Physiologically Adaptive Systems Across the Mixed Reality Continuum

Francesco Chiossi

2025

Physiologically Adaptive Systems Across the Mixed Reality Continuum

Francesco Chiossi

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

vorgelegt von
Francesco Chiossi
aus Modena, Italien

München, den 15.07.2024

Erstgutachter: Prof. Dr. Albrecht Schmidt
Zweitgutachter: Prof. Dr. Michael Sedlmair
Drittgutachter: Prof. Dr. Tilman Dingler

Tag der mündlichen Prüfung: 11.11.2024

Abstract

As Mixed Reality (MR) technologies progressively enter our daily lives, blending digital information seamlessly with our physical world, they unlock new potential for enhancing productivity and engagement across various activities, from sketching to typing and object manipulation. However, this fusion of virtual and physical elements also elevates the complexity of our visual environments, challenging our capacity to navigate and process an increasingly congested information space without becoming overwhelmed. This thesis contends that the dynamic nature of MR systems, which blurs the lines between the real and virtual worlds, necessitates the development of adaptive MR environments. Such environments are designed to intelligently modulate and tailor content in real-time, responding to users' cognitive states and environmental contexts to optimize user engagement and minimize cognitive strain.

Centered on the premise that effective interaction within these richly blended settings hinges on users' ability to manage attentional resources efficiently, this research explores the integration of physiological computing into MR. By leveraging real-time monitoring of physiological signals, MR systems can gain implicit insights into the user's attentional state and cognitive load, enabling them to adapt dynamically to support the user's immediate needs and objectives. By examining how various degrees of virtuality affect users' engagement, attentional allocation, and workload, this thesis systematically investigates the potential of adaptive MR systems to enhance user experience without compromising productivity.

The investigation extends to designing, evaluating, and implementing physiological computing systems within MR environments, offering novel insights into supporting task engagement and managing attentional load. Moreover, it tackles the technical challenges of embedding physiological sensors within MR hardware, proposing a groundbreaking approach to unobtrusive user state monitoring.

This thesis makes significant strides in bridging the gap between physiological computing and MR, laying the groundwork for future research in ubiquitous computing, pervasive computing, and affective computing. Its findings highlight the critical role of interdisciplinary collaboration in understanding and realizing implicit interaction in adaptive MR systems, where intelligent environments enhance our daily lives by being responsive, intuitive, and seamlessly integrated.

Zusammenfassung

Da Mixed Reality (MR)-Technologien zunehmend in unser tägliches Leben Einzug halten und digitale Informationen nahtlos mit unserer physischen Welt verschmelzen, eröffnen sie neues Potenzial zur Steigerung der Produktivität und des Engagements bei verschiedenen Aktivitäten, von Skizzieren bis hin zu Tippen und Objektmanipulation. Diese Verschmelzung von virtuellen und physischen Elementen erhöht jedoch auch die Komplexität unserer visuellen Umgebungen und stellt unsere Fähigkeit auf die Probe, in einem zunehmend überfüllten Informationsraum zu navigieren und Informationen zu verarbeiten, ohne überwältigt zu werden. Diese Arbeit argumentiert, dass die dynamische Natur von MR-Systemen, die die Grenzen zwischen realer und virtueller Welt verwischen, die Entwicklung adaptiver MR-Umgebungen erforderlich macht. Solche Umgebungen sind darauf ausgelegt, Inhalte in Echtzeit intelligent zu modulieren und anzupassen, um auf die kognitiven Zustände und Umweltkontexte der Nutzer zu reagieren, um das Engagement zu optimieren und die kognitive Belastung zu minimieren.

Ausgehend von der Prämisse, dass eine effektive Interaktion in diesen reichhaltig verschmolzenen Umgebungen von der Fähigkeit der Nutzer abhängt, ihre Aufmerksamkeitsressourcen effizient zu verwalten, untersucht diese Forschung die Integration physiologischer Datenverarbeitung in MR. Durch die Nutzung der Echtzeitüberwachung physiologischer Signale können MR-Systeme implizite Einblicke in den Aufmerksamkeitszustand und die kognitive Belastung des Nutzers gewinnen und sich dynamisch anpassen, um die unmittelbaren Bedürfnisse und Ziele des Nutzers zu unterstützen. Indem untersucht wird, wie verschiedene Grade der Virtualität das Engagement der Nutzer, die Aufmerksamkeitsverteilung und die Arbeitsbelastung beeinflussen, untersucht diese Arbeit systematisch das Potenzial adaptiver MR-Systeme zur Verbesserung der Nutzererfahrung, ohne die Produktivität zu beeinträchtigen.

Die Untersuchung erstreckt sich auf das Design, die Evaluierung und die Implementierung physiologischer Datenverarbeitungssysteme innerhalb von MR-Umgebungen und bietet neuartige Einblicke in die Unterstützung der Aufgabenbindung und das Management der Aufmerksamkeitsbelastung. Darüber hinaus werden die technischen Herausforderungen der Einbettung physiologischer Sensoren in MR-Hardware angegangen und ein bahnbrechender Ansatz zur unaufdringlichen Überwachung des Nutzerzustands vorgeschlagen.

Diese Arbeit macht bedeutende Fortschritte bei der Überbrückung der Kluft zwischen physiologischer Datenverarbeitung und MR und legt den Grundstein für zukünftige Forschung in den Bereichen allgegenwärtiges Computing, Pervasive Computing und Affective Computing. Die Ergebnisse heben die entscheidende Rolle der interdisziplinären Zusammenarbeit beim Verständnis und der Verwirklichung impliziter Interaktionen in adaptiven MR-Systemen hervor, in denen intelligente Umgebungen unser tägliches Leben durch Reaktionsfähigkeit, Intuition und nahtlose Integration verbessern.

Acknowledgements

I often imagined the moment when I would sit down to write the acknowledgments of my thesis, and now that day has arrived. It has been a wonderful journey, filled with challenges, learning, and growth, and none of it would have been the same without a group of truly special people who walked this path with me.

First, I want to thank **Albrecht Schmidt** for teaching me how to push ideas forward with creativity and persistence while giving me the freedom to explore and the opportunity to travel the world. Above all, I treasure his inspiring community — a space of trust, curiosity, and connection that I was lucky to be part of. Then, this thesis would not be completed without **Michael Sedlmair** and **Tilman Dingler**. You dedicated time, thoughtful feedback, and support in evaluating my work. I am especially grateful to **Sven Mayer**, with whom I was lucky to share both friendship and guidance — your supervision turned my commitment into a successful and fulfilling path, teaching me how to grow as a researcher and how to support others in turn. After two challenging years, starting to work with you made me feel motivated, focused, and hungry for scientific work, and your energy and proactiveness made the lab a better place for everyone — turning this journey into one of learning, growth, and enthusiasm. I am deeply thankful to **Luke Haliburton** for always spreading positivity and for being a model of how to navigate the challenges of a PhD — a great co-author, and an open ear in difficult moments. This journey would not have been the same without starting as colleagues and finding a true friend along the way. Thank you **Steeven Villa Salazar**, , for sharing our path to disentangle the endless mess of German bureaucracy, for always being there — whether for a good bullshitting session, laughing over research quirks, or sharing beers and stories. You showed me how many values and friendships can grow between people from all over the world, and I'm grateful for all the fun along the way. **Carl Oechsner**, first of all, thanks for being a friend, before being. For always saying yes to adventures, the birthday trips, for moving Renaissance paintings in Venice Biennale, funk concerts surrounded by teenagers, music festivals, grilled oysters on a pick-up truck in New Orleans, always saying yes to adventures, and for all the open, honest chats where we skip the small talk and talk straight to the good and bad, important things. It's been awesome sharing all of that with you. Thank you **Julian Rasch** — for sharing the mustache office, for all the knowledge worker jokes, 3 AM bench conversations, and the silent complicity in late-night CHI paper writing sessions with Boiler Room sets in the background. From spontaneous jungle hikes in Hawaii, climbing lianas, and guiding us through the Hamburg club scene. Thank **Jan Leusmann**, the youngest member of the morning crew. Thanks for all the interior design talks, for helping process the collective grief over losing Patagonia merch, for the "opinion-free" carbonara, and for showing me how optimization applies to almost everything in life (even if, sorry, I might still reject it). Thank you **Changkun Ou** for too many insightful discussions and countless laughs and midnight bursts of research passion. You shaped some of the most enjoyable moments of my PhD, from late-night walks home talking about careers and values in life, preferences, your "I see where you are coming from," and showing me how to switch the locus of control when it matters most. Thank you, **Matthias Schmidtmaier**, for bringing back the light of reason and sanity and normality in an office full of HCI nerds, for sharing good aesthetic taste, for showing me how much fun Oktoberfest can be, for climbing walls in Venice and admiring the lagoon, and — of course — for enjoying together the worst American pizzas. Thank you, **Thomas Kosch**, for being my first office mate, and first kicking me out of my comfort zone with the loudest typing and the heaviest metal music. For being one of the highest-energy people I've ever met, for open chats about careers, great memes, endless stickers — and I still hope that one day you'll stop joking about Pizza Hawaii. Thank you, **Florian Bemmann** — for bringing so much positivity with your sunny smile and our friendly morning chats, for being a true cannolo and dolce vita enjoyer. Modena is waiting for you and for a very productive remote working session. Thank you, **Teodora Mitrevska** — for all the good laughs and those open, empathetic chats about starting EEG research (you will rock it!). Thanks for challenging me and helping me grow in supporting and co-supervising you. Thank you **Lewis Chuang**, for having pushed and enabled a neuroscientist (or so I believed I was) to find his place in an HCI group. Thank you **Sarah Völkel** — for always being kind and available for my naïve questions about breaking into industry, for having one of the brightest senses of humor I've ever seen, for organizing fun Christmas parties, and for including me in planning events I had no idea how to handle. The lab still misses you! **Maxi Windl** — for the unstoppable, restless energy you always bring to the lab.

Matthias Hoppe, for streaming sumo fights, Godzilla and Japanese anime chats, and epic Nerf gun battles in the lab. **Jesse Grootjen** – for showing your amazing baking skills, great organization in teaching Human Factors and hosting the best warm-up parties with a view over Viktualienmarkt. **Jakob Karolus**, for the loudest laughs in the lab. **Florian Lang** — for being (most of the time) patient with my endless bureaucracy requests, for comforting me and Steeven after our on-site interviews, for your wonderfully relaxing (at least to me!) complaints and German bureaucracy jokes, and for always wearing the coolest Dragon Ball T-shirts. **Fiona Draxler**— for being such a relaxing presence in the lab, with your fancy bajour lights creating the perfect chill writing vibes, and for your great taste in Italian indie music. Thank you **Clara Sayffaerth** — for keeping the Japanese lore alive in the office and always bringing good dank jokes. **Robin Welsch** — for showing how a psychologist can break into HCI and absolutely dominate it, for hosting an amazing wedding party, and for teaching me how to bring more solid, good science into my work. Thank you **Henrike Weingärtner** — for all your effort and dedication in making LMU a more inclusive and diverse research group. **Yannick Weiss** — for the great New Orleans and Australian adventures… but we're still waiting for that mullet cut, tough. Thank you **Heiko Drewes** — thank you for your always out-of-the-box comments and way of thinking, your stylish and crazy life stories, and, of course, Chinese cigarettes. **Sophia Sakel** — for organizing fancy IDCs and upcoming TikTok projects. Thank you **Kathrin Schnitzer** — for bringing so much positivity, big smiles, and literal kilos of fusilli to fill up my desk. Thank you **Pascal Knierim** for hill-areous Nerf fights. Thank you - **Thomas Weber** — for bringing to life the ideas that became my PhD hat and for being the lab's BBQ master, bartender, and best AI-memer all in one. Thank you **Rifat Amin** for awesome biryani and positive energy, and great gelato in Liguria. Thank you, **Andreas**

Butz, for always being kind and welcoming, stepping in as senior author for last-minute submissions. You've shown me how important it is to balance research with dedicating time and love to family. Thank you **Heinrich Hussmann** — for putting the first brick of the Media Informatics group and always being a warm, smiling presence in the halls. You are missed.

My work would not have been possible without a series of people who brought new expertise, introduced new methods, and helped me make my work more flexible and well-rounded. I am deeply grateful to my co-authors — **Felix Putze**, your deep knowledge of neuroergonomics and brain-computer interfaces pushed my thinking forward and expanded the scientific depth of our collaborations. **Uwe Gruenefeld**, thank you for your creativity and solid engineering mindset — working with you always felt fun, efficient and inspiring. Joshua Newn, your sense of humor even in the most stressful moments, your eye tracking modeling methods, our shared appreciation for aesthetic photography, and your delicious Melbourne food tips made working with you an absolute joy. **Ludwig Sidenmark**, I've always admired your calm, critical eye, and methodological sharpness — your feedback helped refine our work at every step. **Florian Müller**, you showed me how HCI can be incredibly fun, and how playful, irreverent ideas can turn into great papers — all while keeping in mind our social and societal responsibilities as scientists. **Linda Hirsch**, your perspective on social and inclusive aspects of HCI brought new dimensions to our discussions and made the research stronger. and **Abdallah El Ali**, your passion, drive, and generosity with ideas and empathy brought great energy into our work together — I learned a lot from your perspective. To all the amazing students I had the chance to supervise along the way — I learned so much from each of you. A special thanks to Ines, Simon, Martina, **Carolina**, **Eric**, **Yassmina**, **Yagiz**, **Rulu**, **Elnur**, and **Xingyu**.

Anja Mebus, thank you for patiently standing by my side in the eternal battle against the German Bureaucracy, rejecting all my Reisentrag refund forms and for keeping things running smoothly. **Rainer Fink**, you are the quiet force behind the scenes — thank you for making sure the lab doesn't fall apart and always being available, kind, and tolerant. **Christa**, **Salsabeel**, and **Franziska** — thank you for all the little-big things you've helped with along the way; your support and kindness made a bigger difference than you might know.

To my friends — thank you for being the constant thread that held everything together when life got intense. During a time of change, uncertainty, and growth, you reminded me of who I am beyond the research, helped me stay grounded, and brought joy when I needed it most. **Giacomo**, — for top-tier brain-rot content, for always being there for a chat. You made even the dullest days entertaining. To **Tancredi** and **Lorenzo** — my Münchener brothers, for making this city feel like home. To **Niklas** — your double coffees gave me the double amount of caffeine I needed to write smart papers (or at least try), and your timely career suggestions always arrived when I needed them most.

To my Modena friends, thank you for never letting distance get in the way and came to celebrate our friendship for my defense. **Guido**, from elementary school to now — you've been there through every phase, and that kind of bond means the world. **Nikola**, for legendary dance moves at post-Oktoberfest parties — few could match that energy. **Uto**, for the daily check-ins and the kind of morning chats that made even the earliest hours fun. **Debbo**, for your quiet strength and kind heart — I'm really grateful to have you as a friend. **Zeno**, you literally biked to Munich — that says it all. You've been an inspiration with your quiet heart choices, thoughtful listening, and the calm you bring wherever you go. **Alberto** for the friendship made of places, textures, and moments that feel like home: Modena, Venezia, Padova, Siracusa, and all the fragments of memory in between. For the aesthetic routines, the tour memories, the music, they filled up the heart.

To my Padua friends, **Stefano**, we met eight years ago, bonded over stats, sarcasm, and shared brain wrinkles. You got your PhD in neuroscience, I got mine in HCI — but in a way, we crossed the finish line together. Thanks for always being there for a laugh, a listen, and a perfectly timed statistical rant. **Fabio**, we met right before I started in Munich, we managed to publish together, and now seeing you thrive as world-class researchers has been incredible. Growing together in this field has been one of the most rewarding parts of the ride. **Patrik**, you were my first HCI mentor — without your support, guidance, and belief in me during the interview process, I probably wouldn't have found this PhD position. I owe you a lot.

Thank you **Johanna** — for standing by me through this milestone with joy, care, and love. Your support through the writing, your calm presence, and all the small, thoughtful actions, made this final stretch not only lighter, but full of love I'm so proud of what we have, and we want to build together, and sharing this day with you - filled with joy, music, laughter, and the people we love - made it even more special. I'm lucky to have you next to me.

To my mum, **Anna** — being the first Chiossi PhD is an honor, but everything I am comes from you. You taught me the importance of being happy, even in the hardest moments, showing me the unstoppable power of motivation and what it means to fight for your goals, no matter what. To my dad, **Roberto**, for your high energy and work ethics — thank you for always being there, believing in me even when I doubted myself.

To my nonni, **Enzo** and **Giordana**, for being there at the beginning and beyond — always present in me, in my choices, and in moments when making decisions was harder.

