machinery, 2022, pp. 1–18. DOI: 10.1145/3517428.3544805.

[101] Ofek, Eyal, Iqbal, Shamsi T., and Strauss, Karin. 'Reducing disruption from subtle information delivery during a conversation: mode and bandwidth investigation.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 3111–3120. DOI: 10.1145/2470654.2466425.

[102] Klein, Amarolinda, Sørensen, Carsten, Freitas, Angilberto Sabino de, Pedron, Cristiane Drebes, and Elaluf-Calderwood, Silvia. 'Understanding controversies in digital platform innovation processes: The Google Glass case.' In: *Technological Forecasting and Social Change* 152 (2020), p. 119883. DOI: 10.1016/j.techfore.2019.119883.

[103] Shultz, Kara. 'Every implanted child a star (and some other failures): 1 Guilt and shame in the cochlear implant debates 2.' In: *Quarterly Journal of Speech* 86.3 (2000). Publisher: NCA Website _eprint: https://doi.org/10.1080/00335630009384296, pp. 251–275. DOI: 10.1080/00335630009384296.

[104] Evans, Leonard. 'ABS and Relative Crash Risk Under Different Roadway, Weather, and Other Conditions.' In: *SAE Transactions* 104 (1995). Publisher: SAE International, pp. 676–686.

[105] Yamamoto, Atsushi and Kimura, Yoshiaki. 'Influence of ABS on rollover accidents.' In: 1996.

[106] Katsiaficas, George. 'The Necessity of Autonomy.' In: *New Political Science* (2001). Publisher: Taylor & Francis Group. DOI: 10.1080/07393140120099633.

[107] Campbell, L. 'Kant, autonomy and bioethics.' In: *Ethics, Medicine and Public Health* 3.3 (2017), pp. 381–392. DOI: 10.1016/j.jemep.2017.05.008.

[108] Brown, Garrett W. 'The idea of autonomy: Accountability, self-determinism and what normative claims about institutional autonomy in global governance should mean.' In: *International Organizations and the Idea of Autonomy*. Num Pages: 16. Routledge, 2011.

[109] Bennett, Dan, Metatla, Oussama, Roudaut, Anne, and Mekler, Elisa D. 'How does HCI Understand Human Agency and Autonomy?' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3580651.

[110] Ryan, Richard M. and Deci, Edward L. 'Self-Determination Theory.' In: *Encyclopedia of Quality of Life and Well-Being Research*. Ed. by Filomena Maggino. Springer Inter-

national Publishing, 2023, pp. 6229–6235. DOI: 10.1007/978-3-031-17299-1_2630.

[111] Ballou, Nick, Deterding, Sebastian, Tyack, April, Mekler, Elisa D, Calvo, Rafael A, Peters, Dorian, Villalobos-Zúñiga, Gabriela, and Turkay, Selen. 'Self-Determination Theory in HCI: Shaping a Research Agenda.' In: *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI EA '22. Association for Computing Machinery, 2022, pp. 1–6. DOI: 10.1145/3491101.3503702.

[112] Faas, Cedric, Bergs, Richard, Sterz, Sarah, Langer, Markus, and Feit, A. 'Give Me a Choice: The Consequences of Restricting Choices Through AI-Support for Perceived Autonomy, Motivational Variables, and Decision Performance.' In: (2024).

[113] Schneider, Sascha, Nebel, Steve, Beege, Maik, and Rey, Günter Daniel. 'The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media.' In: *Learning and Instruction* (2018). DOI: 10.1016/J.LEARNINSTRUC.2018.06.006.

[114] Rekimoto, Jun and Nagao, Katashi. 'The World through the Computer: Computer Augmented Interaction with Real World Environments.' In: *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology*. Uist '95. Association for Computing Machinery, 1995. DOI: 10.1145/215585.215639.

[115] Cumiskey, Kathleen M. 'Mobile symbiosis: a precursor to public risk-taking behavior?' In: *Mobile Communication*. Routledge, 2017, pp. 27–31.

[116] Low, Banus Kam Leung and Chan, Alan Hoi Shou. 'Identification of Observable Risk-Taking Behaviours Among SCUBA Divers.' In: *Convergence of Ergonomics and Design*. Ed. by Alma Maria Jennifer Gutierrez, Ravindra S. Goonetilleke, and Rex Aurellius C. Robielos. Springer International Publishing, 2021, pp. 102–111.

[117] Borenstein, Jason, Wagner, Alan R, and Howard, Ayanna. 'Overtrust of pediatric health-care robots: a preliminary survey of parent perspectives.' In: *IEEE Robotics & Automation Magazine* 25.1 (2018), pp. 46–54.

[118] Bredereke, Jan and Lankenau, Axel. 'A rigorous view of mode confusion.' In: *International Conference on Computer Safety, Reliability, and Security*. Springer. 2002, pp. 19–31.

[119] Stirling, Leia, Siu, Ho Chit, Jones, Eric, and Duda, Kevin. 'Human factors considerations for enabling functional use of exosystems in operational environments.' In: *IEEE Systems Journal* 13.1 (2018), pp. 1072–1083.

[120] Shafir, Eldar, Simonson, Itamar, and Tversky, Amos. 'Reason-based choice.' In: *Cognition* 49.1 (1993), pp. 11–36. DOI: https://doi.org/10.1016/0010-0277(93)90034-S.

[121] Buelow, Melissa T. and Blaine, Amber L. 'The assessment of risky decision making: a factor analysis of performance on the Iowa Gambling Task, Balloon Analogue Risk Task, and Columbia Card Task.' In: *Psychological assessment* 27 3 (2015), pp. 777–85.

[122] Deck, Cary, Lee, Jungmin, Reyes, Javier, and Rosen, Chris. 'Risk-Taking Behavior: An Experimental Analysis of Individuals and Dyads.' In: *Southern Economic Journal* 79.2 (2012), pp. 277–299.

[123] Fromme, Kim, Katz, Elizabeth C., and Rivet, Kathy. In: *Cognitive Therapy and Research* 21.4 (1997), pp. 421–442.

[124] Nicholson, Nigel, Soane, Emma, Fenton-O'Creevy, Mark, and Willman, Paul. 'Personality and domain-specific risk taking.' In: *Journal of Risk Research* 8 (2005), pp. 157–176.

[125] Poon, Kean. 'Understanding Risk-taking Behavior in Bullies, Victims, and Bully Victims Using Cognitive- and Emotion-Focused Approaches.' In: *Frontiers in Psychology* 7 (2016).

[126] Hanoch, Yaniv and Gummerum, Michaela. 'A comparison of the risk-taking behaviors of prisoners and non-prisoners.' In: *Journal of Behavioral Decision Making* 24.4 (2010), pp. 431–442.

[127] McKenna, Frank P. and Horswill, Mark S. 'Risk taking from the participant's perspective: the case of driving and accident risk.' In: *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 25 2 (2006), pp. 163–70.

[128] Buelow, Melissa T. and Suhr, Julie A. 'Construct Validity of the Iowa Gambling Task.' In: *Neuropsychology Review* 19 (2009), pp. 102–114.

[129] Lejuez, Carl W., Read, Jennifer P., Kahler, Christopher W., Richards, Jerry B., Ramsey, Susan E., Stuart, Gregory L., Strong, David R., and Brown, Richard A. 'Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART).' In: *Journal of experimental psychology. Applied* 8 2 (2002), pp. 75–84.

[130] Figner, Bernd, Mackinlay, Rachael J., Wilkening, Friedrich, and Weber, Elke U. 'Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task.' In: *Journal of experimental psychology. Learning, memory, and cognition* 35 3 (2009), pp. 709–30.

[131] Holper, Lisa and Murphy, Ryan O. 'Hemodynamic and affective correlates assessed during performance on the Columbia Card Task (CCT).' In: *Brain Imaging and Behavior* 8 (2013), pp. 517–530.

[132] Moore, James W. 'What Is the Sense of Agency and Why Does it Matter?' In: *Frontiers in Psychology* 7 (2016). DOI: 10.3389/fpsyg.2016.01272.

[133] Limerick, Hannah, Coyle, David, and Moore, James W. 'The experience of agency in human-computer interactions: a review.' In: *Frontiers in Human Neuroscience* 8 (2014), p. 643. DOI: 10.3389/fnhum.2014.00643.

[134] Jeannerod, Marc. 'The mechanism of self-recognition in humans.' In: *Behavioural Brain Research* 142.1 (2003), pp. 1–15. DOI: https://doi.org/10.1016/S0166-4328(02)00384-4.

[135] Tapal, Adam, Oren, Ela, Dar, Reuven, and Eitam, Baruch. 'The Sense of Agency Scale: A Measure of Consciously Perceived Control over One's Mind, Body, and the Immediate Environment.' In: *Frontiers in Psychology* 8 (2017). DOI: 10.3389/fpsyg.2017.01552.

[136] Haggard, Patrick, Clark, Sam, and Kalogeras, Jeri. 'Voluntary action and conscious awareness.' In: *Nature neuroscience* 5.4 (2002), pp. 382–385. DOI: 10.1038/nn827.

[137] Wolpe, Noham and Rowe, James B. 'Beyond the "urge to move": objective measures for the study of agency in the post-Libet era.' In: *Frontiers in Human Neuroscience* 8 (2014). DOI: 10.3389/fnhum.2014.00450.

[138] Jeunet, Camille, Albert, Louis, Argelaguet, Ferran, and Lécuyer, Anatole. '"Do You Feel in Control?": Towards Novel Approaches to Characterise, Manipulate and Measure the Sense of Agency in Virtual Environments.' In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (2018), pp. 1486–1495. DOI: 10.1109/TVCG.2018.2794598.

[139] Kang, Suk Yun, Im, Chang-Hwan, Shim, Miseon, Nahab, Fatta B, Park, Jihye, Kim, Do-Won, Kakareka, John, Miletta, Nathanial, and Hallett, Mark. 'Brain Networks Responsible for Sense of Agency: An EEG Study.' In: *PLoS ONE* 10.8 (2015), e0135261. DOI: https://doi.org/10.1371/journal.pone.0135261.

[140] Klimesch, Wolfgang. 'Alpha-band oscillations, attention, and controlled access to stored information.' In: *Trends in Cognitive Sciences* 16.12 (2012), pp. 606–617. DOI: https://doi.org/10.1016/j.tics.2012.10.007.

[141] Bu-Omer, Hani M, Gofuku, Akio, Sato, Kenji, and Miyakoshi, Makoto. 'Parieto-occipital alpha and low-beta EEG power reflect sense of agency.' In: *Brain Sciences* 11.6 (2021), p. 743.

[142] Buchholz, Verena N, David, Nicole, Sengelmann, Malte, and Engel, Andreas K. 'Belief of agency changes dynamics in sensorimotor networks.' In: *Scientific Reports* 9.1 (2019), p. 1995.

[143] Goering, Sara, Klein, Eran, Specker Sullivan, Laura, Wexler, Anna, Arcas, Blaise Agüera y, Bi, Guoqiang, Carmena, Jose M, Fins, Joseph J, Friesen, Phoebe, Gallant, Jack, et al. 'Recommendations for responsible development and application of neurotechnologies.' In: *Neuroethics* 14.3 (2021), pp. 365–386. DOI: 10.1007/s12152-021-09468-6.

[144] Haselager, Pim. 'Did I do that? Brain–computer interfacing and the sense of agency.' In: *Minds and Machines* 23 (2013), pp. 405–418. DOI: 10.1007/s11023-012-9298-7.

[145] Reath, Andrews. 'Hedonism, Heteronomy, and Kant's Principle of Happiness.' In: *Agency and Autonomy in Kant's Moral Theory: Selected Essays*. Ed. by Andrews Reath. Oxford University Press, 2006, p. 0. DOI: 10.1093/0199288836.003.0003.

[146] Stensson, Patrik and Jansson, Anders. 'Autonomous technology - sources of confusion: a model for explanation and prediction of conceptual shifts.' In: *Ergonomics* 57.3 (2014), pp. 455–470. DOI: 10.1080/00140139.2013.858777.

[147] Axelsson, Carl-Anton Werner, Löscher, Ida, Jansson, Anders, and Tschirner, Simon. Authority and Level of Automation: Lessons to Be Learned in Design of In-vehicle Assistance Systems. 2014. DOI: 10.1007/978-3-319-07227-2_39.

[148] Kant, Immanuel. Fundamental Principles of the Metaphysic of Morals. Binker North, 1785.

[149] Koelle, Marion, Ananthanarayan, Swamy, and Boll, Susanne. 'Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Chi '20. Association for Computing Machinery, 2020, 1–19. DOI: 10.1145/3313831.3376162.

[150] Kocielnik, Rafal, Amershi, Saleema, and Bennett, Paul N. 'Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-User Expectations of AI Systems.' In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Chi '19. Association for Computing Machinery, 2019, 1–14. DOI: 10.1145/3290605.3300641.

[151] Shneiderman, Ben. 'Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy.' In: *International Journal of Human–Computer Interaction* 36.6 (2020), pp. 495–504. DOI: 10.1080/10447318.2020.1741118. eprint: https://doi.org/10.1080/10447318.2020.1741118.

[152] Horstmann, Aike C. and Krämer, Nicole C. 'Expectations vs. actual behavior of a social robot: An experimental investigation of the effects of a social robot's interaction skill level and its expected future role on people's evaluations.' In: *Plos One* 15.8 (2020), pp. 1–18. DOI: 10.1371/journal.pone.0238133.

[153] Bhimani, Hardik, Mention, Anne-Laure, and Barlatier, Pierre-Jean. 'Social media and innovation: A systematic literature review and future research directions.' In: *Technological Forecasting and Social Change* (2019).

[154] Do, Tiffany D., McMahan, Ryan P., and Wisniewski, Pamela J. 'A New Uncanny Valley? The Effects of Speech Fidelity and Human Listener Gender on Social Perceptions of a Virtual-Human Speaker.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. Chi '22. Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3517564.

[155] Taherdoost, Hamed. 'A review of technology acceptance and adoption models and theories.' In: *Procedia Manufacturing* 22 (2018), pp. 960–967.

[156] Shinohara, Kristen and Wobbrock, Jacob O. 'In the shadow of misperception: assistive technology use and social interactions.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011).

[157] Li, Franklin Mingzhe, Chen, Di Laura, Fan, Mingming, and Truong, Khai N. '"I Choose Assistive Devices That Save My Face": A Study on Perceptions of Accessibility and Assistive Technology Use Conducted in China.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021).

[158] Lang, Florian, Schmidt, Albrecht, and Machulla, Tonja. 'Augmented Reality for People with Low Vision: Symbolic and Alphanumeric Representation of Information.' In: *Computers Helping People with Special Needs* 12376 (2020), pp. 146–156.