Thank you.

TABLE OF CONTENTS

1 INTRODUCTION

"Reality is created by the mind, we can change our reality by changing our mind."

– Plato.
***The Republic*. 375 BC.**

Mixed Reality (MR) systems are dynamic computing systems that blur the lines between the real and virtual worlds, creating environments where digital and physical elements coexist and interact in real time. As MR systems increasingly enter our everyday lives, the integration of digital information expands both the complexity and diversity of the stimuli we encounter. This shift unlocks a broad spectrum of applications, such as sketching [322], typing [380], and manipulating objects [545], where virtual and physical components merge seamlessly to enhance productivity and engagement. However, the influx of digital content also introduces challenges in navigating visually dense environments where relevant and irrelevant information coexist, often leading to cognitive overload.

The immersive nature of MR environments challenges users' attentional management, as they must continuously process multiple layers of virtual and physical information. Unlike traditional screen-based interfaces, where content is typically constrained to a fixed display, MR experiences integrate digital elements directly into the user's surroundings. This blending of real and virtual content increases cognitive demands, requiring users to allocate attention effectively to relevant information while filtering out distractions. In high-density MR environments, excessive or poorly managed visual stimuli can lead to attentional overload, reducing efficiency and increasing cognitive strain. Users must constantly process multiple layers of virtual and physical information, making attentional management crucial for effective interaction. In collaborative augmented reality (AR) workspaces, designers may struggle to manage an overwhelming number of holograms, comments, and user-interface elements, leading to mental fatigue [169, 461]. Similarly, in virtual reality (VR) training simulations, such as surgical training, excessive visual stimuli - physiological feedback in real time, vitality of patients, and dynamic 3D models - can increase cognitive load, reducing task efficiency and performance [10, 192]. In the context of VR gaming, where players must simultaneously process environmental signals, head-up displays, and real-time movement tracking, an excessive number of visual elements can affect reaction times and immersion [407, 570]. These examples highlight how an abundance of virtual content, if not properly managed, can negatively impact goal-oriented stimulus selection and the overall user experience.

Traditional MR interfaces often rely on a standardized design that does not account for individual differences in cognitive load tolerance or experience level. While some systems provide manual customization options, users without the necessary expertise may struggle to optimize interface settings based on their cognitive needs. Intelligent adaptive MR systems address this limitation by dynamically adjusting content presentation and interface complexity based on real-time cognitive state monitoring. This ensures that digital content remains supportive rather than overwhelming, allowing for a more efficient and user-friendly experience.

The ability of MR systems to dynamically tailor content presentation presents a significant opportunity to improve usability and accessibility. By leveraging physiological sensing and computational models, adaptive systems can detect the user's attentional state and environmental conditions, modifying content accordingly. During high-cognitive-load tasks, these systems can automatically prioritize essential information while suppressing distractions, preventing mental fatigue and enhancing task performance.

Achieving this vision requires an understanding of how users allocate attention and cognitive resources within MR environments. Human-Computer Interaction (HCI) extends the communication bandwidth between users and computational systems [152], facilitating more natural interactions. Instead of relying on explicit user input, physiological computing enables implicit adaptation by analyzing electrodermal activity (EDA), electroencephalography (EEG), and heart rate variability (HRV) [178]. This creates a more symmetrical interaction, where systems dynamically adjust based on cognitive fluctuations rather than requiring users to manually configure settings.

Developing such systems involves defining key cognitive parameters for user-adaptive MR interfaces. Thus, this thesis explores the integration of physiological computing in MR from three perspectives: (I) how different degrees of virtuality, whether task-relevant or irrelevant, impact users' state of engagement and their attentional allocation, (II) the design and evaluation of engagement and attention-aware systems in VR environments, and (III) the construction of physiological sensing systems directly embedded into the MR environment for unobtrusive, real-time, intelligent monitoring of user states.

1.1 Research Questions

Before integrating attention- and engagement-awareness into physiological computing MR systems, three fundamental aspects must be considered: the users and their surroundings, the available workload sensing modalities, and the task set. Our visual and attentional systems might face increased demands when blending virtual and physical elements. This suggests the need for MR systems to account for these perceptual differences in their design to enhance usability [303, 416]. In this thesis, we first identify three primary research questions (**RQ**) that are then investigated in depth regarding the different contributions.

To systematically investigate adaptive MR, we begin by exploring how attention and engagement fluctuate across different degrees of virtuality (**RQ1–RQ3**). Specifically, we first examine electrophysiological correlates of attention allocation and workload in MR environments

Table 1.1: Overview of the investigated research questions. The first part investigated how attention and engagement can fluctuate across different steps of the MR continuum. The second part examined physiological signals for evaluation and input in attention and engagement-aware adaptive systems. Finally, the third part combines both concepts and presents a system for contribution to designing wearable intelligent adaptive systems that are embedded in a MR setup.

	Research Question	**Paper**	**Chapter**
	How do attention allocation and task engagement fluctuate across the Mixed Reality Continuum?	-	3
RQ1	Do different MR manifestations impact performance and attention allocation?	[98]	3.1
RQ2	Does blending of virtual and physical objects impact our attention allocation?	[107]	3.2
RQ3	How do physiological correlates of attention and engagement vary across the continuum while engaged in a productivity task?	[97]	3.3
	Can we support user performance by considering physiological correlates of engagement and attention allocation as input for a Virtual Reality adaptive system?	-	4
RQ4	Do adaptations of secondary task difficulty based on physiological arousal modeled after task engagement support user experience?	[113]	4.1
RQ5	How do adaptations of secondary task difficulty impact physiological correlates of attention allocation?	[103]	4.2
RQ6	Do adaptations of environmental visual complexity based on physiological arousal modeled after engagement support task performance?	[108]	4.3
RQ7	How do adaptations of environmental visual complexity impact physiological correlates of attention allocation?	[104]	4.4
RQ8	Can we support task performance by designing an adaptive system that balances EEG correlates of attention allocation?	[102]	4.5
RQ9	Can embedded physiological sensing systems in MR interfaces be designed for signal reliability and validity?	[100]	5

using multimodal quantitative measures, providing insights into the design of engagement- and attention-aware systems (**RQ1**). We then investigate how users identify relevant blended information across the continuum MR and whether virtual and physical objects relevant to the task differ in search complexity depending on the nature of their surroundings (**RQ2**). Finally, we explore a more ecologically valid task—typing—to understand how different MR actualities influence attention allocation, engagement, and workload across the continuum (**RQ3**). This phase allows us to determine appropriate input channels for designing physiologically adaptive systems.

Building on this foundational understanding, we then shift our focus to developing adaptive Virtual Reality (VR) systems (**RQ4–RQ8**). Recognizing that attention allocation varies across the MR continuum, we investigate the feasibility of modulating secondary task difficulty and overall visual complexity to support primary task performance, while exploring a variety of input modalities (**RQ4** and **RQ6**). To achieve this, we employ a multimodal approach to assess the impact of these adaptations on various physiological channels, further refining the sensor space for future adaptive systems (**RQ5** and **RQ7**). The insights gained from these studies inform the design of an adaptive system that dynamically responds to users' attention and engagement states to enhance task performance, which we evaluate in a user study (**RQ8**).

Finally, we address the technical and methodological challenges of integrating physiological sensors into MR hardware (**RQ9**). To advance intelligent MR systems, we design and engineer a physiological sensing system embedded directly into a VR controller, enhancing the usability and interpretability of physiological data collected within MR environments.

1.2 Research Context

This thesis is founded on research undertaken within the Human-Centered Ubiquitous Media group at LMU Munich, spanning four and a half years. I performed my PhD through the primary supervision of Prof. Albrecht Schmidt, the leader of the research group and Prof.Sven Mayer. Throughout this duration, we collaborated with researchers and project partners from various institutions, contributing to organizing events and co-authoring papers.

Quantitative Methods for Visual Computing - SFB TRR 161 The majority of the work in this thesis was conducted in association with the Sonderforschungsbereiche (SFB) 161 – Quantitative Methods for Visual Computing. SFB TRR - 161 Collaborative Research Centre connected University of Stuttgart, University of Konstanz, Ulm University, and the LMU Munich. In the SFB-TRR 161, around 40 scientists in computer science, visualization, computer vision, HCI, linguistics, and applied psychology are working together to improve the quality of future visual computing methods. Project C06 focused on exploring the adaptability of MR systems to user states as determined by physiological measurements. This project investigated the dynamic blending of physical and virtual content in MR environments based on users' physiological data, such as gaze behavior, peripheral physiology (e.g., electrodermal activity Electrodermal Activity (EDA) and electrocardiography (Electrocardiography (ECG), and brain activity (i.e., Electroencephalography (EEG). The aim was to assess the feasibility and utility of MR scenarios that adjust the integration of virtual and physical components in real-time, according to user physiology.

Humane AI The HumanE-AI-Net project, funded by the EU, unites premier European research institutions, universities, and companies within a network dedicated to excellence. This initiative fostered collaboration between world-renowned AI labs and leading entities in human-computer interaction and the cognitive, social, and computing sciences. The project encourages researchers to broaden their perspectives, facilitating connections across the broader Artificial Intelligence (Artificial Intelligence (AI) landscape. The goal is to create AI systems that are robust, reliable, and capable of "understanding" human states, adapting to the complexities of real-world environments, and navigating through social interactions. HumanE-AI-Net is set to establish the groundwork for a new scientific discipline that aligns AI with European values, making it more relatable and beneficial for European citizens.

University Collaborations This thesis extensively investigates how physiological signals can be interpreted for usability evaluation and as an active input for interaction. With those application areas in mind, we also contributed to the organization of two workshops and an open document for preregistation of EEG experiments across more than 20 universities. The first workshop centered on the emerging field of Cognitive Personal Informatics (Personal Informatics (CPI)), focusing on wearable neurotechnologies for monitoring and interpreting cognitive activities [506, 507]. This initiative aims to enhance personal well-being and is underpinned by critically examining the ethical dimensions. The second workshop broadened the discourse to include integrating physiological signals into HCI systems, highlighting the importance of usability, ethical considerations, and interdisciplinary collaboration to address the complexities of incorporating physiological signals into interactive technologies [106, 394, 537]. Lastly, we supported the writing and implementation of an EEG preregistration template. This template guides researchers who intend to preregister their EEG projects, particularly those investigating event-related potentials (Event-Related Potential (ERP) [213]. Our involvement in creating this template shows our commitment to promoting rigorous and transparent research practices in the field of HCI for physiological computing.

1.3 Summary and Overview of the Thesis

This thesis comprises five chapters. The first part denotes background knowledge and differentiates existing work from the novel contributions presented in this thesis. The second part introduces the necessary background on physiological computing and mixed reality that laid this thesis project's foundation. The third part provides studies investigating how attentional and engagement states, based on a multimodal evaluation, can change the MR continuum across task settings with different levels of ecological validity. The fourth part introduces studies that designed and evaluated the sensor space for adaptive systems in Virtual Reality. The fifth chapter introduces a design approach to engineer and validate a physiological sensing system embedded in an MR system for physiological computing applications. Finally, the sixth part concludes with a summary of the contribution of this thesis and provides a framework for attention-aware systems in the domain of computational interaction. The seventh and last part includes the bibliography and supplementary material.

Chapter 1 - Introduction The first chapter describes the motivation and vision for adaptive mixed reality interfaces, highlighting the context and collaborations in which this thesis was conducted.

Chapter 2 - Background The second chapter introduces the physiological signals investigated in this thesis, their physiological inference, and the concept of attention allocation and engagement. We present state-of-the-art measures that facilitate explicit and implicit engagement and attention allocation measures. This is followed by an introduction by the guiding perspective of this thesis, i.e., physiological computing and recent developments and applications that support the vision of adaptive mixed reality.

Chapter 3 - Attention and Engagement in Mixed Reality Interacting in MR means engaging with enriched environments that might increase users' attentional requirements. Therefore, we first focus on how users discriminate target information from distractors across the MR continuum and then with a varying level of virtuality. Lastly, we evaluate how users are engaged in an ecological productivity setting while exposed to different actualities within the MR continuum.

Chapter 4 - Physiological Computing for Adaptive Virtual Reality This chapter provides an overview of the sensor design space for adapting and evaluating the visual complexity of a virtual environment based on physiological signal input. Moreover, by multimodal analysis, we propose a final system that can balance both attentional and engagement states.

Chapter 5 - Towards Wearable Physiological Computing With intelligent systems in mind, designing proper algorithms and physiological systems might still not be enough to reach a greater community. Thus, in this chapter, we propose a sensing device embedded in an MR system that can be employed for both physiological evaluation and interaction. We thus validate the system across various task settings, showing comparable signal validity as in medical-grade devices.

Chapter 6 - Discussion & Future Work In this chapter, we summarize and conclude the findings of this thesis and revisit the research questions stated in the beginning. We reflect on the research approach and present a framework for adaptive mixed reality. Lastly, we discuss limitations, opportunities, and future work for transitional interfaces and user state modeling that open up future applications in adaptive MR systems.

Chapter 7 - Conclusion This chapter consolidates the findings of the thesis, summarizing the contribution to user behavior and experience in MR, validation of adaptive VR systems, and the development of reliable physiological sensing within MR interfaces, highlighting the benefits of context-aware adaptations for enriched user interactions in immersive technologies.

2 RELATED WORK

> *"I know of no time in human history where ignorance was better than knowledge."*
>
> **– Neil deGrasse Tyson**

This chapter introduces three central themes in this thesis: physiological inference, its implications for physiological computing, and adaptive MR. We first provide an overview of attention, engagement, and their psychophysiological correlates and give context by discussing their applications in physiological computing in HCI. I then expand on adaptive applications in MR, primarily establishing the conceptual framework of MR and discussing previous work in adaptive MR.

This chapter is based on the following publication.

Francesco Chiossi, Johannes Zagermann, Jakob Karolus, Nils Rodrigues, Priscilla Balestrucci, Daniel Weiskopf, Benedikt Ehinger, Tiare Feuchtner, Harald Reiterer, Lewis L. Chuang, Marc Ernst, Andreas Bulling, Sven Mayer, and Albrecht Schmidt. 2022. Adapting visualizations and interfaces to the user. In *it-Information Technology, 64(4-5)*, 133-143.
`https://doi.org/10.1515/itit-2022-0035`

2.1 Attention and Engagement in Mixed Reality

MR environments challenge human attention and engagement due to the integration of virtual and physical elements. Unlike traditional computing interfaces, where interaction is confined to a fixed display, MR dynamically overlays digital content onto the real world, requiring users to allocate cognitive resources efficiently. Successful interaction in MR depends on a user's ability to focus on relevant information while filtering out distractions selectively. This makes attention allocation a critical factor in ensuring both usability and immersion. Similarly, engagement, which reflects the depth of a user's involvement in an experience, impacts performance and overall user satisfaction.

This section explores the theoretical foundations of attention and engagement in MR, examining how these constructs influence interaction quality and cognitive load.

2.1.1 Attention Allocation

Attention allocation refers to the process by which cognitive resources are distributed across multiple tasks and stimuli, enabling individuals to selectively process relevant information while filtering out distractions. In traditional computing environments, attention is primarily directed towards a single screen or interface. However, MR introduces additional complexity by seamlessly blending digital and physical elements, requiring users to dynamically manage attentional resources across multiple layers of information. The ability to effectively allocate attention in MR environments is critical for maintaining performance, usability, and immersion.

The interaction between virtual and real-world stimuli in MR demands a fine balance between perception and cognitive load. Users must engage with virtual objects, interfaces, and overlays while simultaneously being aware of their physical surroundings. This interaction is further complicated by the presence of both goal-driven (top-down) and stimulus-driven (bottom-up) attention mechanisms [117]. The extent to which users can effectively navigate these competing demands determines their ability to interact efficiently with MR systems. Poorly designed virtual content can lead to cognitive overload, causing distractions that reduce task efficiency and overall user experience.

Attentional States

Internal and External Attention Attention in Mixed Reality (MR) environments can be categorized into external attention, which focuses on processing sensory input from the environment, and internal attention, which governs cognitive states such as working memory, task goals, and response selection [117]. External attention allows users to engage with virtual and physical stimuli dynamically, directing perceptual resources toward elements like visual cues or auditory signals. In contrast, internal attention regulates mental processes, such as maintaining focus on a goal while interacting with a virtual interface. The interplay between external and internal attention becomes particularly relevant in MR systems, where the environment continuously shifts between virtual and real elements.