[159] Laat, F. A. J. de, Heerebeek, Bart van, and Netten, Jaap J. van. 'Advantages and disadvantages of interdisciplinary consultation in the prescription of assistive technologies for mobility limitations.' In: *Disability and Rehabilitation: Assistive Technology* 14 (2019), pp. 386–390.

[160] Marangunić, Nikola and Granić, Andrina. 'Technology acceptance model: a literature review from 1986 to 2013.' In: *Universal Access in the Information Society* 14.1 (2015), pp. 81–95. DOI: 10.1007/s10209-014-0348-1.

[161] Kelly, Norene and Gilbert, Stephen B. 'The WEAR Scale: Developing a Measure of the Social Acceptability of a Wearable Device.' In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2016).

[162] Woźniak, Paweł W., Karolus, Jakob, Lang, Florian, Eckerth, Caroline, Schöning, Johannes, Rogers, Yvonne, and Niess, Jasmin. 'Creepy Technology:What Is It and How Do You Measure It?' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Chi '21. Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445299.

[163] Muehlhaus, Marie, Steimle, Jürgen, and Koelle, Marion. 'Feather Hair: Interacting with Sensorized Hair in Public Settings.' In: *Designing Interactive Systems Conference*. DIS '22. Association for Computing Machinery, 2022, 1228–1242. DOI: 10.1145/3532106.3533527.

[164] Lu, Xi, Thomaz, Edison, and Epstein, Daniel A. 'Understanding People's Perceptions of Approaches to Semi-Automated Dietary Monitoring.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6.3 (2022). DOI: 10.1145/3550288.

[165] Nittala, Aditya Shekhar and Steimle, Jürgen. 'Next Steps in Epidermal Computing: Opportunities and Challenges for Soft On-Skin Devices.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3517668.

[166] Fitz, Nicholas S, Nadler, Roland, Manogaran, Praveena, Chong, Eugene WJ, and Reiner, Peter B. 'Public attitudes toward cognitive enhancement.' In: *Neuroethics* 7.2 (2014), pp. 173–188.

[167] Scheske, Christel and Schnall, Simone. 'The ethics of "smart drugs": Moral judgments about healthy people's use of cognitive-enhancing drugs.' In: *Basic and applied social psychology* 34.6 (2012), pp. 508–515.

[168] Dubljević, Veljko. 'Cognitive enhancement, rational choice and justification.' In: *Neuroethics* 6.1 (2013), pp. 179–187.

[169] Davis, Nick J. 'A taxonomy of harms inherent in cognitive enhancement.' In: *Frontiers in Human Neuroscience* 11 (2017).

[170] Bostrom, Nick and Sandberg, Anders. 'Cognitive enhancement: methods, ethics, regulatory challenges.' In: *Science and engineering ethics* 15.3 (2009), pp. 311–341.

[171] Conrad, Erin C, Humphries, Stacey, and Chatterjee, Anjan. 'Attitudes toward cognitive enhancement: the role of metaphor and context.' In: *AJOB neuroscience* 10.1 (2019), pp. 35–47.

[172] ROUSI, Rebekah and Renko, Roni. 'Emotions toward Cognitive Enhancement Technologies and the body–attitudes and willingness to use.' In: *International Journal of Human-Computer Studies* (2020), p. 102472.

[173] Conrad, Erin C., Humphries, Stacey, and Chatterjee, Anjan. 'Attitudes Toward Cognitive Enhancement: The Role of Metaphor and Context.' In: *AJOB Neuroscience* 10.1 (2019). Pmid: 31070552, pp. 35–47. DOI: 10.1080/21507740.2019.1595771. eprint: https://doi.org/10.1080/21507740.2019.1595771.

[174] Breivik, Gunnar, Hanstad, Dag Vidar, and Loland, Sigmund. 'Attitudes towards use of performance-enhancing substances and body modification techniques. A

comparison between elite athletes and the general population.' In: *Sport in Society* 12.6 (2009), pp. 737–754. DOI: 10.1080/17430430902944183. eprint: https://doi.org/10.1080/17430430902944183.

[175] Dijkstra, Anne M. and Schuijff, Mirjam. 'Public opinions about human enhancement can enhance the expert-only debate: A review study.' In: *Public Understanding of Science* 25 (2016), pp. 588–602.

[176] Petróczi, Andrea and Aidman, Eugene. 'Measuring explicit attitude toward doping: Review of the psychometric properties of the Performance Enhancement Attitude Scale.' In: *Psychology of Sport and Exercise* 10.3 (2009), pp. 390–396. DOI: https://doi.org/10.1016/j.psychsport.2008.11.001.

[177] Folkerts, Dirk, Loh, Roland, Petróczi, Andrea, and Brueckner, Sebastian. 'The performance enhancement attitude scale (PEAS) reached 'adulthood': Lessons and recommendations from a systematic review and meta-analysis.' In: *Psychology of Sport and Exercise* 56 (2021), p. 101999. DOI: https://doi.org/10.1016/j.psychsport.2021.101999.

[178] Elprama, Shirley A., Vanderborght, Bram, and Jacobs, An. 'An industrial exoskeleton user acceptance framework based on a literature review of empirical studies.' In: *Applied ergonomics* 100 (2021), p. 103615.

[179] Beck, Owen N., Taboga, Paolo, and Grabowski, Alena M. 'Sprinting with prosthetic versus biological legs: insight from experimental data.' In: *Royal Society Open Science* 9 (2022).

[180] Wager, Tor D and Atlas, Lauren Y. 'The neuroscience of placebo effects: connecting context, learning and health.' In: *Nature Reviews Neuroscience* 16.7 (2015), pp. 403–418.

[181] Colzato, Lorenza S., Barone, Hayley, Sellaro, Roberta, and Hommel, Bernhard. 'More attentional focusing through binaural beats: evidence from the global–local task.' In: *Psychological Research* 81 (2017), pp. 271–277.

[182] Møller, Henrik and Pedersen, Christian. 'Hearing at low and infrasonic frequencies.' In: *Noise & health* 6 (2004), pp. 37–57.

[183] Clements-Cortés, Amy, Ahonen, Heidi, Evans, Michael, Freedman, Morris, and Bartel, Lee R. 'Short-Term Effects of Rhythmic Sensory Stimulation in Alzheimer's Disease: An Exploratory Pilot Study.' In: *Journal of Alzheimer's disease : JAD* 52 2 (2016), pp. 651–60.

[184] Kosch, Thomas, Welsch, Robin, Chuang, Lewis, and Schmidt, Albrecht. 'The Placebo Effect of Artificial Intelligence in Human-Computer Interaction.' In: *ACM Transactions on Computer-Human Interaction* (2022).

[185] Geers, Andrew L, Weiland, Paul E, Kosbab, Kristin, Landry, Sarah J, and Helfer, Suzanne G. 'Goal activation, expectations, and the placebo effect.' In: *Journal of personality and social psychology* 89.2 (2005), p. 143.

[186] Gama, Pedro A Magalhães De Saldanha da, Slama, Hichem, Caspar, Emilie A, Gevers, Wim, and Cleeremans, Axel. 'Placebo-suggestion modulates conflict resolution in the Stroop Task.' In: *PloS one* 8.10 (2013), e75701.

[187] Groot, Kristel de and Van Strien, Jan W. 'Event-related potentials in response to feedback following risk-taking in the hot version of the Columbia Card Task.' In: *Psychophysiology* 56.9 (2019), e13390.

[188] Panno, Angelo. 'Trait Emotional Intelligence Is Related to Risk Taking when Adolescents Make Deliberative Decisions.' In: *Games* 7.3 (2016). DOI: 10.3390/g7030023.

[189] Hart, Sandra G. and Staveland, Lowell E. 'Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.' In: *Human Mental Workload*. Ed. by Peter A. Hancock and Najmedin Meshkati. Vol. 52. Advances in Psychology. North-Holland, 1988, pp. 139–183. DOI: https://doi.org/10.1016/S0166-4115(08)62386-9.

[190] Sochat, Vanessa. 'The Experiment Factory: Reproducible Experiment Containers.' In: *The Journal of Open Source Software* 3 (2018), p. 521. DOI: 10.21105/joss.00521.

[191] de Cheveigné, Alain and Nelken, Israel. 'Filters: When, Why, and How (Not) to Use Them.' In: *Neuron* 102.2 (2019), pp. 280–293. DOI: https://doi.org/10.1016/j.neuron.2019.02.039.

[192] Acunzo, David J., MacKenzie, Graham, and van Rossum, Mark C.W. 'Systematic biases in early ERP and ERF components as a result of high-pass filtering.' In: *Journal of Neuroscience Methods* 209.1 (2012), pp. 212–218. DOI: https://doi.org/10.1016/j.jneumeth.2012.06.011.

[193] Jas, Mainak, Engemann, Denis A., Bekhti, Yousra, Raimondo, Federico, and Gramfort, Alexandre. 'Autoreject: Automated artifact rejection for MEG and EEG data.' In: *NeuroImage* 159 (2017), pp. 417–429. DOI: https://doi.org/10.1016/j.neuroimage.2017.06.030.

[194] Kay, Matthew, Nelson, Gregory L, and Hekler, Eric B. 'Researcher-centered design of statistics: Why Bayesian statistics better fit the culture and incentives of HCI.' In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 4521–4532.

[195] Kay, Matthew, Haroz, Steve, Guha, Shion, and Dragicevic, Pierre. 'Special interest group on transparent statistics in hci.' In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2016, pp. 1081–1084.

[196] Urbaniak, Rafal, Tempska, Patrycja, Dowgiałło, Maria, Ptaszyński, Michał, Fortuna, Marcin, Marcińczuk, Michał, Piesiewicz, Jan, Leliwa, Gniewosz, Soliwoda, Kamil, Dziublewska, Ida, Sulzhytskaya, Nataliya, Karnicka, Aleksandra, Skrzek, Paweł, Karbowska, Paula, Brochocki, Maciej, and Wroczyński, Michał. 'Namespotting: Username toxicity and actual toxic behavior on Reddit.' In: *Computers in Human Behavior* 136 (2022), p. 107371. DOI: https://doi.org/10.1016/j.chb.2022.107371.

[197] Ackermans, Kevin, Rusman, Ellen, Nadolski, Rob, Specht, Marcus, and Brand-Gruwel, Saskia. 'Video-or text-based rubrics: What is most effective for mental model growth of complex skills within formative assessment in secondary schools?' In: *Computers in Human Behavior* 101 (2019), pp. 248–258. DOI: https://doi.org/10.1016/j.chb.2019.07.011.

[198] Gueron-Sela, Noa, Shalev, Ido, Gordon-Hacker, Avigail, Egotubov, Alisa, and Barr, Rachel. 'Screen media exposure and behavioral adjustment in early childhood during and after COVID-19 home lockdown periods.' In: *Computers in Human Behavior* 140 (2023), p. 107572. DOI: https://doi.org/10.1016/j.chb.2022.107572.

[199] Weller, Joshua A, King, Marcie L, Figner, Bernd, and Denburg, Natalie L. 'Information use in risky decision making: Do age differences depend on affective context?' In: *Psychology and aging* 34.7 (2019), p. 1005.

[200] Somerville, Leah H, Haddara, Nadia, Sasse, Stephanie F, Skwara, Alea C, Moran, Joseph M, and Figner, Bernd. 'Dissecting "peer presence" and "decisions" to deepen understanding of peer influence on adolescent risky choice.' In: *Child development* 90.6 (2019), pp. 2086–2103.

[201] Schad, Daniel J, Betancourt, Michael, and Vasishth, Shravan. 'Toward a principled Bayesian workflow in cognitive science.' In: *Psychological methods* 26.1 (2021), p. 103.

[202] Schoot, Rens van de, Depaoli, Sarah, King, Ruth, Kramer, Bianca, Märtens, Kaspar, Tadesse, Mahlet G, Vannucci, Marina, Gelman, Andrew, Veen, Duco, Willemsen, Joukje, et al. 'Bayesian statistics and modelling.' In: *Nature Reviews Methods Primers* 1.1 (2021), pp. 1–26.

[203] Dix, Alan. 'Bayesian statistics.' In: *Bayesian Methods for Interaction and Design* (2022), p. 81.

[204] Bürkner, Paul-Christian. 'brms: An R package for Bayesian multilevel models using Stan.' In: *Journal of statistical software* 80.1 (2017), pp. 1–28.

[205] Carpenter, Bob, Gelman, Andrew, Hoffman, Matthew D, Lee, Daniel, Goodrich, Ben, Betancourt, Michael, Brubaker, Marcus, Guo, Jiqiang, Li, Peter, and Riddell, Allen. 'Stan: A probabilistic programming language.' In: *Journal of statistical software* 76.1 (2017).

[206] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2021.

[207] Wickham, Hadley, Averick, Mara, Bryan, Jennifer, Chang, Winston, McGowan, Lucy D'Agostino, François, Romain, Grolemund, Garrett, Hayes, Alex, Henry, Lionel, Hester, Jim, Kuhn, Max, Pedersen, Thomas Lin, Miller, Evan, Bache, Stephan Milton, Müller, Kirill, Ooms, Jeroen, Robinson, David, Seidel, Dana Paige, Spinu, Vitalie, Takahashi, Kohske, Vaughan, Davis, Wilke, Claus, Woo, Kara, and Yutani, Hiroaki. 'Welcome to the tidyverse.' In: *Journal of Open Source Software* 4.43 (2019), p. 1686. DOI: 10.21105/joss.01686.

[208] Bates, Douglas, Maechler, Martin, and Jagan, Mikael. Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.4-1. 2022.

[209] Müller, Kirill. here: A Simpler Way to Find Your Files. R package version 1.0.1. 2020.

[210] Wickham, Hadley, Miller, Evan, and Smith, Danny. haven: Import and Export SPSS, Stata and SAS Files. R package version 2.5.0. 2022.

[211] Wickham, Hadley, Hester, Jim, and Bryan, Jennifer. readr: Read Rectangular Text Data. R package version 2.1.2. 2022.

[212] Bürkner, Paul-Christian. brms: Bayesian Regression Models using Stan. R package version 2.17.0. 2022.

[213] Bürkner, Paul-Christian. 'brms: An R Package for Bayesian Multilevel Models Using Stan.' In: *Journal of Statistical Software* 80.1 (2017), pp. 1–28. DOI: 10.18637/jss.v080.i01.

[214] Bates, Douglas, Mächler, Martin, Bolker, Ben, and Walker, Steve. 'Fitting Linear Mixed-Effects Models Using lme4.' In: *Journal of Statistical Software* 67.1 (2015), pp. 1–48. DOI: 10.18637/jss.v067.i01.

[215] Kuznetsova, Alexandra, Brockhoff, Per B., and Christensen, Rune H. B. 'lmerTest Package: Tests in Linear Mixed Effects Models.' In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26. DOI: 10.18637/jss.v082.i13.

[216] Revelle, William. psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.2.5. 2022.

[217] Eddelbuettel, Dirk and François, Romain. 'Rcpp: Seamless R and C++ Integration.' In: *Journal of Statistical Software* 40.8 (2011), pp. 1–18. DOI: 10.18637/jss.v040.i08.

[218] Kay, Matthew. tidybayes: Tidy Data and Geoms for Bayesian Models. R package version 3.0.2. 2022.

[219] Gromer, Daniel. apa: Format Outputs of Statistical Tests According to APA Guidelines. R package version 0.3.3. 2020.

[220] Makowski, Dominique, Lüdecke, Daniel, Ben-Shachar, Mattan S., Patil, Indrajeet, Wilson, Michael D., and Wiernik, Brenton M. bayestestR: Understand and Describe Bayesian Models and Posterior Distributions. R package version 0.12.1. 2022.

[221] Wickham, Hadley. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

[222] Xie, Yihui. Dynamic Documents with R and knitr. 2nd. Isbn 978-1498716963. Chapman and Hall/CRC, 2015.

[223] Chang, Winston. extrafont: Tools for Using Fonts. R package version 0.18. 2022.

[224] Pedersen, Thomas Lin. patchwork: The Composer of Plots. R package version 1.1.1. 2020.

[225] Lüdecke, Daniel, Patil, Indrajeet, Ben-Shachar, Mattan S., Wiernik, Brenton M., Waggoner, Philip, and Makowski, Dominique. 'see: An R Package for Visualizing Statistical Models.' In: *Journal of Open Source Software* 6.64 (2021), p. 3393. DOI: 10.21105/joss.03393.

[226] Lüdecke, Daniel. sjstats: Collection of Convenient Functions for Common Statistical Computations. R package version 0.18.1. 2021.

[227] Gelman, Andrew, Rubin, Donald B, et al. 'Inference from iterative simulation using multiple sequences.' In: *Statistical science* 7.4 (1992), pp. 457–472.

[228] Vehtari, Aki, Gelman, Andrew, and Gabry, Jonah. 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC.' In: *Statistics and computing* 27.5 (2017), pp. 1413–1432.

[229] Hoijtink, Herbert and Schoot, Rens van de. 'Testing small variance priors using prior-posterior predictive p values.' In: *Psychological Methods* 23.3 (2018), p. 561.

[230] Meng, Xiao-Li et al. 'Posterior predictive p-values.' In: *The Annals of Statistics* 22.3 (1994), pp. 1142–1160.

[231] Shi, Haolun and Yin, Guosheng. 'Reconnecting p-value and Posterior Probability under One-and Two-sided Tests.' In: *The American Statistician* (2020), pp. 1–11.

[232] Judd, Charles M, Westfall, Jacob, and Kenny, David A. 'Experiments with more than one random factor: Designs, analytic models, and statistical power.' In: *Annual review of psychology* 68 (2017), pp. 601–625.

[233] Hedges, Larry V. 'Effect sizes in cluster-randomized designs.' In: (2007).

[234] Beedie, Christopher J, Benedetti, Fabrizio, Barbiani, Diletta, Camerone, Eleanora, Lindheimer, Jacob, and Roelands, Bart. 'Incorporating methods and findings from neuroscience to better understand placebo and nocebo effects in sport.' In: *European Journal of Sport Science* 20.3 (2019), pp. 313–325.

[235] Oken, Barry S., Flegal, Kristin, Zajdel, Daniel, Kishiyama, Shirley, Haas, Mitchell, and Peters, Dawn. 'Expectancy effect: Impact of pill administration on cognitive performance in healthy seniors.' In: *Journal of Clinical and Experimental Neuropsychology* 30.1 (2007), pp. 7–17.

[236] Beedie, Christopher J., M., Stuart Elizabeth, Colemean, Damian A., and Foad, Abigail J. 'Placebo Effects of Caffeine on Cycling Performance.' In: *Medicine & Science in Sports &p; Exercise* 38.12 (2006), pp. 2159–2164.

[237] Weger, Ulrich W and Loughnan, Stephen. 'Rapid communication: Mobilizing unused resources: Using the placebo concept to enhance cognitive performance.' In: *Quarterly Journal of Experimental Psychology* 66.1 (2013), pp. 23–28.

[238] Rozenkrantz, Liron, Mayo, Avraham E, Ilan, Tomer, Hart, Yuval, Noy, Lior, and Alon, Uri. 'Placebo can enhance creativity.' In: *PloS one* 12.9 (2017), e0182466.

[239] Price, Donald D., Finniss, Damien G., and Benedetti, Fabrizio. 'A comprehensive review of the placebo effect: recent advances and current thought.' In: *Annual review of psychology* 59 (2008), pp. 565–90.

[240] Gray, Heather M., Ambady, Nalini, Lowenthal, William T., and Deldin, Patricia. 'P300 as an index of attention to self-relevant stimuli.' In: *Journal of Experimental Social Psychology* 40.2 (2004), pp. 216–224. DOI: https://doi.org/10.1016/S0022-1031(03)00092-1.

[241] Wang, Lei, Zheng, Jiehui, Huang, Shenwei, and Sun, Haoye. 'P300 and Decision Making under Risk and Ambiguity.' In: *Computational Intelligence and Neuroscience* 2015 (2015).

[242] Kirsch, Irving. 'Response expectancy and the placebo effect.' In: *International review of neurobiology* 138 (2018), pp. 81–93.

[243] Cave, Stephen, Coughlan, Kate, and Dihal, Kanta. '"Scary Robots": Examining Public Responses to AI.' In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. AIES '19. Association for Computing Machinery, 2019, 331–337. DOI: 10.1145/3306618.3314232.

[244] King, William R. and He, Jun. 'A meta-analysis of the technology acceptance model.' In: *Information & Management* 43.6 (2006), pp. 740–755. DOI: https://doi.org/10.1016/j.im.2006.05.003.

[245] Stewart-Williams, Steve and Podd, John. 'The placebo effect: dissolving the expectancy versus conditioning debate.' In: *Psychological bulletin* 130.2 (2004), p. 324.

[246] Jong, Peter J de, Baast, Robert van, Arntz, Amoud, and Merckelbach, Harald. 'The placebo effect in pain reduction: the influence of conditioning experiences and response expectancies.' In: *International journal of behavioral medicine* 3.1 (1996), pp. 14–29. DOI: `https://doi.org/10.1207/s15327558ijbm0301_2`.

[247] Denisova, Alena and Cook, Eliott. 'Power-Ups in Digital Games: The Rewarding Effect of Phantom Game Elements on Player Experience.' In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. Chi Play '19. Association for Computing Machinery, 2019, 161–168. DOI: `10.1145/3311350.3347173`.

[248] Denisova, Alena and Cairns, Paul. 'The placebo effect in digital games: Phantom perception of adaptive artificial intelligence.' In: *Proceedings of the 2015 annual symposium on computer-human interaction in play*. 2015, pp. 23–33.

[249] Vaccaro, Kristen, Huang, Dylan, Eslami, Motahhare, Sandvig, Christian, Hamilton, Kevin, and Karahalios, Karrie. 'The illusion of control: Placebo effects of control settings.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–13.

[250] Page, Matthew J., McKenzie, Joanne E., Bossuyt, Patrick M., Boutron, Isabelle, Hoffmann, Tammy C., Mulrow, Cynthia D., Shamseer, Larissa, Tetzlaff, Jennifer M., Akl, Elie A., Brennan, Sue E., Chou, Roger, Glanville, Julie, Grimshaw, Jeremy M., Hróbjartsson, Asbjørn, Lalu, Manoj M., Li, Tianjing, Loder, Elizabeth W., Mayo-Wilson, Evan, McDonald, Steve, McGuinness, Luke A., Stewart, Lesley A., Thomas, James, Tricco, Andrea C., Welch, Vivian A., Whiting, Penny, and Moher, David. 'The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.' In: *International Journal of Surgery* 88 (2021), p. 105906. DOI: `https://doi.org/10.1016/j.ijsu.2021.105906`.

[251] Dumas, Guillaume, Martinerie, Jacques, Soussignan, Robert, and Nadel, Jacqueline. 'Does the brain know who is at the origin of what in an imitative interaction?' In: *Frontiers in human neuroscience* 6 (2012), p. 128.

[252] Software, Swiss Academic. Citavi. The only all-in-one writing and referencing solution. `https://www.citavi.com/en`. 2020-10-29.

[253] Coyle, David, Moore, James, Kristensson, Per Ola, Fletcher, Paul, and Blackwell, Alan. 'I Did That! Measuring Users' Experience of Agency in Their Own Actions.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Association for Computing Machinery, 2012, 2025–2034. DOI: `10.1145/2207676.2208350`.

[254] Kasahara, Shunichi, Takada, Kazuma, Nishida, Jun, Shibata, Kazuhisa, Shimojo, Shinsuke, and Lopes, Pedro. 'Preserving Agency During Electrical Muscle Stimulation Training Speeds up Reaction Time Directly After Removing EMS.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445147.

[255] Tajima, Daisuke, Nishida, Jun, Lopes, Pedro, and Kasahara, Shunichi. 'Whose Touch is This?: Understanding the Agency Trade-Off Between User-Driven Touch vs. Computer-Driven Touch.' In: *ACM Trans. Comput.-Hum. Interact.* 29.3 (2022). DOI: 10.1145/3489608.

[256] Shahu, Ambika, Wintersberger, Philipp, and Michahelles, Florian. 'Scenario-Based Investigation of Acceptance of Electric Muscle Stimulation.' In: *Proceedings of the Augmented Humans International Conference 2022*. AHs '22. Association for Computing Machinery, 2022, 184–194. DOI: 10.1145/3519391.3519416.

[257] Venot, Tristan, Desbois, Arthur, Corsi, Marie Constance, Hugueville, Laurent, Saint-Bauzel, Ludovic, and De Vico Fallani, Fabrizio. 'Exploring strategies for multimodal BCIs in an enriched environment.' In: *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*. IEEE, 2022, pp. 685–690. DOI: 10.1109/MetroXRAINE54828.2022.9967624.

[258] Zolyomi, Annuska and Snyder, Jaime. 'Early Adopters of a Low Vision Head-Mounted Assistive Technology.' In: *SIGACCESS Access. Comput.* (122 2020), p. 1. DOI: 10.1145/3386410.3386412.

[259] Cinel, Caterina, Valeriani, Davide, and Poli, Riccardo. 'Neurotechnologies for human cognitive augmentation: current state of the art and future prospects.' In: *Frontiers in human neuroscience* 13 (2019), p. 13.

[260] Schneegass, Christina, Wilson, Max L, Maior, Horia A, Chiossi, Francesco, Cox, Anna L, and Wiese, Jason. 'The Future of Cognitive Personal Informatics.' In: *Proceedings of the 25th International Conference on Mobile Human-Computer Interaction*. 2023, pp. 1–5.

[261] Schneegass, Christina, Wilson, Max L, Shaban, Jwan, Niess, Jasmin, Chiossi, Francesco, Mitrevska, Teodora, and Woźniak, Paweł W. 'Broadening the mind: how emerging neurotechnology is reshaping HCI and interactive system design.' In: *i-com* 0 (2024). DOI: 10.1515/icom-2024-0007.

[262] Martinez, Wendy, Benerradi, Johann, Midha, Serena, Maior, Horia A., and Wilson, Max L. 'Understanding the Ethical Concerns for Neurotechnology in the Future of Work.' In: *2022 Symposium on Human-Computer Interaction for Work*. CHIWORK 2022. Association for Computing Machinery, 2022. DOI: 10.1145/3533406.3533423.

[263] Hougaard, Bastian Ilsø, Rossau, Ingeborg Goll, Czapla, Jedrzej Jacek, Mikó, Mózes Adorján, Skammelsen, Rasmus Bugge, Knoche, Hendrik, and Jochumsen, Mads.

'Who Willed It? Decreasing Frustration by Manipulating Perceived Control through Fabricated Input for Stroke Rehabilitation BCI Games.' In: *Proc. ACM Hum.-Comput. Interact.* 5.CHI PLAY (2021). DOI: 10.1145/3474662.

[264] Hougaard, Bastian I., Knoche, Hendrik, Kristensen, Mathias Sand, and Jochumsen, Mads. 'Modulating Frustration and Agency Using Fabricated Input for Motor Imagery BCIs in Stroke Rehabilitation.' In: *IEEE Access* 10 (2022), pp. 72312–72327. DOI: 10.1109/ACCESS.2022.3188103.

[265] Mercado-García, V. R., Alonso-Valerdi, L. M., Salas-Garza, M., López-Montoya, A., Garza-Ibarra, E., and Chávez-Madero, C. 'Do User-Centered Designed Paradigms for BCIs improve the Modulation of EEG Signals?' In: *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2021, pp. 499–502. DOI: 10.1109/NER49283.2021.9441423.

[266] Raisamo, Roope, Rakkolainen, Ismo, Majaranta, Päivi, Salminen, Katri, Rantala, Jussi, and Farooq, Ahmed. 'Human augmentation: Past, present and future.' In: *International Journal of Human-Computer Studies*. 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies 131 (2019), pp. 131–143. DOI: 10.1016/j.ijhcs.2019.05.008.

[267] Mueller, Florian Floyd, Semertzidis, Nathan, Andres, Josh, Marshall, Joe, Benford, Steve, Li, Xiang, Matjeka, Louise, and Mehta, Yash. 'Towards Understanding the Design of Intertwined Human-Computer Integrations.' In: *ACM Trans. Comput.-Hum. Interact.* (2023). Just Accepted. DOI: 10.1145/3590766.

[268] Gauthier, Baptiste, Albert, Louis, Martuzzi, Roberto, Herbelin, Bruno, and Blanke, Olaf. 'Virtual Reality Platform for Functional Magnetic Resonance Imaging in Ecologically Valid Conditions.' In: *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. VRST '21. Association for Computing Machinery, 2021. DOI: 10.1145/3489849.3489894.

[269] Zhang, Jingjing, Huang, Mengjie, Tang, Xiaohang, Wang, Yiqi, and Yang, Rui. 'How Virtual Body Continuity with Different Hand Representations Influence on User Perceptions and Task Performance.' In: *2022 15th International Conference on Human System Interaction (HSI)*. IEEE, 2022, pp. 1–5. DOI: 10.1109/HSI55341.2022.9869486.

[270] Jun, Joohee, Jung, Myeongul, Kim, So-Yeon, and Kim, Kwanguk (Kenny). 'Full-Body Ownership Illusion Can Change Our Emotion.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Association for Computing Machinery, 2018, 1–11. DOI: 10.1145/3173574.3174175.