Endogenous and Exogenous Attention The mechanisms guiding attention allocation can be broadly divided into endogenous (top-down) and exogenous (bottom-up) attention [257]. Endogenous attention is voluntarily controlled and directed based on goals and intentions, such as when a user deliberately focuses on a specific object in an MR interface to retrieve information. Exogenous attention, in contrast, is involuntarily captured by salient stimuli, such as sudden changes in brightness, movement, or auditory alerts within the virtual space. In MR interactions, balancing these two attentional mechanisms is critical to ensuring usability. Highly salient virtual stimuli might disrupt a user's focus, causing frequent exogenous shifts of attention. Conversely, interfaces designed to enhance top-down control can facilitate immersive and goal-oriented engagement, improving task performance and situational awareness.

MR environments inherently engage both attentional mechanisms, as users must dynamically shift between goal-driven focus and stimulus-driven reorientations. According to the Corbetta and Shulman (2002) model of attention, these processes are governed by two distinct neural networks [123]. The dorsal frontoparietal network, comprising the intraparietal sulcus (IPS) and frontal eye fields (FEF), is responsible for top-down attention. This network enables users to direct focus toward relevant elements, such as virtual objects requiring interaction or interface controls for manipulating digital content. In contrast, the ventral frontoparietal network, including the temporo-parietal junction (TPJ) and inferior frontal gyrus (IFG), acts as a circuit breaker, detecting unexpected but behaviorally relevant stimuli. This bottom-up mechanism ensures that critical alerts or sudden environmental changes capture user attention when necessary.

In MR design, an optimal balance between these attentional systems is essential to maintain usability and minimize cognitive fatigue. Excessive reliance on bottom-up cues—such as intrusive pop-ups, unnecessary animations, or frequent alerts—can overload the ventral attention network, leading to distraction and performance degradation. Conversely, systems that do not adequately support top-down attention may reduce efficiency, as users struggle to find and engage with task-relevant information.

A well-designed MR interface should enhance goal-driven (top-down) attention by structuring the interface to support user objectives while minimizing unnecessary distractions. At the same time, it must regulate stimulus-driven (bottom-up) attention by ensuring that environmental cues capture focus only when contextually relevant. To achieve this, MR systems should incorporate adaptive attention control, dynamically adjusting the prominence of virtual elements based on real-time physiological and behavioral indicators. By carefully managing the interplay between these attentional processes, MR interfaces can reduce cognitive overload, facilitate task engagement, and improve user experience.

Selective Attention in MR Systems Given the high-density information environments created by MR, selective attention is essential for filtering relevant stimuli and suppressing irrelevant distractions [117]. Selective attention mechanisms help users focus on goal-relevant information while ignoring competing stimuli that could disrupt task performance. Without effective attentional filtering, users may struggle to maintain engagement, leading to cognitive overload and decreased usability.

Theories of selective attention propose that attention acts as a bottleneck, allowing only the most pertinent stimuli to be processed at higher cognitive levels. In MR interactions, this suggests that systems should provide context-aware filtering ensuring that users are guided toward the most relevant virtual and physical elements based on their task context. Several influential models of selective attention provide foundational insights into how attentional mechanisms operate and how they can be leveraged in MR system design.

Early theories, such as Broadbent's Filter Theory [73], proposed that attention functions as a strict filter that blocks unattended information from reaching higher processing stages. This model suggests that only one stream of information can be processed at a time, with other stimuli being completely ignored. While effective in explaining early-stage attentional selection, this model does not account for the ability to process multiple competing stimuli simultaneously, a common requirement in MR interactions.

In contrast, Treisman's Attenuation Theory [552] introduced the concept of an attenuation filter rather than a strict gate. This theory suggests that while the most relevant stimuli receive full processing, unattended stimuli are not completely blocked but instead undergo weakened processing. In the context of MR, this model supports the design of adaptive interfaces where background elements remain perceptible without distracting from primary tasks. For instance, semi-transparent overlays in MR interfaces can ensure that secondary information remains available without overwhelming the user's primary focus.

A more dynamic perspective is provided by Lavie's Load Theory of Attention [334], which argues that attentional selection is influenced by both perceptual load and cognitive control mechanisms. According to this model, when perceptual load is high (i.e., when the visual field is cluttered with task-relevant elements), the ability to process distractors is reduced. Conversely, under low perceptual load, distractors can more easily capture attention. This theory has direct implications for MR environments, where balancing visual complexity and attentional load is crucial. Interfaces can adapt by dynamically adjusting stimulus saliency based on real-time user workload, preventing unnecessary distractions when cognitive demand is already high.

Building on these models, Feature Integration Theory [553] emphasizes the role of bottom-up attention in selectively processing visual features such as color, shape, and motion. In MR environments, this suggests that saliency maps—highlighting high-contrast elements—can guide users toward task-relevant objects while ensuring non-essential features remain unobtrusive.

2.1.2 Engagement and Workload

Engagement in HCI is a multifaceted construct that has been conceptualized through various theoretical lenses. [413] define engagement as "a quality of user experiences with technology that is characterized by challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control, awareness, motivation, interest, and affect." This definition emphasizes engagement as both a process and an outcome of interaction.

From a process perspective, [518] describe engagement as "the process by which two (or more) participants establish, maintain, and end their perceived connection." This perspective highlights the temporal aspects of engagement, encompassing initiation, maintenance, and disengagement phases. More recent work identifies three primary dimensions of engagement. The first dimension, *cognitive engagement*, involves sustained mental effort, energy expenditure, and attention [510]. The second dimension, *emotional engagement*, includes interest, values, attitudes, and affective responses [67]. The third dimension, *behavioral engagement*, is reflected in observable actions and active participation in the interaction process [18].

Within neuroergonomics, engagement is linked to attentional mechanisms and arousal states, where optimal engagement is associated with a balance between cognitive load and user motivation [138]. This conceptualization aligns with contemporary research in adaptive interfaces that seek to enhance engagement dynamically by monitoring physiological and behavioral cues in real-time.

Mental workload refers to the cognitive demands imposed on an individual during task performance. [415] define workload as "the cost of performing a task in terms of attentional and cognitive resources." This construct has been central to human factors research, particularly in safety-critical environments where excessive workload can degrade performance.

One of the most influential theoretical frameworks for understanding workload is *Wickens' Multiple Resource Theory (MRT)* [592]. This theory posits that humans possess multiple, distinct pools of cognitive resources, and performance decrements occur when concurrent tasks compete for the same resource channels, leading to cognitive interference.

Beyond resource-based models, contemporary research in *neuroergonomics* has expanded our understanding of workload by examining neural correlates of performance [138]. [137] propose that performance decrements under high workload arise from multiple interacting neural mechanisms, including *prefrontal cortex inhibition, competition between neural networks, attentional shifts, and modulation of arousal states*. These insights emphasize that cognitive workload is not solely dictated by the availability of resources but also by how neural systems dynamically adapt to task demands.

Additionally, the *Dynamic Adaptive Theory (DAT)* [233] extends this perspective by arguing that the brain seeks homeostasis in response to varying task demands. According to DAT, both under-stimulation (*hypostress*) and overload (*hyperstress*) can degrade adaptability and impair performance. This framework reinforces the notion that workload is a non-linear construct influenced by environmental, cognitive, and physiological factors.

The Relationship Between Engagement and Workload

Engagement is defined as an effortful investment in the service of task or cognitive goals [376, 447], whereas arousal represents a state of physiological readiness to respond to external contingencies [178]. In interactive and adaptive systems, engagement reflects how users allocate cognitive and attentional resources toward a task, while workload describes the cognitive demands placed on them during performance. Achieving an optimal balance between these two constructs is crucial in HCI and neuroergonomics, as both under-stimulation and excessive cognitive demand can degrade user experience and performance.

Recent research suggests that the engagement-workload relationship is dynamic and non-linear, modulated by cognitive and neurophysiological factors [137]. Under moderate workload conditions, engagement is sustained, supporting attentional focus, task persistence, and goal-directed behavior. However, when cognitive demands exceed available resources, users experience attentional lapses, task disengagement, and degraded performance. Conversely, low workload conditions—such as prolonged passive monitoring—fail to sustain engagement, leading to mind wandering and reduced vigilance [525]. This variability in engagement highlights the necessity of adaptive systems capable of regulating workload dynamically based on user state.

2.1.3 Neurocognitive Mechanisms in Engagement and Workload

Engagement and workload are mediated by multiple, interacting neural networks. Research in neuroergonomics identifies the prefrontal cortex (PFC) as a critical hub for managing cognitive effort, attention allocation, and executive control [137]. Under high workload, functional disengagement of the PFC can occur, leading to attentional lapses, inefficient decision-making, and reduced situational awareness. This phenomenon, termed cognitive tunneling, results in users becoming overly fixated on a single strategy while ignoring environmental cues [86].

Beyond the PFC, the anterior cingulate cortex (ACC) is key in detecting cognitive conflicts and modulating task engagement based on perceived difficulty [376]. Increased workload is associated with heightened ACC activity, signaling the need for adaptive control mechanisms to sustain engagement. However, executive overload can cause functional disengagement under extreme task demands, leading to decision errors and task abandonment [137].

Another key mechanism influencing engagement is arousal regulation, governed by the locus coeruleus-norepinephrine (LC-NE) system [19]. The LC-NE system modulates attentional effort based on task demands, ensuring that cognitive resources are allocated efficiently. In adaptive MR systems, monitoring pupil dilation, heart rate variability, and EEG markers of LC-NE activity can provide real-time insights into cognitive engagement states, enabling interfaces to adjust workload dynamically.

2.1.4 Cognitive and Environmental Factors Influencing Engagement and Workload

Several cognitive and environmental factors modulate how workload affects engagement. Tasks with moderate variability in difficulty promote sustained engagement by providing a continuous cognitive challenge. However, highly repetitive or excessively complex tasks can lead to vigilance decrement or overload, respectively [137]. Individual differences such as cognitive capacity, prior expertise, and motivational states significantly influence workload tolerance and engagement levels. Users with higher working memory capacity can maintain engagement for longer under higher workload conditions, whereas individuals with lower cognitive resources may experience faster disengagement [376]. Environmental stressors, including distractions, time pressure, and multi-tasking demands, further modulate workload-engagement relationships. High-noise environments or time constraints can increase cognitive effort while reducing perceived control, leading to stress-induced performance degradation [86].

2.2 Psychophysiology

The foundation of physiological computing is significantly derived from the discipline of psychophysiology. Psychophysiology is the study of the interrelationships between physiological and cognitive processes. Psychophysiology synthesizes elements from anatomy, physiology, and cognitive science and is closely related to behavioral neuroscience [83].

A central concept in physiological computing is its application in HCI research, particularly in understanding users' physiological reactions to multimedia and multisensory content. This encompasses interactions with virtual environments, physical interfaces, haptic feedback, and more diverse forms of interaction. This represents a crucial part of the field, as interacting with computers often involves multimodal stimuli, which are more complex than the simple cues used in traditional psychophysiological studies [470].

2.2.1 Mapping Cognitive Processes to Physiological Signals

In this thesis, we focus on four physiological signals, i.e., electrodermal activity, electrocardiography, electroencephalography, and electroencephalography, exploring their analysis and interpretation. While this work is not centered on signal analysis, understanding these signals is crucial for physiological computing and identifying specialized applications.

Electrodermal Activity

EDA, also termed Skin Conductance (SC) or galvanic skin response (GSR), is a critical measure of the skin's electrical changes, primarily influenced by sweat secretion. The human skin features two distinct sweat gland types: the apocrine and eccrine glands. The eccrine glands, strategically located on the palms and soles, are pivotal in psychophysiological studies, as their activation is linked to emotional or arousing stimuli, distinguishing them from the thermally-regulated apocrine glands. This distinction is relevant for HCI researchers who choose EDA to evaluate the impact of interaction with machines level of psychological, physiological, and emotional arousal [12]. Additionally, EDA is associated with responses to specific short-term events, including reactions to novel stimuli, mental workload, and cognitive appraisal of stimuli. The EDA signal comprises two key components: the tonic signal, i.e., Skin Conductance Level (SCL), which is slower and reflects the user's affective state or stress level, and the phasic component, i.e., Skin Conductance Response (SCR), characterized by faster, event-related spikes. Further measures of tonic EDA have been suggested - such as the frequency of Not-Specific Skin Conductance Response (NSSCR)(typically 1-3 per/min during rest and over 20 per/min in high arousal situations). The intensity of a phasic spike generally corresponds to the strength of the stimulus that triggered it. Analyzing EDA involves separating these components, often starting with the removal of the tonic activity. This process can be challenging, especially when multiple phasic spikes overlap, as they build upon each other. While deconvolution-based algorithms can isolate these spikes, their processing speed may not be ideal for real-time applications.

In the studies reported in this thesis, EDA was used to predict engagement in section 4.1 and in section 4.2 [108, 113]); as one of the signals used to evaluate the effect of adaptations in VR [103, 104], see section 4.3 and section 4.4. Lastly, we validated EDA recordings from a medical-grade device to the ones acquired with sensors embedded in a MR system [100], see chapter 5.

Electrocardiography

Electrocardiography (ECG) records the heart's electrical activity through chest electrodes. In ECG research, measurements are typically evaluated over several minutes, thus without requiring high-frequency data recording. From the ECG signal, the intervals between heartbeats are analyzed, and these inter-beat intervals are crucial for deriving metrics like Heart Rate Variability (Heart Rate Variability (HRV)) or heart rate (Heart Rate (HR)). These metrics are instrumental in assessing cognitive states and processes in psychophysiology. There are multiple techniques for HR measurement. Photoplethysmography (Photoplethysmography (PPG)) involves using light transmission through tissues like a finger, while ballistocardiography captures heart rate by detecting subtle body movements caused by the heartbeat. Another innovative method uses low-cost webcams to track heart rate through facial video analysis, similar to PPG's light-based approach. However, ECG remains the most prevalent method, capturing heart rate through the heart's electrical activity. ECG's practicality is shown by its widespread use in consumer fitness devices. Unlike EDA, the heart uniquely responds to the sympathetic and parasympathetic nervous systems. The sympathetic nervous system, known for initiating fight-or-flight reactions, typically increases the HR and, conversely, decreases the HRV. In contrast, the parasympathetic system, which facilitates restful states, often reduces it.

In the studies reported in this thesis, ECG was employed to evaluate physiological responses of stress and arousal to adaptations of visual complexity [103, 104], see section 4.3 and section 4.4. ECG, as HR, was then chosen to evaluate engagement in a typing task in MR [97], see section 3.3. Last, as for EDA, we validated HR acquired from PPG recordings, comparing medical-grade data to the ones acquired with sensors embedded in an MR system [100], see chapter 5.

Eye Tracking

Eye-tracking research, tracing its roots to before the 19th century, initially served to decipher perceptual and cognitive aspects of eye behavior. In HCI, it's a means of interaction through eye movements. Human eye movements consist of fixations, where one stares steadily at a point to process a still image, and saccades, which are rapid movements between points, during which visual intake is paused [472]. Most modern trackers use optical methods, projecting infrared light onto the eye and analyzing the reflected light to determine eye rotation.

The incorporation of this technology into MR systems has found applications for market research [4], user interface design [196], and the study of human factors like mental workload and engagement [466]. Many modern head-mounted displays (HMDs) are equipped with eye-tracking capabilities to introduce new interaction methods and optimize the computationally intensive process of stereoscopic rendering.

Eye tracking generates two principal types of data: fixations, where the gaze is stable, allowing for detailed perception of a visual scene, and saccades, quick eye movements that scan the environment [255]. These movements are critical for analyzing how users navigate and interact with MR environments, providing insights into visual interest, attention allocation [56], and cognitive load [614]. Advanced techniques like video-oculography offer high-resolution gaze tracking, opening up possibilities for designing attention-aware user interfaces and enhancing the realism and responsiveness of MR systems [572].

Moreover, the measurement of pupillary responses, or pupillometry, offers an additional layer of insight. Pupillary changes, driven by external stimuli, such as light exposure, and internal cognitive and emotional states, can proxy for cognitive load and task difficulty [331]. Variations in pupil size reflect the intensity of mental effort and processing demands, offering a real-time gauge of user engagement and cognitive state [458]. However, interpreting these changes requires careful consideration of environmental factors affecting pupillary responses.

Eye movement metrics, particularly fixations and saccades, are indispensable for evaluating search behavior in MR [186, 384]. These metrics reveal the cognitive strategies users employ to locate targets, reflecting the efficiency of information processing and attention allocation. Fixation duration and count, for example, provide clues about the cognitive effort involved in identifying targets among distractors [78]. Similarly, saccade patterns offer insights into the navigational strategies and the influence of task complexity on visual search efficiency [166]. Understanding these dynamics is crucial for designing MR environments that support effective visual exploration and interaction.