[271] Seinfeld, Sofia, Feuchtner, Tiare, Pinzek, Johannes, and Müller, Jörg. 'Impact of Information Placement and User Representations in VR on Performance and Embodiment.' In: *IEEE Transactions on Visualization and Computer Graphics* 28.3 (2022), pp. 1545–1556. DOI: 10.1109/TVCG.2020.3021342.

[272] Takada, Kazuma, Kawaguchi, Midori, Uehara, Akira, Nakanishi, Yukiya, Armstrong, Mark, Verhulst, Adrien, Minamizawa, Kouta, and Kasahara, Shunichi. 'Parallel Ping-Pong: Exploring Parallel Embodiment through Multiple Bodies by a Single User.' In: *Proceedings of the Augmented Humans International Conference 2022*. AHs '22. Association for Computing Machinery, 2022, 121–130. DOI: 10.1145/3519391.3519408.

[273] Miura, Reiji, Kasahara, Shunichi, Kitazaki, Michiteru, Verhulst, Adrien, Inami, Masahiko, and Sugimoto, Maki. 'MultiSoma: Distributed Embodiment with Synchronized Behavior and Perception.' In: *Proceedings of the Augmented Humans International Conference 2021*. AHs '21. Association for Computing Machinery, 2021, 1–9. DOI: 10.1145/3458709.3458878.

[274] Sun, Wenxin, Huang, Mengjie, Wu, Chenxin, and Yang, Rui. 'Sense of Agency on Handheld AR for Virtual Object Translation.' In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2022, pp. 838–839. DOI: 10.1109/VRW55335.2022.00270.

[275] Sun, Wenxin, Huang, Mengjie, Wu, Chenxin, and Yang, Rui. 'Sense of Agency in Mid-Air Translation with Head-Mounted AR.' In: *2022 1st IEEE International Conference on Cognitive Aspects of Virtual Reality (CVR)*. IEEE, 2022, pp. 000061–000064. DOI: 10.1109/CVR55417.2022.9967506.

[276] Dingler, Tilman, Agroudy, Passant El, Rzayev, Rufat, Lischke, Lars, Machulla, Tonja, and Schmidt, Albrecht. 'Memory Augmentation Through Lifelogging: Opportunities and Challenges.' In: *Technology-Augmented Perception and Cognition*. Springer International Publishing, 2021, pp. 47–69. DOI: 10.1007/978-3-030-30457-7_3.

[277] Xu, Xuhai, Yu, Anna, Jonker, Tanya R., Todi, Kashyap, Lu, Feiyu, Qian, Xun, Evangelista Belo, João Marcelo, Wang, Tianyi, Li, Michelle, Mun, Aran, Wu, Te-Yen, Shen, Junxiao, Zhang, Ting, Kokhlikyan, Narine, Wang, Fulton, Sorenson, Paul, Kim, Sophie, and Benko, Hrvoje. 'XAIR: A Framework of Explainable AI in Augmented Reality.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581500.

[278] Wang, Qiaosi, Camacho, Ida, Jing, Shan, and Goel, Ashok K. 'Understanding the Design Space of AI-Mediated Social Interaction in Online Learning: Challenges and Opportunities.' In: *Proc. ACM Hum.-Comput. Interact.* 6.CSCW1 (2022). DOI: 10.1145/3512977.

[279] Thieme, Anja, Hanratty, Maryann, Lyons, Maria, Palacios, Jorge, Marques, Rita Faia, Morrison, Cecily, and Doherty, Gavin. 'Designing Human-Centered AI for Mental Health: Developing Clinically Relevant Applications for Online CBT Treatment.' In: *ACM Trans. Comput.-Hum. Interact.* 30.2 (2023). DOI: 10.1145/3564752.

[280] Sali, Serdar, Liu, Ronald, Wardrip-Fruin, Noah, Mateas, Michael, and Kurniawan, Sri. 'Getting a Word in: Adding Artificial Pauses to Natural Language Interaction.' In: *Proceedings of the International Conference on the Foundations of Digital Games*. FDG '12. Association for Computing Machinery, 2012, 57–64. DOI: 10.1145/2282338.2282353.

[281] Sun, Yuan, Song, Qiurong, Gui, Xinning, Ma, Fenglong, and Wang, Ting. 'AutoML in The Wild: Obstacles, Workarounds, and Expectations.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581082.

[282] Ahmad, Imtiaz, Akter, Taslima, Buher, Zachary, Farzan, Rosta, Kapadia, Apu, and Lee, Adam J. 'Tangible Privacy for Smart Voice Assistants: Bystanders' Perceptions of Physical Device Controls.' In: *Proc. ACM Hum.-Comput. Interact.* 6.CSCW2 (2022). DOI: 10.1145/3555089.

[283] Semertzidis, Nathan Arthur, Li Pin Hiung, Annaelle, Vranic-Peters, Michaela Jayne, and Mueller, Florian 'Floyd'. 'Dozer: Towards Understanding the Design of Closed-Loop Wearables for Sleep.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581044.

[284] Friedman, Batya and Kahn, Peter H. 'Human agency and responsible computing: Implications for computer system design.' In: *Journal of Systems and Software* 17.1 (1992). Computer Ethics, pp. 7–14. DOI: https://doi.org/10.1016/0164-1212(92)90075-U.

[285] Yun, Sonmin, Wen, Wen, An, Qi, Hamasaki, Shunsuke, Yamakawa, Hiroshi, Tamura, Yusuke, Yamashita, Atsushi, and Asama, Hajime. 'Investigating the Relationship Between Assisted Driver's SoA and EEG.' In: *Converging Clinical and Engineering Research on Neurorehabilitation III*. Ed. by Lorenzo Masia, Silvestro Micera, Metin Akay, and José L. Pons. Biosystems & Biorobotics. Springer International Publishing, 2019, pp. 1039–1043. DOI: 10.1007/978-3-030-01845-0_208.

[286] Chen, Yu, Chaudhary, Shefali, and Li, Chiang-Shan R. 'Shared and distinct neural activity during anticipation and outcome of win and loss: A meta-analysis of the monetary incentive delay task.' In: *Neuroimage* 264 (2022), p. 119764.

[287] Buchholz, Verena N., David, Nicole, Sengelmann, Malte, and Engel, Andreas K. 'Belief of agency changes dynamics in sensorimotor networks.' In: *Scientific Reports*

9.1 (2019). Number: 1 Publisher: Nature Publishing Group, p. 1995. DOI: 10.1038/s41598-018-37912-w.

[288] Freeman, Frederick G, Mikulka, Peter J, Prinzel, Lawrence J, and Scerbo, Mark W. 'Evaluation of an adaptive automation system using three EEG indices with a visual tracking task.' In: *Biological psychology* 50.1 (1999), pp. 61–76. DOI: 10.1016/S0301-0511(99)00002-2.

[289] Prinzel III, Lawrence J, Scerbo, Mark W, Freeman, Frederick G, and Mikulka, Peter J. 'A bio-cybernetic system for adaptive automation.' In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 39. 21. SAGE Publications Sage CA: Los Angeles, CA. 1995, pp. 1365–1369. DOI: 10.1177/154193129503902102.

[290] Kosch, Thomas, Karolus, Jakob, Zagermann, Johannes, Reiterer, Harald, Schmidt, Albrecht, and Woźniak, Paweł W. 'A Survey on Measuring Cognitive Workload in Human-Computer Interaction.' In: *ACM Comput. Surv.* 55.13s (2023). DOI: 10.1145/3582272.

[291] Salvi, Carola and Bowden, Edward M. 'Looking for Creativity: Where Do We Look When We Look for New Ideas?' In: *Frontiers in Psychology* 7 (2016). DOI: 10.3389/fpsyg.2016.00161.

[292] Janaka, Nuwan, Haigh, Chloe, Kim, Hyeongcheol, Zhang, Shan, and Zhao, Shengdong. 'Paracentral and near-peripheral visualizations: Towards attention-maintaining secondary information presentation on OHMDs during in-person social interactions.' In: *CHI Conference on Human Factors in Computing Systems*. ACM, 2022, pp. 1–14. DOI: 10.1145/3491102.3502127.

[293] Sweeney, Emma and Hua, Zhu. 'Discourse Completion Tasks.' In: *Research Methods in Intercultural Communication*. Section: 14 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781119166283.ch14. John Wiley & Sons, Ltd, 2015, pp. 212–222. DOI: 10.1002/9781119166283.ch14.

[294] Billmyer, K. 'Investigating instrument-based pragmatic variability: effects of enhancing discourse completion tests.' In: *Applied Linguistics* 21.4 (2000), pp. 517–552. DOI: 10.1093/applin/21.4.517.

[295] Golato, A. 'Studying Compliment Responses: A Comparison of DCTs and Recordings of Naturally Occurring Talk.' In: *Applied Linguistics* 24.1 (2003), pp. 90–121. DOI: 10.1093/applin/24.1.90.

[296] Roever, Carsten. 'Validation of a web-based test of ESL pragmalinguistics.' In: *Language Testing - LANG TEST* 23 (2006), pp. 229–256. DOI: 10.1191/0265532206lt329oa.

[297] Sara, Ghiat. 'Integrating Dynamic Assessment in Oral Courses to Develop EFL Students' Pragmatic Competence The Case of Third Year Students at the Department of English Batna2 University. Doctorat thesis, (2022) Université de Batna 2.' In: (2022).

[298] Nikpoo, Iman and Allami, Hamid. 'Exploring (Im)Politeness Strategies in Indirect Reports in Persian across Genders.' In: *Education Research International* 2024 (2024). Publisher: Hindawi, e4642380. DOI: 10.1155/2024/4642380.

[299] Hart, Sandra G. 'Nasa-Task Load Index (NASA-TLX); 20 Years Later.' In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.9 (2006), pp. 904–908. DOI: 10.1177/154193120605000909. eprint: https://doi.org/10.1177/154193120605000909.

[300] Brooke, John. 'SUS: A quick and dirty usability scale.' In: *Usability Eval in Industry* 189 (1995).

[301] Brooke, John. 'SUS: A Retrospective.' In: *J. Usability Studies* 8.2 (2013), 29–40.

[302] Ye, Le, Wang, Zhixuan, Liu, Ying, Chen, Peiyu, Li, Heyi, Zhang, Hao, Wu, Meng, He, Wei, Shen, Linxiao, Zhang, Yihan, Tan, Zhichao, Wang, Yangyuan, and Huang, Ru. 'The Challenges and Emerging Technologies for Low-Power Artificial Intelligence IoT Systems.' In: *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I-REGULAR PAPERS* 68.12 (2021). Num Pages: 14 Place: Piscataway Publisher: Ieee-Inst Electrical Electronics Engineers Inc Web of Science ID: WOS:000724482800004, pp. 4821–4834. DOI: 10.1109/TCSI.2021.3095622.

[303] Wu, Yawen and Hu, Jingtong. 'Towards Independent On-device Artificial Intelligence.' In: *2022 IEEE COMPUTER SOCIETY ANNUAL SYMPOSIUM ON VLSI (ISVLSI 2022)*. Num Pages: 1 Series Title: IEEE Computer Society Annual Symposium on VLSI Proceedings Web of Science ID: WOS:000886230500064. IEEE, 2022, pp. 351–351. DOI: 10.1109/ISVLSI54635.2022.00077.

[304] Vantrepotte, Quentin, Berberian, Bruno, Pagliari, Marine, and Chambon, Valérian. 'Leveraging human agency to improve confidence and acceptability in human-machine interactions.' In: *Cognition* 222 (2022), p. 105020. DOI: 10.1016/j.cognition.2022.105020.

[305] Pagliari, Marine, Chambon, Valérian, and Berberian, Bruno. 'What is new with Artificial Intelligence? Human–agent interactions through the lens of social agency.' In: *Frontiers in Psychology* 13 (2022). DOI: 10.3389/fpsyg.2022.954444.

[306] Schmidmaier, Matthias, Rupp, Jonathan, Cvetanova, Darina, and Mayer, Sven. 'Perceived Empathy of Technology Scale (PETS): Measuring Empathy of Systems Toward the User.' In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '24. Association for Computing Machinery, 2024, pp. 1–18. DOI: 10.1145/3613904.3642035.

[307] Woźniak, Paweł W., Karolus, Jakob, Lang, Florian, Eckerth, Caroline, Schöning, Johannes, Rogers, Yvonne, and Niess, Jasmin. 'Creepy Technology:What Is It and How Do You Measure It?' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Association for Computing Machinery, 2021, pp. 1–13. DOI: 10.1145/3411764.3445299.

[308] Bentvelzen, Marit, Niess, Jasmin, Woźniak, Mikołaj P., and Woźniak, Paweł W. 'The Development and Validation of the Technology-Supported Reflection Inventory.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Association for Computing Machinery, 2021, pp. 1–8. DOI: 10.1145/3411764.3445673.

[309] Boateng, Godfred O., Neilands, Torsten B., Frongillo, Edward A., Melgar-Quiñonez, Hugo R., and Young, Sera L. 'Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer.' In: *Frontiers in Public Health* 6 (2018), p. 149. DOI: 10.3389/fpubh.2018.00149.

[310] Davis, Fred D. 'A technology acceptance model for empirically testing new end-user information systems : theory and results.' Accepted: 2005-08-08T16:11:32Z. Thesis. Massachusetts Institute of Technology, 1985.

[311] Venkatesh, Viswanath, Morris, Michael G., Davis, Gordon B., and Davis, Fred D. 'User Acceptance of Information Technology: Toward a Unified View.' In: *MIS Quarterly* 27.3 (2003). Publisher: Management Information Systems Research Center, University of Minnesota, pp. 425–478. DOI: 10.2307/30036540.

[312] Madigan, Ruth, Louw, Tyron, Wilbrink, Marc, Schieben, Anna, and Merat, Natasha. 'What influences the decision to use automated public transport? Using UTAUT to understand public acceptance of automated road transport systems.' In: *Transportation Research Part F: Traffic Psychology and Behaviour* 50 (2017), pp. 55–64. DOI: 10.1016/j.trf.2017.07.007.

[313] Maillet, Éric, Mathieu, Luc, and Sicotte, Claude. 'Modeling factors explaining the acceptance, actual use and satisfaction of nurses using an Electronic Patient Record in acute care settings: An extension of the UTAUT.' In: *International Journal of Medical Informatics* 84.1 (2015), pp. 36–47. DOI: 10.1016/j.ijmedinf.2014.09.004.

[314] Venkatesh, Viswanath, Thong, James Y. L., and Xu, Xin. 'Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology.' In: *MIS Quarterly* 36.1 (2012). Publisher: Management Information Systems Research Center, University of Minnesota, pp. 157–178. DOI: 10.2307/41410412.