In the studies reported in this thesis, eye tracking was employed to evaluate visual search efficiency in MR [98], see section 3.1, and when searching for physical and virtual distractors [107], see section 3.2. Finally, we evaluated cognitive workload based on the Index of Pupillary Activity (IPA) in an MR typing task [97], see section 3.3.

Electroencephalography

Directly recording brain activity allows applications that depend on mental or cognitive metrics, such as perceived relevance. Electroencephalography (EEG) is pivotal in this context, offering a non-invasive method to capture the brain's electrical activity through electrodes placed on the scalp. This technique measures the cumulative post-synaptic potentials from numerous neurons, providing insights into the brain's functioning with high temporal accuracy. The technology's origin traces back to 1924 when Hans Berger, who also introduced the term "electroencephalogram," made significant advancements in measuring human brain activity using scalp electrodes. However, the first recordings of electrical brain activity were conducted by Richard Caton in 1875, observing the activity in monkeys and rabbits. EEG's foundation lies in the understanding that communication between neurons occurs through neurotransmitters or ions, and the excitation of action potentials, which results in detectable electrical impulses and current changes. EEG signals are generated by the synchronized activity of millions of electrically charged neurons, creating ion waves that interact with electrodes. This interaction can either transfer or withdraw electrons from the electrode's metal, and the variance in electron movement is captured as electrical voltage over time. These voltages, typically ranging between -100 and 100 μV when measured at the scalp, are recorded and can be visually represented, with time on the x-axis and voltage on the y-axis for each electrode.

In addition to analyzing general patterns of brain activity across different regions and over time, specific neuronal phenomena can be examined through the frequency domain, including time-locked or evoked potentials. In the context of this thesis we will review EEG correlates of attention allocation and distractor suppression.

Frequency Domain EEG frequencies, the rhythmic electrical activity generated by the brain, are crucial markers of cognitive states and processes. These frequencies are segmented into distinct bands, each correlating with different brain function and attention aspects [245]. The segmentation facilitates the understanding of human cognition, particularly how attention is allocated and managed across various tasks and environments, including immersive technologies like VR and MR [529].

When analyzing EEG data in the frequency domain, various frequency bands are typically examined: Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-14 Hz), Beta (14-30 Hz), and Gamma (30-70 Hz) [245]. The classification of these bands is derived from empirical research and may show minor variations across different studies. The Delta band, associated with deep sleep and unconscious processing, is crucial in memory consolidation and indicates relaxation or meditation in waking states [314]. The Theta band is linked to drowsiness, early sleep stages, and deep meditation [294]. Still, it is also tied to attention-related activities such as memory encoding and problem-solving, reflecting the brain's engagement in complex cognitive tasks [312, 325]. Alpha frequencies, considered the brain's resting state, are deeply involved in attention dynamics, decreasing during mental tasks or when the eyes are open to indicate a shift from internal focus to engagement with external stimuli [311]. This shift is particularly relevant in VR, where Alpha activity modulation signals how external stimuli affect attention. The Beta band represents active, conscious thought and is associated with alertness and decision-making, with high Beta levels pointing to excitement or anxiety during demanding tasks [245]. Lastly, the Gamma band, correlating with higher mental activities like perception and consciousness, signifies the brain's ability to focus intensely and process complex stimuli, underscoring its role in integrating sensory information [35].

Within the scope of this thesis, attention in VR and MR can be broadly classified into two categories: internal and external, see Figure 2.1, following the taxonomy of Chun [117]. Internal attention, or endogenous attention, involves the brain's focus on internal states, such as recalling memories or solving mental calculations. This form of attention is crucial for tasks that require reflection, introspection, or manipulating information within the users' mind.

External attention, or exogenous attention, on the other hand, is directed towards the environment. In VR and MR, this type of attention is engaged by the sensory stimuli the virtual or augmented reality system provides. The immersive nature of these technologies makes understanding external attention particularly important, as it influences how users interact with and perceive virtual environments. Research has underscored the significance of Alpha and Theta frequency bands in attention [120, 245]. The Alpha band is traditionally linked to the brain's idle state but decreases in amplitude in response to visual engagement or when attention is externally directed. This phenomenon supports the notion that Alpha activity modulation is a mechanism by which the brain filters relevant from irrelevant stimuli. Theta activity, conversely, has been associated with cognitive effort or engagement, particularly in tasks that demand working memory or sustained attention. An increase in Theta power, especially in frontal regions of the brain, signifies intense cognitive engagement, often related to internal attention tasks. In VR and MR settings, understanding the dynamics of Alpha and Theta bands can inform the design of adaptive systems that optimize user experience based on attentional state. For instance, systems can adjust the complexity of a task or the intensity of sensory stimuli based on real-time EEG data, tailoring the experience to maintain a balance between engagement and cognitive load. Research has explored the role of these frequency bands in scenarios where users must divide their attention between internal tasks (such as mental arithmetic) and the external, immersive environment. Findings indicate that detailed VR environments can capture external attention like traditional reading tasks, as evidenced by changes in Alpha power [364, 365].

Event-Related Potentials Event-related potentials (ERPs) represent significant brain electrical activity tied to specific internal or external events, including sensory, cognitive, or motor occurrences [358]. Essentially, ERPs are variations in brain electricity synchronously linked to specific events, allowing for time-domain analysis of EEG signals post-stimulus introduction. The waveform of an ERP signal comprises sequential positive and negative voltage shifts, reflecting various underlying brain responses. These responses are critical for interpreting the brain's reaction to stimuli, categorized by amplitude and latency, informing on a individual and group-level information processing.

ERPs are typically named for their polarity (positive or negative) and the timing of their occurrence post-stimulus or by their sequence in the signal's waveform. For instance, an early negative peak roughly 100 milliseconds after a stimulus might be labeled N100 or N1 [109, 588], illustrating the standardized nomenclature based on polarity and timing or sequence. The timing of ERP components, especially those linked to cognitive stimulus processing, can vary widely. The P300 component, known to emerge upon detecting infrequent targets,

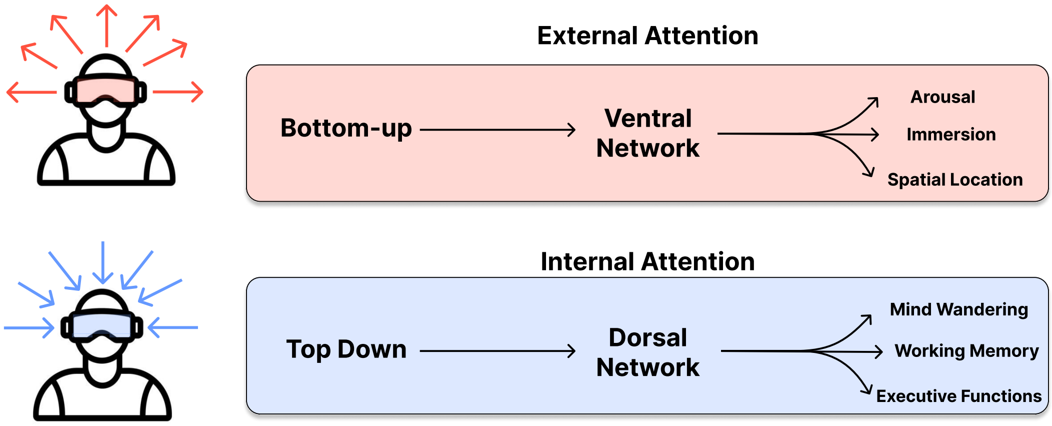


Figure 2.1: This taxonomy integrates traditional concepts of attention, distinguishing between bottom-up and top-down attention. The former is external, where environmental stimuli direct attention, often linked to higher immersion due to digital or virtual stimuli, initiating an 'arousal' response as the first perception. The latter, internal attention, is 'executive' and involves cognitive selection, typical in high cognitive load scenarios but also when external stimuli achieve their focus objective. This dual path illustrates attention's complex nature, as depicted in Chun et al. [117].

exemplifies this variance with peak timings between 250 and 700 ms [445].

Investigating how attentional resources are allocated in both real-world and Mixed Reality settings is vital for visual neuroscience and HCI. The high temporal resolution of EEG makes it an ideal tool for detailed examination of attentional dynamics, especially in MR environments where user interaction and experience are dynamic and multifaceted.

In this thesis, we investigated visual search in MR, [98, 107], see section 3.1 and section 3.2, where understanding visual attention and distraction with high overlap of virtual and physical information. Here, ERPs offer a window into these processes by revealing specific brain activity patterns in response to stimuli or events. In particular, the distractor positivity (Distractor Positivity (PD)) component, a late positive voltage deflection observed after the onset of a visual search display, serves a key role in the cognitive control mechanisms that suppress distracting stimuli to support visual search efficiency [277, 347]. The PD component is particularly significant in scenarios requiring feature-based search over singleton detection, where increased PD amplitudes denote greater cognitive effort to manage or inhibit distractions [197].

Despite its relevance, the exploration of PD and its role in distractor suppression within MR environments is relatively uncharted. Initial research by Marini et al. [373], which compared the brain dynamics elicited by real objects against those by planar images, indicated that real objects provoke more intense and sustained neural activity related to visuomotor planning. This suggests a profound impact of object physicality and affordance on attentional resource distribution and distraction inhibition, hinting that in MR settings—where real and virtual elements intertwine—the PD response could significantly deviate from traditional physical contexts. This variability indicates that distractor positivity as a promising indicator when probing distractor suppression mechanisms [197].

Fixation-Related Potentials Fixation-related potentials (Fixation-Related Potentials (FRP)) are a subset of ERPs, centered on brain activity tied to eye fixations. Here, eye fixations mark the event for time-locked electrophysiological activity analysis [136]. This method necessitates the integration of EEG and eye-tracking technologies to synchronize and correlate eye fixations with EEG data accurately. FRPs are instrumental in gauging task difficulty [479], delineating periods of heightened attentional effort, and differentiating exploratory phases from active interaction periods [300].

This allows for investigating attentional processes as eyes move to assimilate information from the environment, enabling the analysis of complex, real-world visual stimuli. Initial FRP components, observed during fixed gaze tasks within the 150-300 ms timeframe, relate to stimulus identification and perceptual processing, offering insights into early cognitive engagements with visual stimuli.

Further research into FRPs, particularly late components associated with target fixations during visual search tasks, connects with established EEG markers of cognitive resource allocation like the P3 component. Studies have demonstrated that FRPs locked to target fixations elicit significant late cognitive potentials akin to the P3, indicative of deeper cognitive processing in controlled and free-viewing paradigms [300, 479, 600]. This connection suggests FRPs can provide implicit insights into the relevance of on-screen items, potentially benefiting online application design. However, the influence of low-level visual features and task-related demands on FRP components remains an area ripe for exploration. Given the inherent visual complexity and information richness of MR environments, FRP activity could reveal varying task demands, affecting FRP component amplitudes in later processing stages. Confirmatory studies in controlled settings have illustrated FRP's capacity to reflect and differentiate cognitive resource allocation across different task conditions [464, 479, 597].

In this thesis, we employed multiple EEG features to inform our research. In section 3.1 we investigated resource allocation using FRP when searching in MR [98], see section 3.1. We used ERPs, specifically PD, to investigate distractor and target processing in MR [107], see section 3.2. Finally, we evaluated allocation of attention and engagement in MR using EEG alpha and theta frequencies, both for evaluation in a MR typing task [97], and in a N-Back visual working memory task in VR [103, 104], see section 4.2 and section 4.3. Lastly, we employed EEG frequencies as input for an adaptive system in [102], see section 4.5.

2.2.2 Wearable Physiological Sensing

Integrating wearable physiological sensing technologies into MR systems presents a unique opportunity to create intelligent adaptive environments. These technologies can offer real-time insights into a user's cognitive and emotional states, enabling MR systems to adjust dynamically to enhance user engagement, improve learning outcomes, and tailor experiences to individual needs. Integrating wearable physiological sensing technologies into MR systems is transformative, primarily because these devices ensure ubiquity and seamlessness in capturing crucial physiological data. The wearable nature of these sensors means they can be comfortably used in a wide range of settings, from clinical environments to everyday scenarios, making physiological monitoring unobtrusive and continuous. This ubiquity is pivotal for developing intelligent MR systems that are not only adaptive within controlled environments but also capable of extending their adaptability to real-world settings.

Moreover, the wearable aspect of these sensors supports the development of ubiquitous intelligent MR systems. Such systems can seamlessly integrate into users' daily lives, offering personalized assistance, learning, and entertainment tailored to their moment-to-moment needs and states. This ubiquitous intelligence paves the way for MR applications that support users in various contexts, from aiding in task performance and decision-making to providing therapeutic interventions tailored to the user's emotional state.

However, advancing the application of physiological sensing in virtual reality demands a comprehensive understanding of both psychophysiological principles and integrating physiological signals into VR environments [6]. However, the current challenge lies in the cumbersome process of implementing and synchronizing these technologies within VR settings, often hindered by impractical hardware-software solutions [438].

Researchers are required to separately prepare and calibrate each sensing device for collecting different physiological signals, which significantly prolongs the setup phase. Furthermore, the immersive nature of VR means users are often unaware of their own bodies and the attached sensing equipment, leading to potential complications from entangled cables or interference with VR tracking, consequently affecting signal quality and user immersion. To overcome these obstacles, there has been a surge in the development of straightforward, wearable, and user-friendly data collection solutions [114, 317, 346].

Nevertheless, commercial-grade wearables frequently restrict access to raw data, relying instead on proprietary metrics defined by the device manufacturers. This limitation has encouraged researchers to pursue custom integration solutions within VR systems. Incorporating physiological sensing directly into VR environments offers several benefits, including a unified device for both hardware and software, resulting in enhanced user comfort, cost-effectiveness, and potentially greater acceptance of the technology [209]. Various VR headsets now feature embedded physiological sensors, supporting gaze and eye tracking [321], along with recording multiple physiological signals [48]. For example, Bernal et al. combined high-quality sensors for capturing a range of signals such as facial electromyography, electroencephalography, electrodermal activity, and eye-tracking data within MR environments [48]. While EEG and surface Electromyogram data benefit from being recorded at optimal locations on the head, EDA signals from the forehead have shown limited correlation with more responsive body sites, reflecting their greater sensitivity to thermoregulatory rather than arousal activities [431]. Unfortunately, these innovative devices have yet to be extensively validated in real-world VR settings, with most comparisons against medical-grade equipment conducted in laboratory environments [387, 392].

The integration of physiological sensors with VR facilitates research into cognitive and emotional factors under conditions that closely mimic real-life scenarios. Devices like EmteqPro [210], Galea [47], and platforms offering accessible data analysis frameworks [459] have become invaluable in measuring psychophysiological aspects of humans and in developing intelligent virtual environments. This ecosystem of commercial and open-source options introduces physical signals into VR, enhancing interactivity and immersion. Innovations from the hardware sector, including EEG sensors designed for compatibility with VR headsets from companies like Galea [47], LooxidLabs, NextMind, and Neurable, alongside other modalities from the EmteqPro mask [210], which offers facial electromyography and PPG, are expanding the horizons of VR experiences. Additionally, certain VR headsets now provide built-in capabilities for eye-tracking and pupillometry, such as the HTC Vive Pro Eye, Varjo XR-4, and Meta Quest Pro, enriching the data available for creating responsive and adaptive MR systems.

In this thesis, we explored the EDA and PPG sensor space and evaluate the feasibility for embedding them in a MR system in [100], see chapter 5.

2.3 Physiological Computing

The first stage of implementing adaptive interaction systems, according to the definition, is to define the desired “goal.” In a system (a set of connected variables), such a goal is a state or multiple states (specified values for those variables at a certain point in time) that are chosen above others [133]. In order to achieve this, such systems require developing a user model that represents preferences, capacities, and affective processes and their relationship with the task at hand. This cybernetic process requires two components: a sensing component that detects the current state of the system and an actuation component that guides the system toward the desired goal depending on the sensing component.

The user provides multimodal information, both explicit and implicit, that can drive an interface or visualization adaptation toward a (shared) goal. To date, adaptive systems have mainly exploited direct user input as interaction modalities: the computer reacts only to explicit commands provided by the user, e.g., mouse, keyboard, speech, touch. However, recent approaches are increasingly considering implicit aspects of the user, such as their cognitive processing capabilities [299] and the user’s physiological state [178]. For example, in the field of visualization, a personalized adaptation based on various cognitive functions (such as perceptual speed and working memory) impacts the user’s performance [549] and different modalities of information processing require different visualizations [535]. This complements traditional approaches of considering implicit user data based on their profile, e.g., interests or prior knowledge.