[315] Schrepp, Martin, Thomaschewski, Jörg, and Hinderks, Andreas. 'Construction of a Benchmark for the User Experience Questionnaire (UEQ).' In: *International Journal*

of Interactive Multimedia and Artificial Intelligence 4.Regular Issue (2017), pp. 40–44. DOI: 10.9781/ijimai.2017.445.

[316] Bangor, Aaron, Kortum, Philip T., and Miller, James T. 'An Empirical Evaluation of the System Usability Scale.' In: *International Journal of Human–Computer Interaction* 24.6 (2008). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447310802205776, pp. 574–594. DOI: 10.1080/10447310802205776.

[317] Lewis, James R. 'The System Usability Scale: Past, Present, and Future.' In: *International Journal of Human–Computer Interaction* 34.7 (2018). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10447318.2018.1455307, pp. 577–590. DOI: 10.1080/10447318.2018.1455307.

[318] Minge, Michael, Thüring, Manfred, Wagner, Ingmar, and Kuhr, Carina V. 'The meCUE Questionnaire: A Modular Tool for Measuring User Experience.' In: *Advances in Ergonomics Modeling, Usability & Special Populations*. Ed. by Marcelo Soares, Christianne Falcão, and Tareq Z. Ahram. Springer International Publishing, 2017, pp. 115–128. DOI: 10.1007/978-3-319-41685-4_11.

[319] Peters, Dorian, Calvo, Rafael A., and Ryan, Richard M. 'Designing for Motivation, Engagement and Wellbeing in Digital Experience.' In: *Frontiers in Psychology* 9 (2018). Publisher: Frontiers. DOI: 10.3389/fpsyg.2018.00797.

[320] Wynd, Christine A., Schmidt, Bruce, and Schaefer, Michelle Atkins. 'Two Quantitative Approaches for Estimating Content Validity.' In: *Western Journal of Nursing Research* 25.5 (2003). Publisher: SAGE Publications Inc, pp. 508–518. DOI: 10.1177/0193945903252998.

[321] Polit, Denise F. and Beck, Cheryl Tatano. 'The content validity index: Are you sure you know what's being reported? critique and recommendations.' In: *Research in Nursing & Health* 29.5 (2006). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nur.20147, pp. 489–497. DOI: 10.1002/nur.20147.

[322] Streiner, David L. 'Figuring Out Factors: The Use and Misuse of Factor Analysis.' In: *The Canadian Journal of Psychiatry* 39 (1994), pp. 135–140.

[323] Horn, John Louis. 'A rationale and test for the number of factors in factor analysis.' In: *Psychometrika* 30 (1965), pp. 179–185.

[324] Cattell, Raymond Bernard and Vogelmann, Silke. 'A Comprehensive Trial Of The Scree And Kg Criteria For Determining The Number Of Factors.' In: *Multivariate behavioral research* 12 3 (1977), pp. 289–325.

[325] Crawford, Charles. 'A comparison of the direct oblimin and primary parsimony methods of oblique rotation.' In: *British Journal of Mathematical and Statistical Psychology* 28.2 (1975), pp. 201–213.

[326] Cortina, Jose M. 'What is coefficient alpha? An examination of theory and applications.' In: *Journal of applied psychology* 78.1 (1993), p. 98.

[327] Rosseel, Yves. 'lavaan: An R package for structural equation modeling.' In: *Journal of statistical software* 48 (2012), pp. 1–36.

[328] Moradbakhti, Laura, Leichtmann, Benedikt, and Mara, Martina. 'Development and Validation of a Basic Psychological Needs Scale for Technology Use.' In: *Psychological Test Adaptation and Development* 5.1 (2024). Publisher: Hogrefe Publishing, pp. 26–45. DOI: 10.1027/2698-1866/a000062.

[329] Rammstedt, Beatrice and John, Oliver P. 'Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German.' In: *Journal of Research in Personality* 41.1 (2007), pp. 203–212. DOI: 10.1016/j.jrp.2006.02.001.

[330] Wilson, Matthew L., Huggins-Manley, Anne Corinne, Ritzhaupt, Albert D., and Ruggles, Krista. 'Development of the Abbreviated Technology Anxiety Scale (ATAS).' In: *Behavior Research Methods* 55.1 (2023), pp. 185–199. DOI: 10.3758/s13428-022-01820-9.

[331] Bland, J Martin and Altman, Douglas G. 'Measuring agreement in method comparison studies.' In: *Statistical Methods in Medical Research* 8 (1999), pp. 135–160.

[332] Welsch, Robin, Schmidt, Alexander F., Turner, Daniel, and Rettenberger, Martin. 'Test–Retest Reliability and Temporal Agreement of Direct and Indirect Sexual Interest Measures.' In: *Sexual Abuse* 33.3 (2021). Pmid: 32036763, pp. 339–360. DOI: 10.1177/1079063220904354. eprint: https://doi.org/10.1177/1079063220904354.

[333] Rönkkö, Mikko and Cho, Eunseong. 'An Updated Guideline for Assessing Discriminant Validity.' In: *Organizational Research Methods* 25.1 (2022). Publisher: SAGE Publications Inc, pp. 6–14. DOI: 10.1177/1094428120968614.

[334] Cheung, Gordon W., Cooper-Thomas, Helena D., Lau, Rebecca S., and Wang, Linda C. 'Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations.' In: *Asia Pacific Journal of Management* 41.2 (2024), pp. 745–783. DOI: 10.1007/s10490-023-09871-y.

[335] Schaik, Paul van and Ling, Jonathan. 'An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia.' In: *Interacting with Computers* 23.1 (2011), pp. 18–32. DOI: 10.1016/j.intcom.2010.07.002.

[336] Thong, James Y. L., Venkatesh, Viswanath, Xu, Xin, Hong, Se-Joon, and Tam, Kar Yan. 'Consumer Acceptance of Personal Information and Communication Technology Services.' In: *IEEE Transactions on Engineering Management* 58.4 (2011). Conference

Name: IEEE Transactions on Engineering Management, pp. 613–625. DOI: 10.1109/TEM.2010.2058851.

[337] Bargas-Avila, Javier A. and Hornbæk, Kasper. 'Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Association for Computing Machinery, 2011, pp. 2689–2698. DOI: 10.1145/1978942.1979336.

[338] Kujala, Sari, Roto, Virpi, Väänänen-Vainio-Mattila, Kaisa, Karapanos, Evangelos, and Sinnelä, Arto. 'UX Curve: A method for evaluating long-term user experience.' In: *Interacting with Computers* 23.5 (2011), pp. 473–483. DOI: 10.1016/j.intcom.2011.06.005.

[339] Kujala, Sari and Miron-Shatz, Talya. 'The evolving role of expectations in long-term user experience.' In: *Proceedings of the 19th International Academic Mindtrek Conference*. AcademicMindTrek '15. Association for Computing Machinery, 2015, pp. 167–174. DOI: 10.1145/2818187.2818271.

[340] Wyatt, Sally. 'Technological determinism is dead; long live technological determinism.' In: *The handbook of science and technology studies* 3 (2008), pp. 165–180.

[341] Syrdal, Dag Sverre, Dautenhahn, Kerstin, Koay, Kheng Lee, and Walters, Michael L. 'The negative attitudes towards robots scale and reactions to robot behaviour in a live human-robot interaction study.' In: *Adaptive and emergent behaviour and complex systems* (2009).

[342] Brown, Susan A., Venkatesh, Viswanath, and Goyal, Sandeep. 'Expectation Confirmation in Information Systems Research: A Test of Six Competing Models.' In: *MIS Quarterly* 38.3 (2014). Publisher: Management Information Systems Research Center, University of Minnesota, 729–A9. DOI: 10.25300/MISQ/2014/38.3.05.

[343] Brown, Susan A., Venkatesh, Viswanath, Kuruzovich, Jason, and Massey, Anne P. 'Expectation confirmation: An examination of three competing models.' In: *Organizational Behavior and Human Decision Processes* 105.1 (2008), pp. 52–66. DOI: 10.1016/j.obhdp.2006.09.008.

[344] Kloft, Agnes Mercedes, Welsch, Robin, Kosch, Thomas, and Villa, Steeven. '"AI enhances our performance, I have no doubt this one will do the same": The Placebo effect is robust to negative descriptions of AI.' In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 2024, pp. 1–24. DOI: 10.1145/3613904.3642633.

[345] Kosch, Thomas, Welsch, Robin, Chuang, Lewis, and Schmidt, Albrecht. 'The Placebo Effect of Artificial Intelligence in Human–Computer Interaction.' In: *ACM Transactions on Computer-Human Interaction* 29.6 (2022), pp. 1–32. DOI: 10.1145/3529225.

[346] Denisova, Alena and Cairns, Paul. 'The Placebo Effect in Digital Games: Phantom Perception of Adaptive Artificial Intelligence.' In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM, 2015, pp. 23–33. DOI: 10.1145/2793107.2793109.

[347] Linxen, Sebastian, Sturm, Christian, Brühlmann, Florian, Cassau, Vincent, Opwis, Klaus, and Reinecke, Katharina. 'How WEIRD is CHI?' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Chi '21. Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445488.

[348] Aguinis, Herman and Bradley, Kyle J. 'Best practice recommendations for designing and implementing experimental vignette methodology studies.' In: *Organizational research methods* 17.4 (2014), pp. 351–371.

[349] Findler, Liora, Vilchinsky, Noa, and Werner, Shirli. 'The multidimensional attitudes scale toward persons with disabilities (MAS) construction and validation.' In: *Rehabilitation Counseling Bulletin* 50.3 (2007), pp. 166–176.

[350] Wobbrock, Jacob O, Findlater, Leah, Gergle, Darren, and Higgins, James J. 'The aligned rank transform for nonparametric factorial analyses using only anova procedures.' In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2011, pp. 143–146.

[351] Yusif, Salifu, Soar, Jeffrey, and Hafeez-Baig, Abdul. 'Older people, assistive technologies, and the barriers to adoption: A systematic review.' In: *International Journal of Medical Informatics* 94 (2016), pp. 112–116. DOI: https://doi.org/10.1016/j.ijmedinf.2016.07.004.

[352] Juengst, Eric and Moseley, Daniel. 'Human enhancement.' In: (2015).

[353] Shneiderman, Ben. 'Human-centered artificial intelligence: Reliable, safe & trustworthy.' In: *International Journal of Human–Computer Interaction* 36.6 (2020), pp. 495–504.

[354] Anderson, Walter Truett. 'Augmentation, symbiosis, transcendence: technology and the future (s) of human identity.' In: *Futures* 35.5 (2003), pp. 535–546.

[355] Sasaki, Tomoya, Saraiji, MHD Yamen, Minamizawa, Kouta, and Inami, Masahiko. 'MetaArms: Body Remapping Using Feet-Controlled Artificial Arms.' In: *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. UIST '18 Adjunct. Association for Computing Machinery, 2018, 140–142. DOI: 10.1145/3266037.3271628.

[356] Findler, Liora, Vilchinsky, Noa, and Werner, Shirli. 'The Multidimensional Attitudes Scale Toward Persons With Disabilities (MAS): Construction and Validation.' In: *Rehabilitation Counseling Bulletin* 50.3 (2007), pp. 166–176. DOI: 10.

1177 / 0034355207050003040l. eprint: https : / / doi . org / 10 . 1177 / 0034355207050003040l.

[357] Moore, James W. 'What Is the Sense of Agency and Why Does it Matter?' In: *Frontiers in Psychology* 7 (2016).

[358] Fiske, Susan T., Cuddy, Amy J. C., Glick, Peter, and Xu, Jun. 'A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition.' In: *Journal of Personality and Social Psychology* 82.6 (2002). DOI: 10.1037/0022-3514.82.6.878.

[359] Fiske, Susan T., Cuddy, Amy J. C., and Glick, Peter. 'Universal dimensions of social cognition: Warmth and competence.' In: *Trends in cognitive sciences* 11.2 (2007). DOI: 10.1016/j.tics.2006.11.005.

[360] Haliburton, Luke, Ghebremedhin, Sinksar, Welsch, Robin, Schmidt, Albrecht, and Mayer, Sven. Investigating Labeler Bias in Face Annotation for Machine Learning. arXiv:2301.09902 [cs]. 2023. DOI: 10.48550/arXiv.2301.09902.

[361] McKee, Kevin, Bai, Xuechunzi, and Fiske, Susan. 'Understanding human impressions of artificial intelligence.' In: (2021).

[362] Bogardus, Emory S. 'A social distance scale.' In: *Sociology & Social Research* (1933).

[363] Corrigan, Patrick W, Edwards, Annette Backs, Green, Amy, Diwan, Sarah Lickey, and Penn, David L. 'Prejudice, social distance, and familiarity with mental illness.' In: *Schizophrenia bulletin* 27.2 (2001), pp. 219–225.

[364] Graham, Jesse, Nosek, Brian A, Haidt, Jonathan, Iyer, Ravi, Spassena, Koleva, and Ditto, Peter H. 'Moral foundations questionnaire.' In: *Journal of Personality and Social Psychology* (2008).

[365] Rousi, Rebekah A. and Renko, Roni. 'Emotions toward cognitive enhancement technologies and the body - Attitudes and willingness to use.' In: *Int. J. Hum. Comput. Stud.* 143 (2020), p. 102472.

[366] Comrey, Andrew L. 'Factor-analytic methods of scale development in personality and clinical psychology.' In: *Journal of consulting and clinical psychology* 56 5 (1988), pp. 754–61.

[367] Zhang, Guangjian and Preacher, Kristopher J. 'Factor Rotation and Standard Errors in Exploratory Factor Analysis.' In: *Journal of Educational and Behavioral Statistics* 40 (2015), pp. 579–603.

[368] Meyer, Bertolt and Asbrock, Frank. 'Disabled or cyborg? How bionics affect stereotypes toward people with physical disabilities.' In: *Frontiers in psychology* 9 (2018), p. 2251.

[369] Schepman, Astrid and Rodway, Paul. 'Initial validation of the general attitudes towards Artificial Intelligence Scale.' In: *Computers in Human Behavior Reports* 1 (2020), p. 100014. DOI: https://doi.org/10.1016/j.chbr.2020.100014.

[370] Lam, Shun Yin, Chiang, Jeongwen, and Parasuraman, A. 'The effects of the dimensions of technology readiness on technology acceptance: An empirical analysis.' In: *Journal of Interactive Marketing* 22.4 (2008), pp. 19–39. DOI: https://doi.org/10.1002/dir.20119.

[371] MacCallum, Robert C., Roznowski, Mary Ann, and Necowitz, Lawrence B. 'Model modifications in covariance structure analysis: the problem of capitalization on chance.' In: *Psychological bulletin* 111 3 (1992), pp. 490–504.

[372] Siedl, Sandra Maria and Mara, Martina. 'What Drives Acceptance of Occupational Exoskeletons? Focus Group Insights from Workers in Food Retail and Corporate Logistics.' In: *International Journal of Human–Computer Interaction* (2022), pp. 1–10.

[373] Bretschneider, Maximilian, Mandl, Sarah, Strobel, Anja, Asbrock, Frank, and Meyer, Bertolt. 'Social perception of embodied digital technologies—a closer look at bionics and social robotics.' In: *Gruppe. Interaktion. Organisation. Zeitschrift Für Angewandte Organisationspsychologie (GIO)* 53.3 (2022), pp. 343–358.

[374] Mandl, Sarah, Bretschneider, Maximilian, Asbrock, Frank, Meyer, Bertolt, and Strobel, Anja. 'The Social Perception of Robots Scale (SPRS): Developing and Testing a Scale for Successful Interaction Between Humans and Robots.' In: *Collaborative Networks in Digitalization and Society 5.0: 23rd IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2022, Lisbon, Portugal, September 19–21, 2022, Proceedings*. Springer. 2022, pp. 321–334.

[375] Parasuraman, A. and Colby, Charles L. 'An Updated and Streamlined Technology Readiness Index: TRI 2.0.' In: *Journal of Service Research* 18.1 (2015), pp. 59–74. DOI: 10.1177/1094670514539730. eprint: https://doi.org/10.1177/1094670514539730.

[376] Parasuraman, A. 'Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies.' In: *Journal of Service Research* 2.4 (2000), pp. 307–320. DOI: 10.1177/109467050024001. eprint: https://doi.org/10.1177/109467050024001.

[377] Profita, Halley, Albaghli, Reem, Findlater, Leah, Jaeger, Paul, and Kane, Shaun K. 'The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use.' In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Association for Computing Machinery, 2016, 4884–4895. DOI: 10.1145/2858036.2858130.

[378] DeVaul, Richard W., Pentland, Alex "Sandy", and Corey, Vicka R. 'The Memory Glasses: Subliminal vs. Overt Memory Support with Imperfect Information.' In: *Proceedings of the 7th IEEE International Symposium on Wearable Computers*. ISWC '03. IEEE Computer Society, 2003, p. 146.

[379] Hodges, Steve, Berry, Emma, and Wood, Ken. 'SenseCam: A wearable camera that stimulates and rehabilitates autobiographical memory.' In: *Memory* 19.7 (2011). PMID: 21995708, pp. 685–696. DOI: 10.1080/09658211.2011.605591. eprint: https://doi.org/10.1080/09658211.2011.605591.

[380] Schmidt, Richard A and Wrisberg, Craig A. Motor learning and performance: A situation-based learning approach. Human kinetics, 2008.

[381] Dharani, Krishnagopal. 'Chapter 3 - Memory.' In: *The biology of thought: A neuronal mechanism in the generation of thought-A new molecular model*. Ed. by Krishnagopal Dharani. Academic Press, 2015, pp. 53–74. DOI: 10.1016/B978-0-12-800900-0.00003-8.

[382] Kim, Sungshin, Ogawa, Kenji, Lv, Jinchi, Schweighofer, Nicolas, and Imamizu, Hiroshi. 'Neural substrates related to motor memory with multiple timescales in sensorimotor adaptation.' In: *PLoS biology* 13.12 (2015), e1002312.

[383] Doyon, Julien, Bellec, Pierre, Amsel, Rhonda, Penhune, Virginia, Monchi, Oury, Carrier, Julie, Lehéricy, Stéphane, and Benali, Habib. 'Contributions of the basal ganglia and functionally related brain structures to motor learning.' In: *Behavioural Brain Research*. Special issue on the role of the basal ganglia in learning and memory 199.1 (2009), pp. 61–75. DOI: 10.1016/j.bbr.2008.11.012.

[384] Luft, Andreas R and Buitrago, Manuel M. 'Stages of motor skill learning.' In: *Molecular neurobiology* 32.3 (2005), pp. 205–216. DOI: 10.1385/MN:32:3:205.

[385] Masaki, Hiroaki and Sommer, Werner. 'Cognitive neuroscience of motor learning and motor control.' In: *The Journal of Physical Fitness and Sports Medicine* 1.3 (2012), pp. 369–380. DOI: 10.7600/jpfsm.1.369.

[386] Fitts, Paul M and Posner, Michael I. 'Human performance.' In: (1967).

[387] Weaver, Janelle. 'Motor learning unfolds over different timescales in distinct neural systems.' In: *PLoS biology* 13.12 (2015), e1002313.

[388] Sigrist, Roland, Rauter, Georg, Riener, Robert, and Wolf, Peter. 'Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review.' In: *Psychonomic bulletin & review* 20.1 (2013), pp. 21–53. DOI: 10.3758/s13423-012-0333-8.

[389] Wulf, Gabriele and Shea, Charles H. 'Principles derived from the study of simple skills do not generalize to complex skill learning.' In: *Psychonomic bulletin & review* 9.2 (2002), pp. 185–211. DOI: 10.3758/BF03196276.

[390] Maquet, Pierre. 'The role of sleep in learning and memory.' In: *science* 294.5544 (2001), pp. 1048–1052. DOI: 10.1126/science.1062856.

[391] Walker, Matthew P, Brakefield, Tiffany, Morgan, Alexandra, Hobson, J Allan, and Stickgold, Robert. 'Practice with sleep makes perfect: sleep-dependent motor skill learning.' In: *Neuron* 35.1 (2002), pp. 205–211. DOI: https://doi.org/10.1016/S0896-6273(02)00746-8.

[392] Machaen, Zandor, Martin, Luis, and Rosales, Jonathan-Hernando. 'Bio-inspired cognitive model of motor learning by imitation.' In: *Cognitive Systems Research* 66 (2021), pp. 134–149.

[393] Faltaous, Sarah, Koelle, Marion, and Schneegass, Stefan. 'From Perception to Action: A Review and Taxonomy on Electrical Muscle Stimulation in HCI.' In: *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*. ACM, 2022, pp. 159–171. DOI: 10.1145/3568444.3568460.

[394] Proteau, Luc, Marteniuk, Ronald G., and Lévesque, Line. 'A Sensorimotor Basis for Motor Learning: Evidence Indicating Specificity of Practice.' In: *The Quarterly Journal of Experimental Psychology Section A* 44.3 (1992). Publisher: Routledge _eprint: https://doi.org/10.1080/14640749208401298, pp. 557–575. DOI: 10.1080/14640749208401298.

[395] Tatsuno, Sho, Hayakawa, Tomohiko, and Ishikawa, Masatoshi. 'Supportive training system for sports skill acquisition based on electrical stimulation.' In: *2017 IEEE World Haptics Conference (WHC)*. 2017, pp. 466–471. DOI: 10.1109/WHC.2017.7989946.

[396] Nishikawa, Yoshiki, Atarashi, Yui, Sekiguchi, Shogo, and Shizuki, Buntarou. 'Effects of Electrical Muscle Stimulation on Memorability of Hand Gestures: A Preliminary Study.' In: *Universal Access in Human-Computer Interaction*. Ed. by Margherita Antona and Constantine Stephanidis. Springer Nature Switzerland, 2024, pp. 88–105. DOI: 10.1007/978-3-031-60875-9_7.

[397] Starch, Daniel. 'A demonstration of the trial and error method of learning.' In: *Psychological Bulletin* 7.1 (1910). Place: US Publisher: The Review Publishing Company, pp. 20–23. DOI: 10.1037/h0063796.

[398] Kiefer, Adam W, Gualberto Cremades, J, and Myer, Gregory D. 'Train the Brain: Novel Electroencephalography Data Indicate Links between Motor Learning and Brain Adaptations.' In: *Journal of novel physiotherapies* 4.2 (2014), p. 198. DOI: 10.4172/2165-7025.1000198.

[399] Rouleau, Isabelle, Salmon, David P., and Vrbancic, Mirna. ‘Learning, Retention and Generalization of a Mirror Tracing Skill in Alzheimer’s Disease.’ In: *Journal of Clinical and Experimental Neuropsychology* 24.2 (2002). Publisher: Routledge _eprint: https://doi.org/10.1076/jcen.24.2.239.997, pp. 239–250. DOI: 10.1076/jcen.24.2.239.997.

[400] Alrubaye, Zainab, Hudhud Mughrabi, Moaaz, Manav, Banu, and Batmaz, Anil Ufuk. ‘Effects of color cues on eye-hand coordination training with a mirror drawing task in virtual environment.’ In: *Frontiers in Psychology* 14 (2024). Publisher: Frontiers. DOI: 10.3389/fpsyg.2023.1307590.

[401] Arya, Kamal Narayan, Pandian, Shanta, Kumar, Dharmendra, and Puri, Vinod. ‘Task-Based Mirror Therapy Augmenting Motor Recovery in Poststroke Hemiparesis: A Randomized Controlled Trial.’ In: *Journal of Stroke and Cerebrovascular Diseases* 24.8 (2015), pp. 1738–1748. DOI: 10.1016/j.jstrokecerebrovasdis.2015.03.026.

[402] Bhushan, Braj, Dwivedi, C. B., Mishra, R., and Mandal, Manas K. ‘Performance on a Mirror-Drawing Task by Non-Right-Handers.’ In: *The Journal of General Psychology* 127.3 (2000). Publisher: Routledge _eprint: https://doi.org/10.1080/00221300009598585, pp. 271–277. DOI: 10.1080/00221300009598585.

[403] Desmottes, Lise, Maillart, Christelle, and Meulemans, Thierry. ‘Mirror-drawing skill in children with specific language impairment: Improving generalization by incorporating variability into the practice session.’ In: *Child Neuropsychology* 23.4 (2017). Publisher: Routledge _eprint: https://doi.org/10.1080/09297049.2016.1170797, pp. 463–482. DOI: 10.1080/09297049.2016.1170797.

[404] Faltaous, Sarah, Abdulmaksoud, Aya, Kempe, Markus, Alt, Florian, and Schneegass, Stefan. ‘GeniePutt: Augmenting human motor skills through electrical muscle stimulation.’ In: *it - Information Technology* 63.3 (2021). Publisher: De Gruyter Oldenbourg, pp. 157–166. DOI: 10.1515/itit-2020-0035.

[405] Hassan, Mahmoud, Daiber, Florian, Wiehr, Frederik, Kosmalla, Felix, and Krüger, Antonio. ‘FootStriker: An EMS-based Foot Strike Assistant for Running.’ In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1.1 (2017), 2:1–2:18. DOI: 10.1145/3053332.

[406] Kasahara, Shunichi, Takada, Kazuma, Nishida, Jun, Shibata, Kazuhisa, Shimojo, Shinsuke, and Lopes, Pedro. ‘Preserving Agency During Electrical Muscle Stimulation Training Speeds up Reaction Time Directly After Removing EMS.’ In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 2021, pp. 1–9. DOI: 10.1145/3411764.3445147.

[407] Tatsuno, Sho, Hayakawa, Tomohiko, and Ishikawa, Masatoshi. 'Trajectory adjustment system for learning based on electrical stimulation.' In: *Proceedings of the 8th Augmented Human International Conference*. AH '17. Association for Computing Machinery, 2017, pp. 1–4. DOI: 10.1145/3041164.3041197.

[408] Sullivan, Jennifer L., Pandey, Shivam, Byrne, Michael D., and O'Malley, Marcia K. 'Haptic Feedback Based on Movement Smoothness Improves Performance in a Perceptual-Motor Task.' In: *IEEE Transactions on Haptics* 15.2 (2022), pp. 382–391. DOI: 10.1109/TOH.2021.3126017.

[409] Miall, R. C. and Cole, J. 'Evidence for stronger visuo-motor than visuo-proprioceptive conflict during mirror drawing performed by a deafferented subject and control subjects.' In: *Experimental Brain Research* 176.3 (2007), pp. 432–439. DOI: 10.1007/s00221-006-0626-0.

[410] Snoddy, G. S. 'Learning and stability: a psychophysiological analysis of a case of motor learning with clinical applications.' In: *Journal of Applied Psychology* 10.1 (1926). Place: US Publisher: American Psychological Association, pp. 1–36. DOI: 10.1037/h0075814.

[411] Stratton, Shannon M., Liu, Yeou-Teh, Hong, Siang Lee, Mayer-Kress, Gottfried, and Newell, Karl M. 'Snoddy (1926) Revisited: Time Scales of Motor Learning.' In: *Journal of Motor Behavior* 39.6 (2007). Publisher: Routledge _eprint: https://doi.org/10.3200/JMBR.39.6.503-516, pp. 503–515. DOI: 10.3200/JMBR.39.6.503-516.

[412] Julius, Mona Sharon and Adi-Japha, Esther. 'A Developmental Perspective in Learning the Mirror-Drawing Task.' In: *Frontiers in Human Neuroscience* 10 (2016), p. 83. DOI: 10.3389/fnhum.2016.00083.

[413] Li, Zong-Ming, Kuxhaus, Laurel, Fisk, Jesse A., and Christophel, Thomas H. 'Coupling between wrist flexion–extension and radial–ulnar deviation.' In: *Clinical Biomechanics* 20.2 (2005), pp. 177–183. DOI: 10.1016/j.clinbiomech.2004.10.002.

[414] Youm, Y., McMurthy, R. Y., Flatt, A. E., and Gillespie, T. E. 'Kinematics of the wrist. I. An experimental study of radial-ulnar deviation and flexion-extension.' In: *JBJS* 60.4 (1978), p. 423. DOI: 10.2106/00004623-197860040-00001.

[415] Lopes, Pedro, Yüksel, Doăa, Guimbretière, François, and Baudisch, Patrick. 'Muscle-plotter: An Interactive System based on Electrical Muscle Stimulation that Produces Spatial Output.' In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 2016, pp. 207–217. DOI: 10.1145/2984511.2984530.

[416] McGorry, Raymond W., Chang, Chien-Chi, and Dempsey, Patrick G. 'A technique for estimation of wrist angular displacement in radial/ulnar deviation and flex-

ion/extension.' In: *International Journal of Industrial Ergonomics* 34.1 (2004), pp. 21–29. DOI: 10.1016/j.ergon.2004.01.005.

[417] Laforce, Robert and Doyon, Julien. 'Differential role for the striatum and cerebellum in response to novel movements using a motor learning paradigm.' In: *Neuropsychologia* 40.5 (2002), pp. 512–517. DOI: 10.1016/S0028-3932(01)00128-2.