Monitoring the user's physiological state to infer and adapt interaction will couple them to the user's goals and, thus, enable developers to design biocybernetic loops. The user's physiological data are processed online and classified to trigger the adaptive system, which is then in charge of performing adaptive actions in the interface [178]. Fairclough [177] proposed first and second-order adaptation loops. The first-order adaptation consists of a loop that begins with monitoring the user's condition. The loop completes by executing adaptive actions. This first-order adaptation necessitates a set of rules that link each user's current condition to at least one adaptive action. The second-order adaptation encompasses detecting changes as a direct result of adaptation. It allows the system to acquire information on the user's state and preferences over multiple iterations. Thus, it allows the system to adjust the actions to a single user. And after a phase of reciprocal coupling, it leads to system and user co-existence.

Biocybernetic approaches do not consider individual users as static entities. Therefore, we describe such dynamic systems as continuously updating systems using incoming environment information and, thus, changing user requirements and goals [171, 605]. Such dynamics are critical, especially when the update of the adaptive interface is not under the explicit control of the user but depends on the characteristics of the interaction itself. Thus, both the adaptive interface and the user learn from each other. Mutual adaptation dynamics can lead to complex interaction patterns, affecting the adaptive system's usability. A paradigmatic example consists of adaptive interfaces designed to improve the user's performance by automatically reducing the error associated with a given task, e.g., an adaptive touch keyboard [185]. If an interface fails to incorporate the user's learning capability, the performance of the joint system will likely be even worse than that of a user in isolation. In such cases, the contribution of the adaptive interface may result in error overcorrection and hence in a potentially unstable interaction [31]. In contrast, the system's features will change according to the characteristics of the user, for example, by providing only partial correction and taking into account the learning rate associated with the user over the entire course of the interaction. In the future, such a joint adaptation approach can positively improve the outcome of the interaction and the user's subjective experience. The next generation of intelligent systems will encompass increased autonomy and adaptability [256] facilitating proactive and implicit interaction with users [1].

This work provides an overview of applications in adaptive content, interaction, and visualization. We specifically address which information researchers used for adaptation, with which purpose of using such information, and lastly, which domains are feasible for adaptive systems in the HCI and the visualization domains.

2.4 Technologies for Adaptation

Systems nowadays may draw from various information to infer the user's current state and environment. Such information ranges from static adaptation using user profiles to context-aware systems [144, 502], using the user's surroundings and even ubiquitous sensing technologies. The last category potentially provides deep insights into a person's cognitive ability or motoric skills. In the following, we provide a short overview regarding currently common physiological measures leveraged in adaptive interfaces.

2.4.1 Extracting Features for Adaptation

Among the psychophysiological measures useful for adaptive interfaces, EEG, ET, EDA, and Electromyogram (EMG) have garnered the most interest. Their sensors' comparatively small size and ability to measure physiological activity non-invasively make them more likely to be incorporated into wearable consumer devices, such as glasses, wristwatches, and headbands.

EEG records electric potentials from the scalp, which reflects brain activity. Machine Learning (ML) can extract event-related activity to estimate cognitive workload [315], attention allocation [571], or affective states [398]. Brain-machine interfaces [90] and user state estimation systems use these ML-generated estimates. However, such systems have to be considered in light of current challenges such as the need for generalizable applications of classification methods online [400], improvement of transfer learning, and application of new approaches such as deep learning or Riemannian geometry-based classifiers [357].

Similarly, gaze behavior can indicate high-level cognitive processes, see early work by Deubel [142] and Hoffmann [253]. Recent work analyzed specific eye movements and gaze patterns to infer, for instance, user activities and cognitive states. Jacob and Karn [274] and Duchowski [159] provide in-depth overviews of this domain.

When the goal is to infer responses to novel stimuli, cognitive workload, and stress, the choice of EDA measures might be preferable as a noninvasive and easy-to-use method. EDA measures are the joined pattern of its phasic and tonic components [453]. Phasic Skin Conductance Response (SCR)s reflect discrete and stimulus-specific responses to evaluate the novelty, importance, and intensity of the stimuli utilized [411]. As indexed via SCL, tonic activity is an inertial and slow response particularly well suited to evaluate the effect of continuous stimuli, i.e., task. Therefore, HCI has used it to quantify, for instance, changes in arousal under high cognitive load [317] or stress [62].

Besides inferring cognitive processes, measuring the user's motoric responses can be especially useful in enabling a system to adapt to the user's abilities and potential actions [389]. By measuring muscular activity, EMG can provide insights into the working mechanism of motor tasks. Using EMG measures for adaptation allows for providing user-tailored feedback, ranging from detecting emotional states through facial EMG [576], over gesture recognition [496], to an adaptive tutoring system for motor tasks [296].

Finally, sensor fusion multi-model adaptive systems can often achieve more robust adaptation. For example, Putze et al. [456] showed that combining EEG recordings with eye-tracking addresses the Midas-Touch problem in gaze-based selection by estimating whether a fixation was purposeful or not. Moreover, combining EMG with EEG, Haufe et al. [241] showed that this leads to faster automatic braking in a driving simulator than using EEG alone.

2.4.2 Adapting the Interface and Visualization

Although there have been earlier models for adaptive systems [234, 503], we consider three critical adaptation elements: content, presentation, and interaction. When adaptive systems adjust their content, which relates to users' preferences and engagement, they must consider the user's prior knowledge and interest. Such dimension might involve notification design and recommendations, especially considering the exploratory visual analytics process [490].

Secondly, presentation adaptation affects User Interface(User Interface (UI)s) or visualizations according to users' spare perceptual capacity, discomfort and, stress level by simplifying displayed information, luminance, or other properties.

Thirdly, interaction adaptation is a broader field as it might encompass different paradigms. For example, in multitasking environments, users might experience tasks being switched off [454], see the number of options change in a decision-making task [448], or modify the interaction modality, i.e., from gesture to hand-free interaction.

2.5 Mixed Reality

Throughout the evolution of computer technology, the interaction between humans and computers has undergone significant transformations approximately every two decades [21]. In the early 1960s, command-line interfaces were the norm, necessitating a familiarity with specific commands and a reliance on text to navigate the digital interface. By the 1980s, graphical user interfaces (Graphical User Interface (GUI)) emerged, introducing the mouse as an input device and significantly lowering the barrier to computer interaction by eliminating the need for command memorization. GUIs allowed users to navigate the digital space visually, though direct interaction with on-screen elements was still mediated through a pointer.

The subsequent shift to natural user interfaces in the early 2000s further democratized digital interaction, enabling users to engage with content directly through touch or stylus, moving closer to a tangible interaction with digital content. Despite these advancements, the interface remained confined within the bounds of a physical screen, a mere "window" into the digital realm.

The advent of Mixed Reality (MR) technology marks the latest paradigm shift, offering a leap through this "window" to merge the physical and digital worlds in unprecedented ways. MR encompasses both Augmented Reality (AR), Augmented Virtuality (AV), and Virtual Reality (VR), technologies that modulate human perception by integrating virtual objects into our environment (AR), integrating physical ones into a virtual environment (AV), or immersing us completely within a digital space (VR). These experiences are enhanced by Head-Mounted Displays (HMDs) that track and respond to head movements, facilitating a mobile and immersive interaction.

The concept of Mixed Reality (MR) was first introduced by Milgram et al. in the mid-1990s [390], defining it as a continuum that ranges from fully physical environments to fully virtual ones, with MR encompassing all intermediate stages. VR was initially positioned at the extreme virtual end of this continuum. However, because VR systems require physical hardware—such as HMDs and controllers—to facilitate user interaction, they maintain a connection to the physical world. This integration within the MR framework is further reinforced by industry leaders like Microsoft, Apple, and Meta, who classify their VR headsets as immersive MR devices. Their categorization highlights how the boundaries between purely virtual experiences and those incorporating real-world elements are becoming increasingly blurred.

In this thesis, VR is considered an integral part of the MR ecosystem, reflecting a broader understanding that any virtual experience necessitates some form of real-world interaction or manifestation.

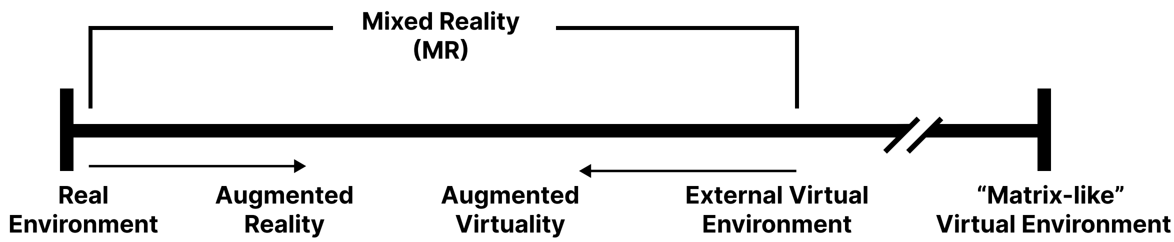


Figure 2.2: Reality-virtuality continuum by Milgram et al. [390] . Here, we consider Virtual Reality as a part of the continuum.

In the following subsection, we will briefly overview the history of VR, Augmented Virtuality (AV) and Augmented Reality (AR) technologies and define the terms. Further, we will highlight the characteristics of all technologies. We will conclude this section with the advantages of adaptive mixed reality and related application space.

2.5.1 Virtual Reality

The concept of VR, once a staple of science fiction like StarTrek’s Holodeck, was formally conceptualized by Ivan Sutherland in 1965 as "the ultimate display," a vision of immersive environments where virtual objects have tangible effects [542]. This ambition traces even further back to the 19th-century stereoscopes, which laid the groundwork for immersive visual experiences, evolving through advancements like the Head-Mounted Displays (LEEP) optics in the late 1970s for wider fields of view and the Sayre Glove in the 1980s for interactive hand gestures in virtual spaces [540].

The 1990s introduced the Cave Automatic Virtual Environment (CAVE) system, immersing users in a digitally constructed space that tracked their movements [126]. Interest in VR declined by the mid-1990s but was revived in the 2000s with the advent of smartphones, which provided the necessary high-resolution displays and sensors. This resurgence was started by Palmer Luckey's 2012 Kickstarter campaign for a consumer VR headset, highlighting the close relationship between VR and mobile technology advancements. In the present day, we see a continuous effort in developing and supplying VR headsets tailored to specific functionalities to the consumer market, from productivity as the Varjo XR-4 to gaming as in PS-VR from Sony.

VR's definition has evolved, focusing on presence as its core characteristic [524]. Immersion and presence, the sensation of being in the virtual environment, are central to VR's effectiveness [598]. The difference between presence and immersion in VR contexts merits clarification. Immersion refers to the technical qualities of a VR system that support an engaging experience, such as high-resolution visuals, broad fields of view, and accurate motion tracking. It revolves around how enveloping the virtual environment is from a sensory perspective. Presence, on the other hand, is a subjective psychological response to immersion. It describes the user's perception of being physically present in the virtual environment, which can occur even if the immersive qualities are not maximally enhanced. Presence is influenced by how interactively and consistently the user's actions are reflected in the VR environment, contributing to a sense of "being there" [522].

Here, attention allocation within VR environments is intricately linked to the levels of immersion and presence experienced by users. As VR commands the visual and auditory senses, it naturally directs users' attention to specific aspects of the virtual environment. This directed attention is critical for maintaining presence, as it guides the user's focus to interactive elements or tasks within the VR experience.

The impact of VR on attention and how it relates to immersion and presence is an area ripe for exploration. Research, including insights from [529], suggests that VR can impact attention location either internally or externally directed. Internal attention involves processing one's thoughts, memories, and internal stimuli, enabling individuals to focus on mental tasks or recall specific information. This aspect of attention is crucial in VR environments when users are required to manipulate objects or solve puzzles based solely on their cognitive capacity, without external cues. External attention, on the other hand, is directed towards stimuli in the external environment. In the context of VR, this can be triggered by the dynamic and interactive elements within the virtual environment, such as moving objects, changes in scenery, or specific audio cues designed to draw the user's focus. Both processes can significantly influence a user's immersive experience and sense of being present in the virtual world. Attention allocation in VR, whether captured involuntarily by vivid elements of the virtual environment or directed voluntarily towards goal-relevant tasks, underscores the dynamic interplay between user attention and VR environment design.

2.5.2 Augmented Virtuality

AV sits uniquely within the MR spectrum, crafting environments where virtual spaces are enhanced with elements from the real world. This approach distinguishes AV from its MR counterparts by enriching digital environments with tangible, real-life inputs, merging the immersive depth of VR with the grounding influence of actual objects and information visualization.

AV leverages advanced technologies to seamlessly integrate real-world elements—such as video feeds, sensor data, and live inputs—into virtual scenarios. This integration facilitates more engaging and interactive virtual experiences, as users can interact with digital representations that reflect their real-world actions or surroundings. For instance, AV applications might project a user's hands into a virtual space, allowing direct manipulation of virtual objects with real movements.

One challenge in AV development is achieving a naturalistic integration of real and virtual elements that maintain user immersion without introducing cognitive dissonance or discomfort. This requires not only technical prowess in terms of tracking and rendering but also a deep understanding of human perceptual and cognitive processes to ensure that the integration enhances rather than detracts from the intended experience.

While the direct literature on AV's impact on attention and task engagement is still developing [529], however emerging research into the impact of AV on attention and engagement is beginning to shed light on its potential benefits and challenges. Regenbrecht et al. [474] explored the utilization of AV for enhancing remote collaboration, presenting a system that simulates a co-located meeting environment by integrating live video streams of participants into a shared virtual space. This approach fosters a sense of presence among remote participants and facilitates a more natural and effective communication dynamic akin to face-to-face interactions. Wang et al. [583] investigate the educational applications of AV, focusing on how integrating real-world data and interactions within virtual learning environments can augment educational outcomes. By blending live environmental data or physical interactions with virtual simulations, AV offers a dynamic and interactive learning experience that can adapt in real time to the learner's actions and the surrounding context. This adaptability enhances the relevance and impact of educational content, potentially improving comprehension and retention by making abstract concepts more tangible and engaging.

Lastly, even tough research on AV is still in its infancy, principles derived from VR and AR research provide a foundation. For example, studies in VR have shown how immersive environments can significantly influence attention by enveloping the user's sensory inputs [269, 437, 610], suggesting that AV's blend of real-world elements could further enhance this effect by anchoring attention to familiar, tangible cues. Similarly, AR research highlights the potential for augmented information to aid memory retention and learning by providing contextual cues, suggesting that AV could leverage these benefits within fully immersive environments [54].

2.5.3 Augmented Reality

In AR, we provide information by overlaying digital content onto our physical surroundings, altering perception through devices like smartphones and AR glasses. While smartphones offer widespread accessibility to AR, specialized AR glasses, such as Microsoft Hololens and MagicLeap, target more niche, enterprise-level applications. The history of AR is rich, with roots extending back to Ivan Sutherland's development of the first head-mounted display in 1968, known as "The Sword of Damocles," which necessitated ceiling mounting due to its weight [541]. The term "Augmented Reality" was coined in 1992 by Tom Caudell and David Mizell, who developed a head-up display for aircraft manufacturing at Boeing [548].

Over the decades, advancements in mobile hardware and tracking solutions have significantly improved AR's fidelity, enabling immersive experiences like location-based games (e.g., Pokémon Go). AR technology has matured from simple wireframe overlays to complex, interactive digital enhancements firmly anchored in physical reality [504, 509, 543].

Ronald Azuma's definition of AR emphasized three core characteristics: the combination of real and virtual worlds, interaction in real-time, and 3D registration of virtual objects [24]. This distinction distinguishes AR from early "smart glasses" and underscores its immersive potential. Despite the technological strides, early AR devices struggled with issues like limited field of view and fidelity, often leading to diminished engagement and usability where the presence of the real world overshadowed the virtual.

AR technologies can now deliver information as a three-dimensional experience directly overlaid on real objects. This blend blurs the lines between the physical and digital realms, offering an "experience" rather than mere visual data, enhancing our ability to leverage the digital world in real-time, interactive manners [390]. Technological components of AR systems, including displays, tracking mechanisms, sensors, and the computer, have evolved to support a wide range of applications. Devices vary from head-mounted displays (HMDs) to handheld devices, each offering different advantages in terms of immersion, accessibility, and user interaction. HMDs, like the Microsoft Hololens 2, utilize optical see-through technology to integrate digital content naturally. At the same time, smartphones have popularized AR through accessibility and convenience, albeit with limitations in immersion and display size.