[418] Serrien, Deborah J. 'Bimanual information processing and the impact of conflict during mirror drawing.' In: *Behavioural Brain Research* 205.2 (2009), pp. 391–395. DOI: 10.1016/j.bbr.2009.07.015.

[419] Marks, Ray. 'Ipsilateral and Contralateral Skill Acquisition following Random Practice of Unilateral Mirror-Drawing.' In: *Perceptual and Motor Skills* 83.3 (1996). Publisher: SAGE Publications Inc, pp. 715–722. DOI: 10.2466/pms.1996.83.3.715.

[420] Zimny, George H. 'Effects of Direction upon Speed of Mirror-Drawing.' In: *The American Journal of Psychology* 76.1 (1963). Publisher: University of Illinois Press, pp. 58–65. DOI: 10.2307/1419999.

[421] Wadden, Katie P., De Asis, Kristopher, Mang, Cameron S., Neva, Jason L., Peters, Sue, Lakhani, Bimal, and Boyd, Lara A. 'Predicting Motor Sequence Learning in Individuals with Chronic Stroke.' In: *Neurorehabilitation and neural repair* 31.1 (2017), pp. 95–104. DOI: 10.1177/1545968316662526.

[422] Ruttle, Jennifer E., Hart, Bernard Marius 't, and Henriques, Denise Y. P. 'Implicit motor learning within three trials.' In: *Scientific Reports* 11.1 (2021). Publisher: Nature Publishing Group, p. 1627. DOI: 10.1038/s41598-021-81031-y.

[423] Heathcote, Andrew, Brown, Scott, and Mewhort, D. J. K. 'The power law repealed: The case for an exponential law of practice.' In: *Psychonomic Bulletin & Review* 7.2 (2000), pp. 185–207. DOI: 10.3758/BF03212979.

[424] Newell, A. and Rosenbloom, Paul. 'Mechanisms of skill acquisition and the law of practice.' In: *Cognitive Skills and Their Acquisition* Vol. 1 (1993).

[425] Hagert, Elisabet, Persson, Jonas K. E., Werner, Michael, and Ljung, Björn-Ove. 'Evidence of Wrist Proprioceptive Reflexes Elicited After Stimulation of the Scapholunate Interosseous Ligament.' In: *The Journal of Hand Surgery* 34.4 (2009), pp. 642–651. DOI: 10.1016/j.jhsa.2008.12.001.

[426] Großhauser, Tobias, Bläsing, Bettina, Spieth, Corinna, and Hermann, Thomas. 'Wearable sensor-based real-time sonification of motion and foot pressure in dance teaching and training.' In: *Journal of the Audio Engineering Society* 60.7/8 (2012), pp. 580–589.

[427] Wu, Di, Ouyang, Jinhui, Dai, Ningyi, Wu, Mingzhu, Tan, Haodan, Deng, Hanhui, Fan, Yongmei, Wang, Dakuo, and Jin, Zhanpeng. 'DeepBrain: Enabling Fine-Grained Brain-Robot Interaction through Human-Centered Learning of Coarse EEG Signals from Low-Cost Devices.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6.3 (2022). DOI: 10.1145/3550334.

[428] Ren, Jie, Gao, Ling, Wang, Xiaoming, Ma, Miao, Qiu, Guoyong, Wang, Hai, Zheng, Jie, and Wang, Zheng. 'Adaptive Computation Offloading for Mobile Augmented Reality.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5.4 (2022). DOI: 10.1145/3494958.

[429] Nabeshima, Junichi, Saraiji, MHD Yamen, and Minamizawa, Kouta. 'Prosthetic Tail: Artificial Anthropomorphic Tail for Extending Innate Body Functions.' In: *Proceedings of the 10th Augmented Human International Conference 2019*. Ah2019. Association for Computing Machinery, 2019. DOI: 10.1145/3311823.3311848.

[430] Maekawa, Azumi, Takahashi, Shota, Saraiji, MHD Yamen, Wakisaka, Sohei, Iwata, Hiroyasu, and Inami, Masahiko. 'Naviarm: Augmenting the Learning of Motor Skills Using a Backpack-Type Robotic Arm System.' In: *Proceedings of the 10th Augmented Human International Conference 2019*. Ah2019. Association for Computing Machinery, 2019. DOI: 10.1145/3311823.3311849.

[431] Yumbla, Emiliano Quinones, Qiao, Zhi, Tao, Weijia, and Zhang, Wenlong. 'Human Assistance and Augmentation with Wearable Soft Robotics: a Literature Review and Perspectives.' In: *Current Robotics Reports* 2.4 (2021), pp. 399–413. DOI: 10.1007/s43154-021-00067-0.

[432] Sadeghian, Shadan and Hassenzahl, Marc. 'From Limitations to "Superpowers": A Design Approach to Better Focus on the Possibilities of Virtual Reality to Augment Human Capabilities.' In: *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. DIS '21. Association for Computing Machinery, 2021, 180–189. DOI: 10.1145/3461778.3462111.

[433] Neuhaus, Robin, Ringfort-Felner, Ronda, Sadeghian, Shadan, and Hassenzahl, Marc. 'To mimic reality or to go beyond? "Superpowers" in virtual reality, the experience of augmentation and its consequences.' In: *International Journal of Human-Computer Studies* 181 (2024), p. 103165. DOI: https://doi.org/10.1016/j.ijhcs.2023.103165.

[434] Mayring, Philipp. 'Qualitative Inhaltsanalyse – Abgrenzungen, Spielarten, Weiterentwicklungen.' In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 20.3 (2019). Number: 3. DOI: 10.17169/fqs-20.3.3343.

[435] Krauß, Veronika, Boden, Alexander, Oppermann, Leif, and Reiners, René. 'Current Practices, Challenges, and Design Implications for Collaborative AR/VR Ap-

plication Development.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445335.

[436] Neuhaus, Robin, Ringfort-Felner, Ronda, Courtney, Daniel, Kneile, Madlen, and Hassenzahl, Marc. 'Virtual Unreality: Augmentation-Oriented Ideation Through Design Cards.' In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. ACM, 2024.

[437] Ziat, Mounia, Balcer, Carrie Anne, Shirtz, Andrew, and Rolison, Taylor. 'A Century Later, the Hue-Heat Hypothesis: Does Color Truly Affect Temperature Perception?' In: *Haptics: Perception, Devices, Control, and Applications*. Ed. by Fernando Bello, Hiroyuki Kajimoto, and Yon Visell. Lecture Notes in Computer Science. Springer International Publishing, 2016, pp. 273–280. DOI: 10.1007/978-3-319-42321-0_25.

[438] Kocur, Martin, Jackermeier, Lukas, Schwind, Valentin, and Henze, Niels. 'The Effects of Avatar and Environment on Thermal Perception and Skin Temperature in Virtual Reality.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 2023, pp. 1–15. DOI: 10.1145/3544548.3580668.

[439] Slobodenyuk, Nadiya, Jraissati, Yasmina, Kanso, Ali, Ghanem, Lama, and Elhajj, Imad. 'Cross-Modal Associations between Color and Haptics.' In: *Attention, Perception, & Psychophysics* 77.4 (2015), pp. 1379–1395. DOI: 10.3758/s13414-015-0837-1.

[440] Wang, Han, Liu, Xin, Jing, Peiguang, and Liu, Yu. 'Airborne Ultrasound Haptics From Amplitude and Spatiotemporal Modulation of Acoustic Vortices.' In: *IEEE Sensors Letters* 7.8 (2023). Conference Name: IEEE Sensors Letters, pp. 1–4. DOI: 10.1109/LSENS.2023.3296345.

[441] Long, Benjamin, Seah, Sue Ann, Carter, Tom, and Subramanian, Sriram. 'Rendering Volumetric Haptic Shapes in Mid-Air Using Ultrasound.' In: *ACM Trans. Graph.* 33.6 (2014). DOI: 10.1145/2661229.2661257.

[442] Almesri, I., Awbi, H. B., Foda, E., and Sirén, K. 'An Air Distribution Index for Assessing the Thermal Comfort and Air Quality in Uniform and Nonuniform Thermal Environments.' In: *Indoor and Built Environment* 22.4 (2013), pp. 618–639. DOI: 10.1177/1420326X12451186. eprint: https://doi.org/10.1177/1420326X12451186.

[443] Witmer, Bob G. and Singer, Michael J. 'Measuring Presence in Virtual Environments: A Presence Questionnaire.' In: *Presence: Teleoperators and Virtual Environments* 7.3 (1998), pp. 225–240. DOI: 10.1162/105474698565686.

[444] Gehrke, Lukas, Akman, Sezen, Lopes, Pedro, Chen, Albert, Singh, Avinash Kumar, Chen, Hsiang-Ting, Lin, Chin-Teng, and Gramann, Klaus. 'Detecting Visuo-Haptic Mismatches in Virtual Reality using the Prediction Error Negativity of Event-Related

Brain Potentials.' In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 2019, pp. 1–11. DOI: 10.1145/3290605.3300657.

[445] Winter, Bodo. Linear models and linear mixed effects models in R with linguistic applications. Publisher: arXiv Version Number: 1. 2013. DOI: 10.48550/ARXIV.1308.5499.

[446] Zuur, Alain F, Ieno, Elena N, Walker, Neil J, Saveliev, Anatoly A, and Smith, Graham M. Mixed effects models and extensions in ecology with R. Vol. 574. Springer, 2009.

[447] Kuznetsova, Alexandra, Brockhoff, Per B., and Christensen, Rune H. B. 'lmerTest Package: Tests in Linear Mixed Effects Models.' In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26. DOI: 10.18637/jss.v082.i13.

[448] Günther, Sebastian, Müller, Florian, Schön, Dominik, Elmoghazy, Omar, Mühlhäuser, Max, and Schmitz, Martin. 'Therminator: Understanding the Interdependency of Visual and On-Body Thermal Feedback in Virtual Reality.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Association for Computing Machinery, 2020, 1–14. DOI: 10.1145/3313831.3376195.

[449] Blaga, Andreea Dalia, Frutos-Pascual, Maite, Creed, Chris, and Williams, Ian. 'Too Hot to Handle: An Evaluation of the Effect of Thermal Visual Representation on User Grasping Interaction in Virtual Reality.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Association for Computing Machinery, 2020, 1–16. DOI: 10.1145/3313831.3376554.

[450] Fang, Cathy, Zhang, Yang, Dworman, Matthew, and Harrison, Chris. 'Wireality: Enabling Complex Tangible Geometries in Virtual Reality with Worn Multi-String Haptics.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Association for Computing Machinery, 2020, 1–10. DOI: 10.1145/3313831.3376470.

[451] UltraLeap. Ultrahaptics Developer Information: Interaction zones and form factors. Accessed: 2022-08-16. 2022.

[452] Howard, Thomas, Marchal, Maud, Lécuyer, Anatole, and Pacchierotti, Claudio. 'PUMAH: Pan-tilt Ultrasound Mid-Air Haptics for larger interaction workspace in virtual reality.' In: *IEEE transactions on haptics* 13.1 (2019), pp. 38–44. DOI: 10.1109/TOH.2019.2963028.

[453] Brice, Daniel, McRoberts, Thomas, and Rafferty, Karen. 'A proof of concept integrated multi-systems approach for large scale tactile feedback in VR.' In: *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer International Publishing, 2019, pp. 120–137. DOI: 10.1007/978-3-030-25965-5_10.

[454] Suzuki, Shun, Fujiwara, Masahiro, Makino, Yasutoshi, and Shinoda, Hiroyuki. 'Midair Ultrasound Haptic Display with Large Workspace.' In: *Haptic Interaction*. Springer Singapore, 2019, pp. 3–5. DOI: 10.1007/978-981-13-3194-7_1.

[455] Johansson, Roland S and Vallbo, Ake B. 'Tactile sensibility in the human hand: relative and absolute densities of four types of mechanoreceptive units in glabrous skin.' In: *The Journal of physiology* 286.1 (1979), pp. 283–300. DOI: 10.1113/jphysiol.1979.sp012619.

[456] Corniani, Giulia and Saal, Hannes P. 'Tactile innervation densities across the whole body.' In: *Journal of Neurophysiology* 124.4 (2020), pp. 1229–1240. DOI: 10.1152/jn.00313.2020.

[457] Peters, Ryan M, McKeown, Monica D, Carpenter, Mark G, and Inglis, J Timothy. 'Losing touch: age-related changes in plantar skin sensitivity, lower limb cutaneous reflex strength, and postural stability in older adults.' In: *Journal of Neurophysiology* 116.4 (2016), pp. 1848–1858. DOI: 10.1152/jn.00339.2016.

[458] Schütt, Heiko H, Harmeling, Stefan, Macke, Jakob H, and Wichmann, Felix A. 'Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data.' In: *Vision research* 122 (2016), pp. 105–123. DOI: 10.1016/j.visres.2016.02.002.

[459] Tatarchuk, Natalya. 'Artist-Directable Real-Time Rain Rendering in City Environments.' In: *ACM SIGGRAPH 2006 Courses*. SIGGRAPH '06. Association for Computing Machinery, 2006, 23–64. DOI: 10.1145/1185657.1185828.

[460] Zhang, Yang, Kienzle, Wolf, Ma, Yanjun, Ng, Shiu S., Benko, Hrvoje, and Harrison, Chris. 'ActiTouch: Robust Touch Detection for On-Skin AR/VR Interfaces.' In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. Association for Computing Machinery, 2019, 1151–1159. DOI: 10.1145/3332165.3347869.

[461] Schrepp, Martin, Hinderks, Andreas, and Thomaschewski, Jörg. 'Construction of a Benchmark for the User Experience Questionnaire (UEQ).' In: *International Journal of Interactive Multimedia and Artificial Intelligence* 4.4 (2017), pp. 40–44. DOI: 10.9781/ijimai.2017.445.

[462] Knierim, Pascal, Schwind, Valentin, Feit, Anna Maria, Nieuwenhuizen, Florian, and Henze, Niels. 'Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Association for Computing Machinery, 2018, 1–9. DOI: 10.1145/3173574.3173919.

[463] Regenbrecht, Holger T., Schubert, Thomas W., and Friedmann, Frank. 'Measuring the Sense of Presence and its Relations to Fear of Heights in Virtual Environments.'

In: *International Journal of Human-Computer Interaction* 10.3 (1998), pp. 233–249. DOI: 10.1207/s15327590ijhc1003_2.

[464] Freeman, Euan, Anderson, Ross, Williamson, Julie, Wilson, Graham, and Brewster, Stephen A. 'Textured Surfaces for Ultrasound Haptic Displays.' In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ICMI '17. Association for Computing Machinery, 2017, 491–492. DOI: 10.1145/3136755.3143020.

[465] Wobbrock, Jacob O., Findlater, Leah, Gergle, Darren, and Higgins, James J. 'The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Association for Computing Machinery, 2011, 143–146. DOI: 10.1145/1978942.1978963.

[466] Gonzalez, Eric J., Abtahi, Parastoo, and Follmer, Sean. 'REACH+: Extending the Reachability of Encountered-type Haptics Devices through Dynamic Redirection in VR.' In: *Proc. of the 33rd Annual ACM Symposium on User Interface Software and Technology*. UIST '20. ACM, 2020, pp. 236–248. DOI: 10.1145/3379337.3415870.

[467] Mercado, Victor, Marchal, Maud, and Lecuyer, Anatole. 'ENTROPiA: Towards Infinite Surface Haptic Displays in Virtual Reality Using Encountered-Type Rotating Props.' In: *IEEE Transactions on Visualization and Computer Graphics* 27.3 (2021), pp. 2237–2243. DOI: 10.1109/TVCG.2019.2963190.

[468] Mercado, Victor Rodrigo, Howard, Thomas, Si-Mohammed, Hakim, Argelaguet, Ferran, and Lécuyer, Anatole. 'Alfred: the Haptic Butler On-Demand Tangibles for Object Manipulation in Virtual Reality using an ETHD.' In: *2021 IEEE World Haptics Conf. (WHC)*. 2021, pp. 373–378. DOI: 10.1109/WHC49131.2021.9517250.

[469] Araujo, Bruno, Jota, Ricardo, Perumal, Varun, Yao, Jia Xian, Singh, Karan, and Wigdor, Daniel. 'Snake Charmer: Physically Enabling Virtual Objects.' In: *Proc. of the TEI '16: Tenth Intl. Conf. on Tangible, Embedded, and Embodied Interaction*. ACM, 2016, pp. 218–226. DOI: 10.1145/2839462.2839484.

[470] Mercado, Víctor Rodrigo, Marchal, Maud, and Lécuyer, Anatole. '"Haptics On-Demand": A Survey on Encountered-Type Haptic Displays.' In: *IEEE Transactions on Haptics* 14.3 (2021). Conf. Name: IEEE Transactions on Haptics, pp. 449–464. DOI: 10.1109/TOH.2021.3061150.

[471] Schubert, Thomas, Friedmann, Frank, and Regenbrecht, Holger. 'The Experience of Presence: Factor Analytic Insights.' In: *Presence: Teleoperators and Virtual Environments* 10.3 (2001), pp. 266–281. DOI: 10.1162/105474601300343603.

[472] Anwar, Ahmed, Shi, Tianzheng, and Schneider, Oliver. 'Factors of Haptic Experience across Multiple Haptic Modalities.' In: *Proc. of the 2023 CHI Conf.* ACM, 2023, pp. 1–12. DOI: 10.1145/3544548.3581514.

[473] Sathiyamurthy, Suji, Lui, Melody, Kim, Erin, and Schneider, Oliver. 'Measuring Haptic Experience: Elaborating the HX model with scale development.' In: *2021 IEEE World Haptics Conf. (WHC)*. IEEE, 2021, pp. 979–984. DOI: 10.1109/WHC49131.2021.9517220.

[474] Shen, Vivian, Shultz, Craig, and Harrison, Chris. 'Mouth Haptics in VR using a Headset Ultrasound Phased Array.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Association for Computing Machinery, 2022, pp. 1–14. DOI: 10.1145/3491102.3501960.

[475] Dorf, Richard C and Bishop, Robert H. Modern control systems. Pearson, 2011.

[476] Ramon Borràs, Marina. 'iMOVE: Development of a hybrid control interface based on sEMG and movement signals for an assistive robotic manipulator.' MA thesis. Universitat Politècnica de Catalunya, 2020.

[477] Zhao, Dongya, Zhao, Yongrui, Li, Shaoyuan, and Zhu, Quanmin. 'A new terminal converging PD control for parallel robotic manipulators with bounded torque.' In: *Proceedings of 2011 International Conference on Modelling, Identification and Control*. IEEE, 2011, pp. 115–121. DOI: 10.1109/ICMIC.2011.5973686.

[478] Ouyang, Puren R, Tang, J, Yue, WH, and Jayasinghe, Shanuka. 'Adaptive PD plus sliding mode control for robotic manipulator.' In: *2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2016, pp. 930–934. DOI: 10.1109/AIM.2016.7576888.

[479] Yang, Chifu, Huang, Qitao, Jiang, Hongzhou, Peter, O Ogbobe, and Han, Junwei. 'PD control with gravity compensation for hydraulic 6-DOF parallel manipulator.' In: *Mechanism and Machine theory* 45.4 (2010), pp. 666–677. DOI: 10.1016/j.mechmachtheory.2009.12.001.

[480] Phuoc, Le Minh, Martinet, Philippe, Lee, Sukhan, and Kim, Hunmo. 'Damped least square based genetic algorithm with Ggaussian distribution of damping factor for singularity-robust inverse kinematics.' In: *Journal of mechanical science and technology* 22.7 (2008), pp. 1330–1338. DOI: 10.1007/s12206-008-0427-4.

[481] Ariga, Kentaro, Fujiwara, Masahiro, Makino, Yasutoshi, and Shinoda, Hiroyuki. 'Workspace Evaluation of Long-Distance Midair Haptic Display Using Curved Reflector.' In: *2021 IEEE World Haptics Conference (WHC)*. IEEE, 2021, pp. 85–90. DOI: 10.1109/WHC49131.2021.9517193.

[482] Suzuki, Shun, Fujiwara, Masahiro, Makino, Yasutoshi, and Shinoda, Hiroyuki. 'Midair Ultrasound Haptic Display with Large Workspace.' In: *Haptic Interaction*. Ed. by Hiroyuki Kajimoto, Dongjun Lee, Sang-Youn Kim, Masashi Konyo, and Ki-Uk Kyung. Springer Singapore, 2019, pp. 3–5.

[483] Brice, Daniel, McRoberts, Thomas, and Rafferty, Karen. 'A Proof of Concept Integrated Multi-systems Approach for Large Scale Tactile Feedback in VR.' In: *Augmented Reality, Virtual Reality, and Computer Graphics*. Ed. by Lucio Tommaso De Paolis and Patrick Bourdot. Springer International Publishing, 2019, pp. 120–137. DOI: 10.1007/978-3-030-25965-5_10.

[484] Hoppe, Matthias, Rossmy, Beat, Neumann, Daniel Peter, Streuber, Stephan, Schmidt, Albrecht, and Machulla, Tonja-Katrin. 'A Human Touch: Social Touch Increases the Perceived Human-Likeness of Agents in Virtual Reality.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Association for Computing Machinery, 2020, 1–11. DOI: 10.1145/3313831.3376719.

[485] Salazar, Steeven Villa, Pacchierotti, Claudio, Tinguy, Xavier de, Maciel, Anderson, and Marchal, Maud. 'Altering the Stiffness, Friction, and Shape Perception of Tangible Objects in Virtual Reality Using Wearable Haptics.' In: *IEEE Transactions on Haptics* 13.1 (2020), pp. 167–174. DOI: 10.1109/TOH.2020.2967389.

[486] Tinguy, Xavier de, Pacchierotti, Claudio, Marchal, Maud, and Lécuyer, Anatole. 'Enhancing the Stiffness Perception of Tangible Objects in Mixed Reality Using Wearable Haptics.' In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 81–90. DOI: 10.1109/VR.2018.8446280.

[487] Tinguy, Xavier de, Pacchierotti, Claudio, Emily, Mathieu, Chevalier, Mathilde, Guignardat, Aurélie, Guillaudeux, Morgan, Six, Chloé, Lécuyer, Anatole, and Marchal, Maud. 'How Different Tangible and Virtual Objects Can Be While Still Feeling the Same?' In: *2019 IEEE World Haptics Conference (WHC)*. IEEE, 2019, pp. 580–585. DOI: 10.1109/WHC.2019.8816164.

[488] Elsayed, Hesham, Weigel, Martin, Müller, Florian, Schmitz, Martin, Marky, Karola, Günther, Sebastian, Riemann, Jan, and Mühlhäuser, Max. 'Vibromap: Understanding the spacing of vibrotactile actuators across the body.' In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4.4 (2020), pp. 1–16.

[489] Cholewiak, Roger W and Craig, James C. 'Vibrotactile pattern recognition and discrimination at several body sites.' In: *Perception & psychophysics* 35.6 (1984), pp. 503–514.

[490] Weinstein, Sidney. 'Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality.' In: *The skin senses* (1968).

[491] Choi, Seungmoon and Kuchenbecker, Katherine J. 'Vibrotactile Display: Perception, Technology, and Applications.' In: *Proceedings of the IEEE* 101.9 (2013), pp. 2093–2104. DOI: 10.1109/jproc.2012.2221071.

[492] Azadi, Mojtaba and Jones, Lynette A. 'Identification of vibrotactile patterns: building blocks for tactons.' In: *2013 World Haptics Conference (WHC)* (2013), pp. 347–352.

[493] Jones, Lynette A. and Singhal, Anshul. 'Perceptual dimensions of vibrotactile actuators.' In: *2018 IEEE Haptics Symposium (HAPTICS)* (2018), pp. 307–312.

[494] Bradley, Margaret M. and Lang, Peter J. 'Measuring emotion: The self-assessment manikin and the semantic differential.' In: *Journal of Behavior Therapy and Experimental Psychiatry* 25.1 (1994), pp. 49–59. DOI: https://doi.org/10.1016/0005-7916(94)90063-9.

[495] Zhang, Xuan, Yu, Hui W., and Barrett, Lisa F. 'How does this make you feel? A comparison of four affect induction procedures.' In: *Frontiers in Psychology* 5 (). DOI: 10.3389/fpsyg.2014.00689.

[496] Siedlecka, Ewa and Denson, Thomas F. 'Experimental Methods for Inducing Basic Emotions: A Qualitative Review.' In: *Emotion Review* 11.1 (2019), pp. 87–97. DOI: 10.1177/1754073917749016. eprint: https://doi.org/10.1177/1754073917749016.

[497] Gross, James J. and Levenson, Robert W. 'Emotion elicitation using films.' In: *Cognition and Emotion* 9.1 (1995), pp. 87–108. DOI: 10.1080/02699939508408966. eprint: https://doi.org/10.1080/02699939508408966.

[498] Schaefer, Alexandre, Nils, Frédéric, Sanchez, Xavier, and Philippot, Pierre. 'Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers.' In: *Cognition and emotion* 24.7 (2010), pp. 1153–1172.

[499] Lobbestael, Jill, Arntz, Arnoud, and Wiers, Reinout W. 'How to push someone's buttons: A comparison of four anger-induction methods.' In: *Cognition & Emotion* 22.2 (2008), pp. 353–373.

[500] Jurásová, Kinga and Spajdel, M. 'Development and assessment of film excerpts used for emotion elicitation.' In: *Activitas Nervosa Superior Rediviva* 55.3 (2013), pp. 135–140.

[501] Gilman, T Lee, Shaheen, Razan, Nylocks, K Maria, Halachoff, Danielle, Chapman, Jessica, Flynn, Jessica J, Matt, Lindsey M, and Coifman, Karin G. 'A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation.' In: *Behavior research methods* 49.6 (2017), pp. 2061–2082.

[502] Gevrey, Muriel, Dimopoulos, Ioannis, and Lek, Sovan. 'Review and comparison of methods to study the contribution of variables in artificial neural network models.' In: *Ecological Modelling* 160.3 (2003). Modelling the structure of acquatic communities: concepts, methods and problems., pp. 249–264. DOI: https://doi.org/10.1016/S0304-3800(02)00257-0.

[503] Van Berkel, Niels and Hornbæk, Kasper. 'Implications of Human-Computer Interaction Research.' In: *Interactions* 30.4 (2023), pp. 50–55. DOI: 10.1145/3600103.

[504] Norman, Donald A. The design of everyday things. Springer, 2013.

[505] Schmidt, Albrecht and Herrmann, Thomas. 'Intervention user interfaces: a new interaction paradigm for automated systems.' In: *interactions* 24.5 (2017), pp. 40–45. DOI: 10.1145/3121357.

[506] Mosqueira-Rey, Eduardo, Hernández-Pereira, Elena, Alonso-Ríos, David, Bobes-Bascarán, José, and Fernández-Leal, Ángel. 'Human-in-the-loop machine learning: a state of the art.' In: *Artificial Intelligence Review* 56.4 (2023), pp. 3005–3054. DOI: 10.1007/s10462-022-10246-w.

[507] ISO/IEC 22989:2022.

[508] Laaksoharju, Mikael. Designing for Autonomy. Illustrated edição. Uppsala Universitet, 2014.

[509] Tan, Felicia Fang-Yi, Gupta, Chitralekha, Rajendran, Dixon Prem Daniel, Maes, Pattie, and Nanayakkara, Suranga. 'Assistive Augmentation: Fundamentally Transforming Human Ability.' In: *interactions* 32.1 (2025), pp. 22–27. DOI: 10.1145/3702656.

[510] Rogers, Yvonne. HCI Theory: Classical, Modern, and Contemporary. Synthesis Lectures on Human-Centered Informatics. Springer International Publishing, 2012. DOI: 10.1007/978-3-031-02197-8.

[511] Popper, Karl Raimund. The Logic of Scientific Discovery. Google-Books-ID: Yq6xeupNStMC. Psychology Press, 2002.

[512] Dienes, Zoltan. Understanding Psychology as a Science: An Introduction to Scientific and Statistical Inference. Google-Books-ID: kVfiEAAAQBAJ. Bloomsbury Publishing, 2008.

[513] Kirsch, Irving. 'Specifying Nonspecifics: Psychological.' In: *The placebo effect: An interdisciplinary exploration* 8 (1999), p. 166.

[514] Yakura, Hiromu, Lopez-Lopez, Ezequiel, Brinkmann, Levin, Serna, Ignacio, Gupta, Prateek, and Rahwan, Iyad. Empirical evidence of Large Language Model's influence on human spoken communication. arXiv:2409.01754 [cs]. 2024. DOI: 10.48550/arXiv.2409.01754.

[515] Rafner, Janet, Dellermann, Dominik, Hjorth, Arthur, Verasztó, Dóra, Kampf, Constance, Mackay, Wendy, and Sherson, Jacob. 'Deskilling, Upskilling, and Reskilling: a Case for Hybrid Intelligence.' In: *Morals & Machines* 1.2 (2021), pp. 24–39. DOI: 10.5771/2747-5174-2021-2-24.

List of Figures

Ends and Means

List of Tables

Context

Autonomy

Heteronomy

Experiencing and Simulating

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig und ohne unerlaubte Beihilfe angefertigt wurde.

München, den 28. Februar 2025

David Steeven Villa Salazar