Apple Vision Pro has recently emerged as a pioneering spatial computing device, challenging these limitations by integrating digital content with the physical world more seamlessly. This advancement allows users to stay connected with their surroundings while engaging with virtual applications, offering a novel interaction that surpasses traditional screen-based interfaces. With Apple Vision Pro, the promise of AR—to blend digital enhancements naturally within our environment—takes a significant leap forward, offering enhanced productivity, entertainment, and educational possibilities.

Promising AR application domains range from gaming and entertainment to education, navigation, collaboration, and task support. These applications demonstrate AR's potential to transform various sectors by enhancing real-world tasks with digital information [53]. The introduction of Apple Vision Pro into this landscape opens new possibilities for AR, bridging the gap between digital data and real-world application with its advanced spatial computing capabilities, pointing towards a future where AR is seamlessly integrated into daily life, enriching our interaction with the physical world.

While AR devices' fidelity and immersive quality have significantly improved, challenges remain in optimizing user interfaces to leverage AR's full potential beyond mimicking traditional desktop scenarios. Integrating real and virtual elements opens vast opportunities for presenting additional information but also introduces challenges like divided attention and increased visual complexity, which can lead to mental fatigue or distraction. Studies have shown that AR's effectiveness hinges on the relevance of displayed content to the user's specific tasks [153, 176], underlining the importance of designing AR systems that enhance task performance without overwhelming users. To address such research gap, Vortmann et al. [571] developed and tested an AR task designed to elicit either internally or externally directed attention while recording EEG data, demonstrating the feasibility of using EEG-based machine learning methods to discern attention states in AR environments with high accuracy. Further, Vortmann et al. [573] have explored the integration of Brain-Computer Interfaces (BCIs) with AR to enhance user interaction by detecting attention focus, aiming to enable BCIs in AR without individual calibration . This approach addresses user-specific needs by dynamically adjusting AR content based on the user's attention state, potentially reducing cognitive load and improving task performance. This suggests that real-time detection of user attention states could significantly enhance AR application interactivity and user experience, making digital enhancements more intuitive and less intrusive. As AR evolves, focusing on unobtrusive designs and minimizing attentional load will be crucial in realizing its full potential to augment human perception and interaction with the digital world seamlessly.

2.5.4 Adaptive Mixed Reality

Adaptive MR emerges as a paradigm shift from conventional screen-based interactions, embedding interface elements directly into our physical environment, making them context-sensitive and inherently tied to the user's surroundings. This evolution presents unique challenges in interface management, particularly in adapting to dynamic usage contexts and ensuring the seamless integration of virtual content with the physical world.

Pioneering efforts in this domain have explored various strategies to manage the complexity of MR interfaces. Early works focused on enhancing visibility and minimizing occlusions through intelligent label placement and clustering techniques, as demonstrated by Azuma and Furmanski [23], and later, by employing machine learning for dynamic label visibility adjustments [200]. The quest for context-awareness led to more sophisticated adaptations, considering environmental geometry and user state to ensure interface elements harmonize with their physical backdrop, as Grasset et al. [218] and Gal et al. [195] explored. This body of research underscores the transition from static to dynamic interfaces, capable of retargeting MR layouts based on the semantic associations between virtual elements and the environment. However, the main limitations of such approaches is that most of the context is derived by image segmentation of users's surroundings and location without considering actually the ongoing user state [505].

Building on this foundation, recent studies focused on included ergonomics for environment-driven adaptations, leveraging physical affordances to guide the layout and interaction techniques within MR interfaces. Techniques have evolved to optimize virtual content placement through algorithms that respect ergonomic principles and spatial constancy, ensuring that interactions remain intuitive and reduce cognitive load [332, 359]. These advancements underscore a shift towards creating MR interfaces that not only adapt to physical spaces but also cater to the ergonomic needs and preferences of users.

The challenge of adaptive MR is thus defined by three core aspects: the user, the environment, and the activity. For the user, considerations range from physical abilities to cognitive load and design preferences. The environment encompasses both physical and virtual elements that support or hinder interaction. The activity defines the goals and tasks users aim to accomplish, shaping the interface's role in facilitating these objectives. Adaptive MR systems must navigate these dimensions, balancing the need for specific and general context understanding.

In envisioning the future of Adaptive Mixed Reality, we argue for systems capable of sensing and adapting to users' current contexts, adjusting interface elements like placement and information density. Such systems could revolutionize how we interact with digital information, making MR an integral part of our daily lives by ensuring that virtual augmentations are helpful and embedded in our real-world experiences. This holistic perspective on adaptive MR, grounded in technological innovation and a deep understanding of human factors, physiology, and cognitive processing, points toward ubiquitous MR environments where interactions are immersive, productive, and intuitive.

2.5.5 Adaptive Interactions and Visualizations

Although there have been earlier models for adaptive systems [234, 503], we consider three critical adaptation elements: content, presentation, and interaction. When adaptive systems adjust their content, which relates to users' preferences and engagement, they must consider the user's prior knowledge and interests. Such dimension might involve notification design and recommendations, especially considering the exploratory visual analytics process [490].

Secondly, presentation adaptation affects UIs or visualizations according to users' spare perceptual capacity, discomfort and, stress level by simplifying displayed information, luminance, or other properties.

Thirdly, interaction adaptation is a broader field as it might encompass different paradigms. For example, in multitasking environments, users might experience tasks being switched off [454], see the number of options change in a decision-making task [448], or modify the interaction modality, i.e., from gesture to hand-free interaction.

2.5.6 Use Cases and Application

Here, we provide a brief overview of adaptive visualization and interfaces with use cases and applications, specifically targeting content-based adaptation from physiological data, adaptation of visualization presentation from physiological data, and interaction adaptation.

Content Adaptation

In the following, we present and discuss systems based on eye-tracking features, that adapt to support language proficiency, increase recommender systems' performance based on inferred users' interest, or help visual analytics.

Adaptive Displays Based on Language Proficiency Globalization means that interfaces are prevalent in a multitude of different languages. Hard-to-access language correction can lead to user aversion. Consequently, there is merit in creating systems capable of estimating a user's language proficiency and displaying content appropriate to the user's abilities.

Recently, Karolus et al. [297] explored the potential of using a user's gaze properties to detect whether the information is presented in a language the user understands (see Figure 2.3). Robustness and feasibility with low-grade eye-tracking equipment were important aspects of this work. They proposed technical specifications for the recording equipment and the interaction period using robust gaze features, including fixation and blink duration. They found that a few seconds of recorded gaze data is sufficient to determine if a user can speak the displayed language.

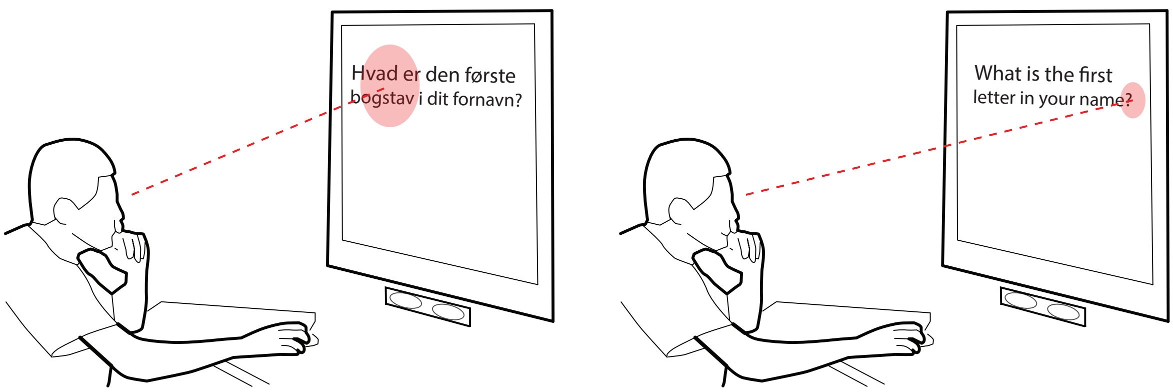


Figure 2.3: User interacting with a language-aware interface [297].

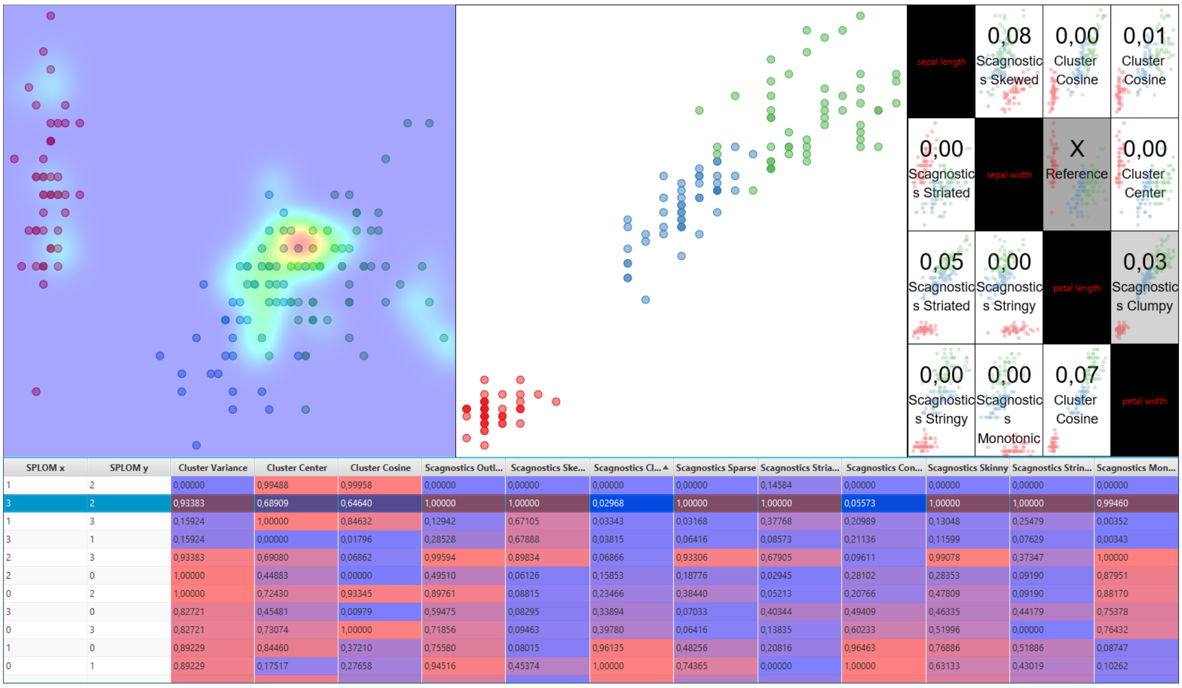


Figure 2.4: Screenshot of a recommender system with eye-tracking support [485]. Gaze is used as an indicator of interest and mapped to the underlying data.

Gaze as Input for Recommender Systems Silva et al. [519] sketched the possibility of back-propagating eye-gaze through the visualization pipeline and mapping it onto the underlying data. According to the eye-mind hypothesis, this viewed data is of interest to the user. Recommender systems, such as the one in Figure 2.4, can take advantage of such implicitly selected data to suggest helpful visualizations [485]. Recommendations based on such data fit the user's current interest and might, by extension, also fit their current task. However, a robust inference of an explicit task is not trivial, but a recommender system based on data interest can suggest the correct views for any generic and unidentified task.

Eye Tracking Support in Visual Analytics Systems Visual analytics is a design framework for interactive visual displays to facilitate the exploration of, and insight into, data sets. They rely on a loop that includes the viewer with all their prior knowledge, interests, and tasks. This allows the user to alter the selection of data, adjust parameters for data processing, and adapt the visualization on-the-fly to cater to current needs.

With the added information from eye-trackers, such visual analytics systems can augment existing interaction techniques [519]. This can include, for instance, gaze as additional cursors for interaction through speech or disambiguation of targets when pushing buttons on hand-held controllers. In addition, with the advent of coarse eye-tracking for devices with front-mounted cameras (e.g., tablets and phones), existing visual analytics software can 'retro-fit' gaze data without changing the actual hardware. For example, law enforcement agents

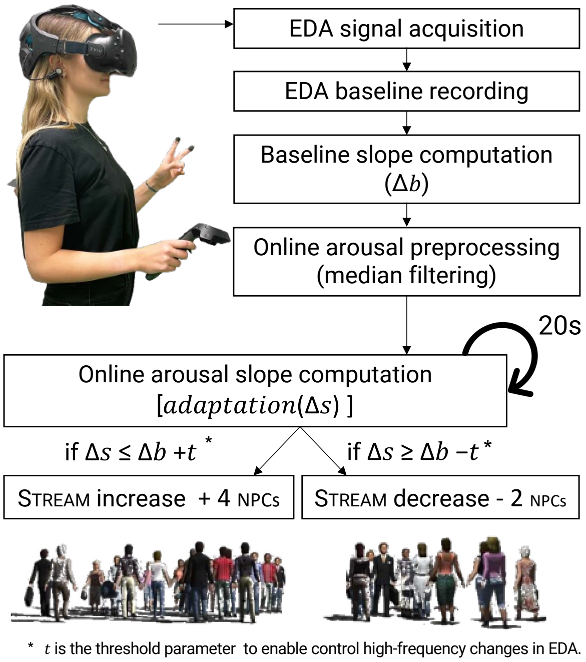


Figure 2.5: Flowchart of the physiological loop used by Chiossi et al. [110]. The visual complexity (in the form of NPCs) adapts according to changes in the EDA calibrated from a baseline recording (Δb). The adaptation function is called every 20 secs.

already use software on car-mounted tablets to provide them with overviews of occurrences in their districts. In such a scenario, even coarse gaze data can check whether relevant events have been overlooked and provide adaptive visualizations to attract the agent's attention.

Presentation Adaptation

In this section, we highlight the work of adaptive systems that adapts presentation based on users' physiological input, such as EDA, to support user experience, or to support processing of relevant information, i.e., notifications.

Adaptation of Virtual Reality Visual Complexity Based on Physiological Arousal VR is rapidly gaining popularity for social or collaborative virtual environment applications. Such settings envision the involvement of realistic Non-Player Characters (Not-Player Charachter (NPC)), such as virtual crowds with human-like behavior. However, highly dynamic environments could provide task-irrelevant elements that negatively increase a user's cognitive load and distractibility. Thus, monitoring users' physiological activity and adapting the interaction is an emerging research trend to optimize user experience or performance.

The goal of physiological control loops is to detect deviations from the optimal physiological state that influence the adaptation of the features of the environment or tasks to drive users towards a more desirable state. Here, Chiossi et al. [113] focused on a peripheral measure of physiological arousal, i.e., EDA. Physiological arousal correlates with task demands and engagement in a multi-component task [179] and can be affected by proxemics of NPCs both in VR [353] and AR [260]. Hence, the stream of NPCs was adapted in response to changing EDA levels of users while being engaged in a dual-task setting. They processed the EDA data only using an average moving window of 20 sec. For user-dependent adaptation, the adaptive algorithm adjusts the visual complexity to a baseline slope recording recorded at the beginning of the experiment. Thus, when the EDA slope was larger than the baseline slope, 2 NPCs were removed, indicating increased arousal. On the contrary, 4 NPCs are added to the environment if the system detects decreased arousal. Figure 2.5 visualizes the adaptation algorithm. Thus, they supported the user experience by leveraging visual complexity, i.e., increased system acceptance, competence, and immersion.

On the surface, the findings of Chiossi et al. [110] only impact the design of social VR scenarios. However, such results can generalize to the design of information visualization patterns to help users in mixed reality (MR) [520]. Thus, physiologically-aware systems can potentially personalize environments for the user, improving the user experience. Furthermore, in the visualization domain, VR complexity can burden users; thus, a physiological-adaptive system varying the information load according to users' workload could foster the viability and usefulness of applications.

Adapting Notifications to Visual Appearance and Human Perception Users benefit from desktop notifications showing them their incoming messages, upcoming calendar events, or other important information. Notifications need to attract and divert attention from a primary task effectively to ensure that users notice important information. At the same time, notifications are embedded into the visual design of the user interface and are subject to aesthetic considerations. However, design decisions that are also currently static, i.e., do not adapt at runtime, can severely impair the user's ability to perceive notifications.

Müller et al. [401] presented a software tool to automatically synthesize realistically looking desktop images for major operating systems and applications. These images allowed them to systematically study the noticeability of notifications during a realistic interaction task. They found that the visual importance of the background at the notification location significantly impacts whether users detect notifications. Their work also introduced the idea of noticeability maps: 2D maps encoding the predicted noticeability across the desktop. The maps inform designers how to trade-of notification design and noticeability. In the future, such automatically predicted noticeability maps could be used in UI design and during runtime to adapt the appearance and placement of desktop notifications to the predicted user noticeability.

Interaction Adaptation

In the last section, we present relevant work that shows how adaptive systems can support interaction for mid-air or multimodal interactions in immersive MR environments.

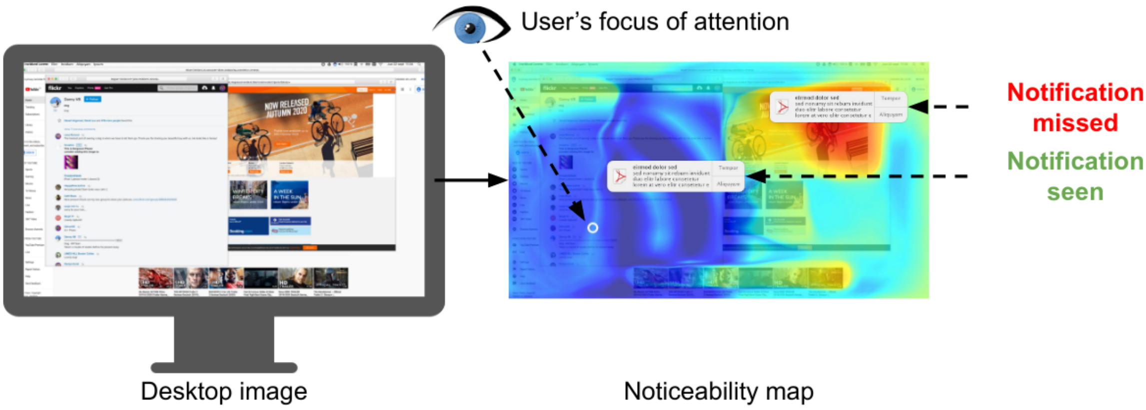


Figure 2.6: Müller et al. [401] collected perceivability and behavioral data on realistically looking synthesized desktop images. They used this data to identify the factors that impact the noticeability of notifications. This allowed them to develop a computational model of noticeability that can predict noticeability maps for a given desktop image and user attention focus. These maps visualize the locations at which a notification is likely to be missed (red) or likely to be seen (green).

Adapting the 3D User Interfaces for Improved Ergonomics Interactive MR applications surround the user with virtual content that can be manipulated directly by reaching for it with the tracked hand or controllers. Such mid-air interaction techniques are beneficial, as they feel natural, but they may lead to physical strain, muscle fatigue, and challenging postures [59, 377]. The XRgonomics toolkit [174] addresses these issues by visualizing the ergonomics of the user's interaction space (see Figure 2.7), allowing UI designers to create interfaces that are convenient and easy to manipulate. Further, it supports the automatic adaptation of UIs so that interactive elements remain within easy reach while the user moves about in a changing physical environment. The ergonomics metrics currently supported in XRgonomics are RULA [378], Consumed Endurance [250], and muscle activation [28].

Prior research has explored ergonomics [28, 250, 378] and while the resulting metrics help evaluate existing UIs, it is difficult to use them for generating novel UI layouts. Further, the formulated design recommendations can be challenging to interpret and apply, particularly if the ideal interaction space is unavailable, e.g., due to the user's physical environment.

To address this, Belo et al. [174] present a toolkit to visualize the interaction cost in the user's entire interaction space by computing ergonomics metrics for each reachable point in space. Their work shows a half-sphere of voxels around the user, color-coded to reveal the ergonomics of reaching for that position. Thus, the toolkit allows UI designers to inspect the interaction space and identify ideal placements for various interactive elements. The toolkit further allows the definition of constraints, e.g., allowing the designer to define areas of the interaction space that are not available for placement of interactive virtual content, for example, due to physical obstacles in the user's environment. Based on that, the toolkit can recommend the ideal position with the best ergonomic properties for reaching with the hand. As this computation is feasible in real-time, the toolkit API can be used for dynamic adaptation of UIs, depending on the user's changing physical environment or varying visible

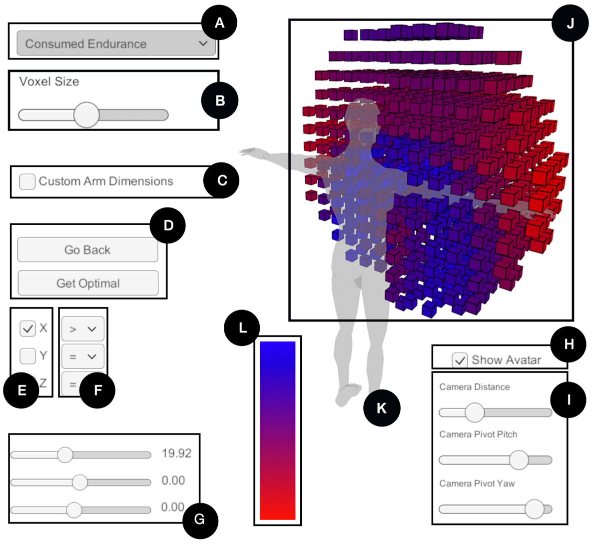


Figure 2.7: The XRgonomics toolkit [174] visualizes the cost of interaction for each reachable point in the user's interaction space, through color coding (K) from blue (most comfortable) to red (least comfortable) (L). The applied metric is selected in a dropdown menu (A), and the computed value can be adapted for the user's arm dimensions (C). For a better visibility, the voxel size can be adapted (B), and the range of values to visualize can be limited along all three axes (E-G) to show only individual regions or slices of the space. Further, the user can retrieve the "optimal" voxel with the lowest ergonomic cost (D). Finally, the visualization of the avatar can be deactivated (H), and three sliders enable control of the perspective (I).

space. For example, consider a UI element that should always remain within the user's field of view in an AR scenario. The user is wearing a HMD, and as they turn their head to look around, using the view frustum of the HMD, constraints arise in the available interaction space. The toolkit automatically computes the most ergonomic placement for the respective UI element within this available volume, keeping it in easy reach for the user.

Beyond improving the ergonomics of mid-air interaction, this approach may be applied to achieve the opposite goal of increasing physical effort to reach a UI element or virtual object, e.g., with the aim to train particular muscles. This may contribute to rehabilitation or be applied in exergame scenarios, as proposed by Munoz et al [404].

Hybrid User Interfaces for Augmented Reality The complexity of interaction in AR environments provides many opportunities for adaptation, such as adapting visualizations based on the user's physical surroundings (e.g., Shin et al. [516]), within situated analytics (e.g., Fleck et al. [190]), or by considering the devices available in the user's workspace (e.g., STREAM [265]). One possibility of adapting visualizations and interfaces to the user can be realized through hybrid user interfaces that combine the advantages of heterogeneous devices (e.g., head-mounted AR devices and handheld tablets), creating the ability to facilitate multiple coordinated views across different realities for visual analytics.

For example, STREAM [265] combines an immersive AR environment using an AR headset

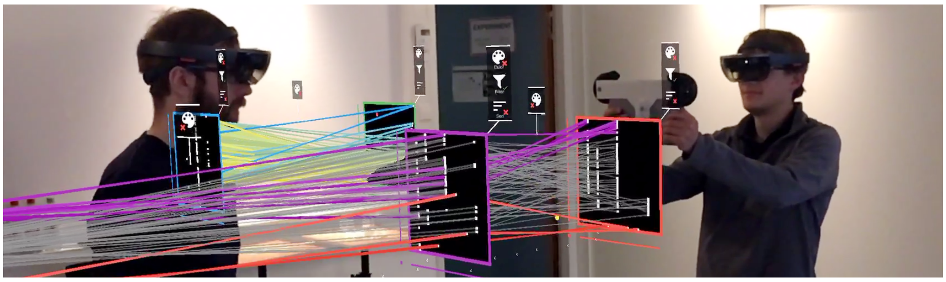


Figure 2.8: STREAM [265] combines spatially-aware tablets with head-mounted AR displays for visual data analysis using a 3D parallel coordinates visualization. STREAM's adaptation mechanisms allow users to seamlessly switch between the AR visualization and the tablet visualization without losing context. For example, the user on the right holds their tablet vertically, allowing STREAM to adapt their AR scatter plot with the tablet's visual space. In contrast, the collaborator's (left user) AR visualization is unaffected.

with a spatially-aware tablet for interacting with 3D parallel coordinates visualizations, consisting of linked 2D scatter plots. Here, the AR headset allows users to see the visualization in stereoscopic 3D space. At the same time, the tablet provides familiar touch interaction on individual 2D components of the visualization, e.g., 2D scatter plots, see Figure 4.30. Furthermore, to reduce the cognitive demand when switching between both interfaces, STREAM automatically adapts the representation of both interfaces to the user's implicit interaction by tracking the tablet's position in space: Once a user holds their tablet in front of them (i.e., indicating that the user wants to switch between devices), the selected 2D component of the visualization in AR (e.g., a 2D scatter plot) rotates toward the user's viewing direction, while the tablet adapts its content to show the same 2D component on screen—effectively merging both visual spaces into one interaction space. This adaptation allows users to seamlessly switch between AR and tablet visualization without losing context.

2.6 Summary

While prior research has extensively explored attention and engagement mechanisms in traditional and MR environments, several gaps remain in understanding how these factors interact dynamically with workload and usability considerations in adaptive MR systems. Previous models of attention allocation and workload often focus on isolated contexts, such as desktop-based HCI interactions or controlled VR environments, without fully addressing the challenges introduced by blending virtual and physical elements. This thesis contributes to filling this gap by systematically investigating electrophysiological correlates of attention allocation and workload in MR through multimodal quantitative measures, providing design implications for engagement- and attention-aware systems (**RQ1**).

Specifically, existing work lacks an integrated perspective on how users identify and prioritize blended information across the MR continuum, particularly when balancing the processing of virtual and physical objects. While selective attention models and perceptual load theories provide a foundation, their application in MR contexts remains underexplored, particularly regarding visual search complexity in mixed environments (**RQ2**). Additionally, prior work does not sufficiently address how these attention mechanisms generalize to more ecologically valid tasks, such as typing, where engagement and workload dynamics differ across MR actualities (**RQ3**).

Building upon these foundational studies, this thesis extends prior work by examining how adaptive systems can leverage physiological data to modulate secondary task difficulty or visual complexity in VR environments, thereby supporting primary task performance (**RQ4** and **RQ6**). Existing adaptive MR research primarily focuses on explicit user input and behavioral markers, often neglecting the potential of physiological signals as implicit indicators of user state. Our approach addresses this by employing a multimodal framework to assess the effects of adaptations on physiological channels (**RQ5** and **RQ7**), providing a more nuanced understanding of sensor-driven adaptation.

Furthermore, the practical implementation of physiological sensing in MR hardware remains an open challenge. While previous studies demonstrate the feasibility of integrating physiological data for interaction, technical constraints such as sensor placement, signal quality, and real-time processing hinder broader adoption. This thesis tackles these challenges by designing and validating an MR-embedded physiological sensing system that enhances both usability and data interpretability, directly contributing to the development of intelligent MR interfaces (**RQ9**).

By addressing these research gaps, this work advances the understanding of physiological computing in MR, paving the way for adaptive systems that respond to users' attentional and engagement states in real time. The findings contribute to both theoretical models of MR interaction and practical applications in designing user-centered, physiologically adaptive environments.

3

ATTENTION AND ENGAGEMENT IN MIXED REALITY

“Gentlemen you had my curiosity ... but now you have my attention”

– Quentin Tarantino
***Django Unchained*. 2012.**

In this chapter, we investigate the impact of different degrees of virtuality in the MR continuum when searching for relevant information among distractors (**RQ1** & **RQ2**) or in ecological productive tasks (**RQ3**). First, we study how users identify relevant information and suppress distracting information across different manifestations of the MR continuum under varying levels of task difficulty. Second, we investigated users’ visual search performance when identifying physical and virtual target information in AR and AV. Here, we aim to understand if virtual and physical information poses different perceptual demands on users and if their virtuality-physicality impacts users’s capacity to ignore distracting information. Our findings indicate that in AR environments, users experience increased cognitive load and scatter visual search patterns. Conversely, AV environments, where virtual elements dominate, facilitate more efficient target processing, demonstrated by distinctive ERP components and eye movement patterns. Further, we investigated the effect of executive a productivity task, i.e., typing, in different MR manifestations, on physiological correlates of attention and engagement (**RQ3**). We found that AV may strike the best balance among the different MR manifestations, shielding users from external distractions while displaying task-relevant real-world content to support typing performance and engagement.

This chapter is based on the following publications.

Francesco Chiossi, Uwe Gruenefeld, Baosheng James Hou, Joshua Newn, Changkun Ou, Rulu Liao, Robin Welsch, and Sven Mayer. 2024. Understanding the Impact of the Reality-Virtuality Continuum on Visual Search Using Fixation-Related Potentials and Eye Tracking Features. In *Proceedings of the ACM on Human-Computer Interaction*, 9(MHCI), 1-32.
https://doi.org/10.1145/3676528

Francesco Chiossi, Ines M. Trauttmansheimer, Uwe Gruenefeld, and Sven Mayer. 2024. Searching Across Realities: Investigating ERPs and Eye-Tracking Correlates of Visual Search in Mixed Reality. In *IEEE Transactions on Visualization and Computer Graphics*, 30(1), 1-14. IEEE, New York, NY, USA.
https://doi.org/10.1109/TVCG.2024.3456172

Francesco Chiossi, Yassmine El Khaoudi, Changkun Ou, Ludwig Sidenmark, Abdelrahman Zaky, Tiare Feuchtner, and Sven Mayer. 2024. Evaluating Typing Performance in Different Mixed Reality Manifestations Using Physiological Features. In *Proceedings of the ACM on Human-Computer Interaction*, 7(ISS), 1-25. ACM, New York, NY, USA.
https://doi.org/10.1145/3698142

3.1 Study 1: Visual Search Across the Mixed Reality Continuum

Our study investigates how users identify relevant information and suppress distracting ones across different manifestations of the MR continuum under varying levels of task difficulty. To achieve this, we choose a visual search task, adaptable to diverse MR settings such as cognitive training [145], information retrieval [55], and daily tasks like finding items while cooking [131]. Despite the established research on visual search in controlled settings [163] with known physiological correlates [254, 290], there is a lack of systematic evaluation of how visual search behavior and its physiological indicators translate across MR manifestations. Thus, to bridge this gap, we designed a study where participants perform a visual search task (as modeled after [145]) across AR, AV, and VR scenarios using the validated MR toolkit VRception [224]. To ensure the external validity of our results, we are not only focusing on the continuum but also incorporating task difficulty in the form of perceptual load as a validation control variable.

Drawing from previous work from visual search, we formulate the following research questions:

RQ1: Do different MR manifestations impact performance and perceived workload differently?

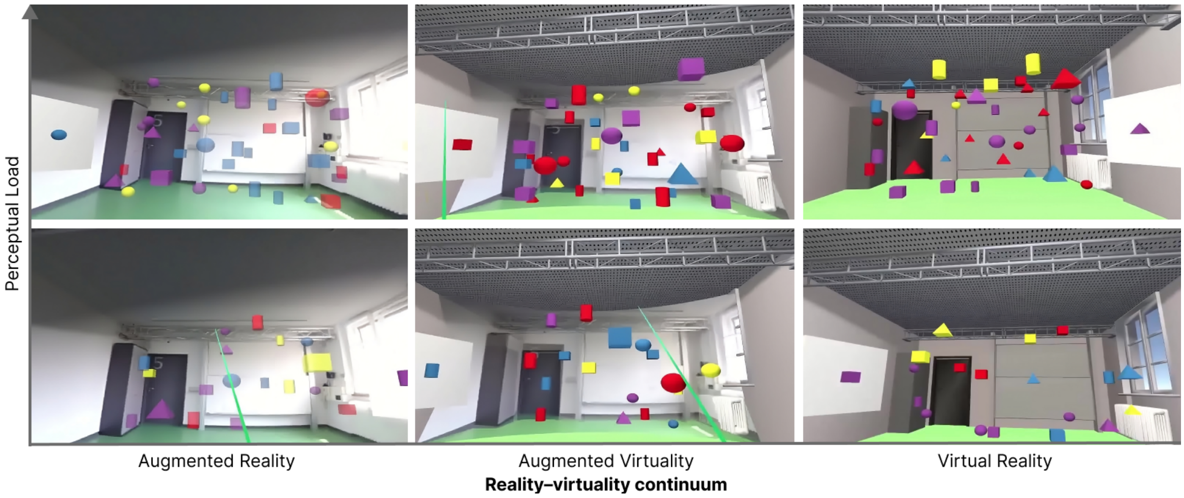


Figure 3.1: *Experiment Conditions.* In our study, participants performed a visual search task across the MR continuum (AR, AV, and VR) with two possible levels of perceptual load (low and high). In the first row, we show the VR condition for the visual search task under high PERCEPTUALLOAD across the virtuality continuum. In the second row, we show the low PERCEPTUALLOAD condition in the different levels of the reality-virtuality continuum.

RQ2: How do cognitive resource allocation in a visual search task vary across the continuum, as indexed by Fixation-Related P3?

RQ3: Does the MR continuum impact eye tracking correlates of visual search efficiency (fixations and saccades), and workload (pupil size)?

3.1.1 Study Design

We used a within-participants experimental design. The independent variables were MANIFESTATION (three levels: AR/AV/VR) and PERCEPTUALLOAD (two levels: Low / High). To avoid learning effects, we counterbalanced the order of conditions in a balanced Latin Williams square design with six levels [582]. Therefore, independent variables were manipulated using a 3×2 experimental design.

Mixed Reality Manifestation Conditions

Here, we describe the implementation of the tree levels of our MANIFESTATION independent variable. For a graphical depiction, we refer the reader to Figure 3.1.

AR condition. In the *AR* conditions, participants execute a visual search task in a real-life 360-video scenario, like if they are asked to look for virtual objects on a shelf. Here, virtual objects have typical AR features. Thus, they are see-through, while the surrounding physical environment maintains the color of the pre-recorded video and transparency. To optimize the visibility of virtual objects against real-world backgrounds, we employed a transparency level of T0.5 (alpha = 0.5) for stimuli in the AR condition, following [268]'s recommendations. We chose to employ a pre-recorded video of the empty 'real' room in the background to control for background luminance as variations in external environment luminance can lead to changes in color perception in AR [251, 617]. Here, the background was maintained fully opaque mimicking a real environment (alpha = 1.0).

AV condition. In the *AV* condition, participants perform the visual search task in an AV environment, where there are shared features elements from both AR and VR scenarios, i.e., the floor is virtual, and parts of the closet are virtual and physical, see Figure 3.1. Virtual objects exhibit VR features, exhibiting an opaque appearance. On the other hand, the physical objects within the environment retain the color of the pre-recorded video blended with the VR scene. Our approach to AV aims to be primarily virtual; however, parts of the physical reality are integrated into the experience. We chose a ratio of real and virtual elements that mirrors the approach in AR, where this ratio is flipped (c.f., AR and AV on the reality-virtuality continuum [390]). Consistent with prior studies on AV like [380] and [486], we incorporated aspects of the actual room into our virtual experience. We aimed to ensure a fair comparison, in which the ratio of real to virtual mirrors AR without focusing too much on specific objects, as depicted in Figure 3.1. Here, we used a transparency level of .5 for the background and 1.0 for the visual search stimuli.

VR condition. In the *VR* condition, participants carry out the visual search task in an immersive 3D rendering of the room used for the AR condition. The virtual objects and background exhibit VR features in this environment, creating a fully opaque appearance by applying a transparency level of 1.0.

Perceptual Load Conditions

Our PerceptualLoad manipulation consisted of 25 (Low) vs 40 (High) search display set, based on previous work that showed how set size yields inefficient performances between 30 and 60 items [248, 267, 425, 515, 556]. Among distractors, we presented a third of objects with the same color, a third with the same shape as the target, and a third that did not share either color or shape features, i.e., 8 in the low PerceptualLoad and 13 in the high PerceptualLoad conditions, respectively.

Dependent Variables

We collected a set of multimodal variables: (i) FRPs (P3b), (ii) eye tracking features, (iii) behavioral accuracy, (iv) Reaction Times (RTs) , and, (v) subjective workload (NASA TLX). We additionally collected three 7-point Likert-scale responses with an unipolar rating scheme [60, 348] (1 : Strongly Disagree, 7 : Strongly Agree) on MR usability, i.e. on how the background was *Distracting* ("The background distracted me strongly from the visual search task"), how

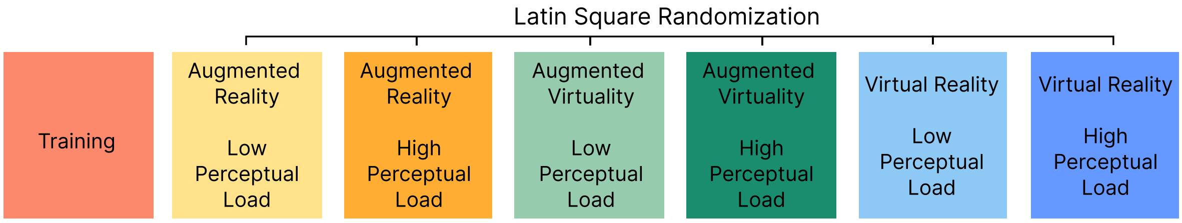


Figure 3.2: *Experiment Procedure.* The experiment encompassed seven different blocks. In the first block, participants performed a training session until they reached an accuracy level of 80 % in the visual search task. Finally, the experimental blocks started, manipulating MANIFESTATION and PERCEPTUALLOAD using a Latin Square randomization. Refer to section 3.1.1 for a complete description of the experimental conditions.

the virtual objects were *Overwhelming* (“I perceived the number of virtual objects as very overwhelming": Overwhelming) and, finally how the background made the task more or less *Difficult* ("The background made it very difficult to distinguish the target object from other objects": Difficult).

Procedure

Upon the participants’ arrival, the experimenter provided them with details about the study’s process and obtained written informed consent. Next, the experimenter set up the water-based EEG recording. Then, participants wore the HTC Vive Pro Eye headset and completed a five-point eye-tracking calibration. The study began with a training phase, where participants familiarized themselves with the Visual Search task in a Training block set in a neutral VR environment, i.e., Unity Default Skybox. Participants performed 20 Visual Search trials with a low PERCEPTUALLOAD search display. To start the experimental phase, they should have reached at least 80% accuracy. Otherwise, they repeated the training block. Finally, the experimental phase started with the first Visual Search block, following a Latin Williams square design [582]. Each Visual Search block followed the trial structure depicted in Figure 3.3 and encompassed 50 trials, as depicted in Figure 3.2. After each block, participants filled in the NASA TLX questionnaire[237] and ad-hoc UX survey. NASA-TLX and UX surveys were administered via the VR Questionnaires Toolkit [181]. On average, the experiment lasted about one hour.

Task

Participants performed a visual search task across three different MR manifestations. Participants were asked to select the target item with the VIVE controller trigger button among various distractors. Participants were instructed on the target features by presenting an image of the target object in the middle of the user’s view. The sides (left-right) where the target object image was presented were randomized across trials to avoid habituation effects in the eye tracking patterns. To select an object, participants were instructed to visually explore the MR environment and use the controller’s ray cast to aim at the target object. Once the target object was identified and aligned with the raycast, participants pressed the Vive controller’s

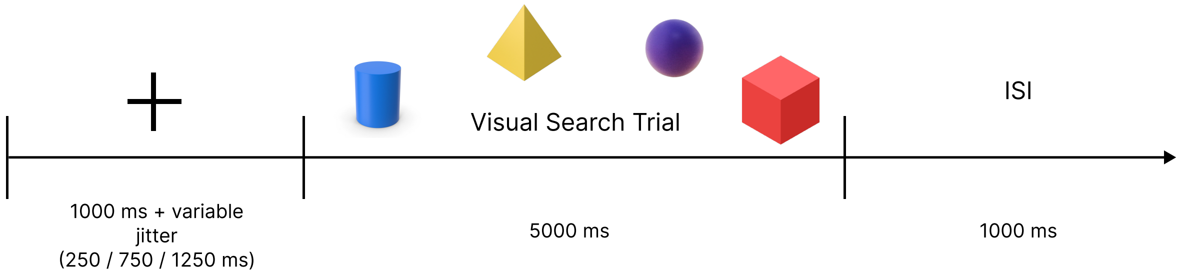


Figure 3.3: *Visual Search Trial Structure.* The visual search trial was structured into three phases: Initially, participants were shown a fixation cross for a baseline duration of 1000 ms, followed by an additional, randomly assigned jitter duration of either 250 ms, 750 ms, or 1250 ms. This meant the total fixation cross-presentation varied between 1250 ms, 1750, and 2250 ms in each trial depending on variable jitter duration with no objects in the background. Subsequently, participants had 5000 ms to identify the target among distractors, followed by a 1000 ms interstimulus interval (ISI). Each condition involved 50 trials per participant.

trigger button. Participants hold only one controller using their dominant hand. They were required to respond to the stimuli as fast and accurately as possible. Participants performed 50 trials per experimental block.

Trial Structure

Our trial structure was inspired by the work of [191], following a real-world visual search task approach [93, 602]. The structure of the task, as depicted in Figure 3.3, was as follows: (1) participants were asked to fixate the fixation cross (+) with a pseudorandom duration (1250, 1500, or 1750 ms) at the center of the camera rig; (2) participants performed the visual search trial; (3) after selection, an Inter-Stimulus Interval (ISI) of 1 second was presented with no cross or objects presented to allow for neural and attentional reset and counteract fatigue effects [20, 606]. Participants had 5000 ms after visual search display onset to select the target among distractors.

Stimuli The stimulus set consisted of five types of objects: a fixation cross (red) and four possible target/distractor objects (cube, cylinder, pyramid, sphere). Colors were defined by the following RGB values: red = (191, 24, 24); blue = (255, 255, 255); purple = (29, 119, 47), yellow = (255, 255, 0), following recommendations from previous work to avoid the influence of complementary color and environment background on the search task [72, 135]. As luminance can significantly impact pupil size and act as a confounding variable [85], we controlled for luminance in the Unity VR Editor, setting brightness at 0.7. We displayed the stimuli using a video see-through technique. Here, we used a pre-recorded video to provide the same real-world experience for all participants in the AR and AV conditions displayed via the same device. Moreover, to validate our approach, we measured luminance across conditions. We measured the light levels inside the headset using a lux meter sensor (LT300, Extech, USA). Using 20 measurements per condition, we found an AR in the average lux was 44.4 (SD=.9), 43.3 (SD=1.0) lux in AV, and 44.1 (SD=$.9$) lux in VR. Those values align with luminance guidelines (below 200 nits) based on eye-tracking best practices to avoid confounders for pupil size computation [85, 375]. Participants were placed at a 1.5-meter distance from the fixation cross, resulting $5.72 \times 5.72 \times 5.72$ degrees visual angle in size. Regardless of the perceptual load condition (High vs. Low), we displayed objects within the participant's frustum, maintaining eccentricity stable ($8°$), allowing for greater ERPs component sensitivity and statistical power [427]. By displaying objects within the frustum, they are positioned within the participant's field of view, making them easily accessible and visible without requiring extensive head or eye movements, minimizing EEG artifacts. Additionally, presenting objects within the frustum aligns with users' natural viewing behavior, enhancing the study's ecological validity and ensuring a more realistic representation of visual search behaviors in everyday MR environments.

Apparatus and Data Recording

The MR Visual Search task was implemented in Unity (Version 2020.3.8 LTS). We presented the MR manifestations using an HTC VIVE Eye Pro headset with a display resolution of 2880×1600 pixels combined (Field of View: 110 °). For environment tracking, we use two HTC Vive lighthouses 2.0. We acquired two physiological measurements: EEG signal using a LiveAmp amplifier connected via Bluetooth (BrainProducts GmbH, Germany, 500 Hz) and eye tracking data via HTC Vive Pro Eye headset (120 Hz). Physiological data were streamed within the Unity VR environment within the Lab Streaming Layer (Lab Streaming Layer (LSL)) framework[1] to the acquisition PC (Windows 10, Intel Core i7-11700K, 3.60 GHz, 16GB RAM). For recording the real-life 360-video used in the AR and AV conditions, we used the Insta360 Pro with a resolution of 3840x1920 pixels (4K) with 120fps.

EEG Recording & Preprocessing We acquired EEG data (sampling rate = 500 Hz) via LiveAmp amplifier (Brain Products, Germany) from 32 water-based electrodes from the R-Net elastic cap (Fp1, F3, F7, F9, FC5, FC1, C3, T7, CP5, CP1, Pz, P3, P7, P9, O1, Oz, O2, P10, P8, P4, CP2, CP6, T8, C4, FC2, FC6, F10, F8, F4, Fp2, Fz). We kept impedance levels below $\leq$20 kΩ. We set the reference at FCz during the recording, while FPz was used as ground. Electrodes were placed using the International 10-20 layout. For time synchronization with the MR environment, we employed the LSL Framework, while for preprocessing and analysis, we used the MNE-Python Toolbox [216]. We first automatically detected bad or outliers channels via random sample consensus (RANdom SAmple Consensus (RANSAC)) method [52] of spherical splines for estimating scalp potential based on algorithms proposed by [436]. We then applied a notch filter (50 Hz) and band-passed the signal between (1-30 Hz) to remove high and low-frequency noise. We then re-referenced to the common average reference (common average reference (CAR)). We applied an Independent Component Analysis (Independent Component Analysis (ICA)) for artifact detection and correction with extended Infomax algorithm [342]. We automatized the labeling and rejection process of ICA components via

[1] `https://github.com/labstreaminglayer/`

the MNE plugin “ICLabel” [280, 443]. Epochs that showed blinks, eye movement, muscle, or single-channel artifacts in any of the electrodes were rejected. Only trials with a correct response and last fixation on the target were used; error trials or distractor fixations were excluded from analyses.

Fixation-Related Potential Analysis

For FRP analyses, we segmented continuous signals between 250 ms before and 500 ms after the onset of each fixation. For the trials considered, no additional saccades were present during this period. We chose this epoching based on previous FRP analyses and eye-tracking research showing that EEG activity can be measured before saccade onset [290] but also posterior to the fixation onset [77]. Since the artifact of the previous saccade was restricted from 100 ms to 0 ms before the fixation onset, the baseline for each channel was defined between 250 ms and 100 ms before the onset of the current fixation and subtracted. For FRP computation, we chose O1, O2, and Oz electrodes based on previous work [406].

Eye Tracking Recording & Preprocessing We collected the head position and directional 3D vectors from Unity. We recorded eye tracking (Eye Tracking (ET)) data using the built-in eye tracker in the HTC Vive Pro Eye headset (120 Hz) through the SRAnipal eye tracking SDK, giving us eye-directional 3D vectors relative to both the world and the head. To facilitate our analysis, we initially converted the head directional 3D vector, eye-in-world directional 3D vector, and eye-in-head directional 3D vector into 2D Fick angles using the Fick-gimbal method [239]. This conversion involved two rotations, one about the vertical axis and the other about the nested horizontal axis, effectively characterizing the position of each vector. These resulting 2D Fick angles for eye and head directions were the basis for our subsequent analyses. For analysis, we focused on the eye-tracking data for each trial until the participant selected the target. Unlike EEG, eye tracking is no longer relevant after the target selection. Given the relatively short duration of our trials, averaging approximately 2.71 ± 1.01 seconds across all trials, we conducted our eye tracking analysis at the trial level rather than examining individual behaviors within each trial.

Fixation-Saccade Analysis

We calculated fixation and saccades using pymovements [323], an open-sourced Python package for analyzing eye-tracking data. We chose pymovements’ implementation of the ID-T algorithm [493] with specific fixation thresholds set at a minimum fixation duration of 83ms and a maximum dispersion of 1.8 degrees [39, 562]. This allowed us to calculate the following fixation-based metrics: total fixation duration, average fixation duration, fixation count, and the time from the start of the visual search trial to the last fixation within the trial. For saccadic analysis, we employed pymovements’ implementation of the microsaccade algorithm [168]. This enabled us to calculate two key saccade metrics: saccade amplitude and saccade frequency. Saccade amplitude was computed as the angular distance in degrees between the saccade onset and offset, while saccade frequency was determined by dividing the number of saccades within a trial by the trial duration.

Pupil Diameter Analysis To compute the average pupil size for each trial [307, 423], we initiated the process by removing the pupil baseline for each eye. This baseline was established as the mean pupil size observed during the phase of the trial when the fixation cross was displayed. Subsequently, we computed the mean of the normalized left and right pupil sizes and then calculated the mean and standard deviation of the combined pupil size to facilitate our analysis at the trial level.

Behavioral Data Analysis

Accuracy Given its dichotomic nature, response accuracy at each trial was analyzed by conducting a Generalized Linear Mixed Model (Generalized Linear Mixed Model (GLMM)) with logit link function using the glmer function from the lme4 library [36]. Behavioral responses and conditions were entered into the model as fixed effects while a random intercept varying among participants was entered into the model as random effects, leading to the formula: `accuracy ∼ Manifestation * PerceptualLoad + (1|participant)`.

Reaction Times Raw RTs were analyzed using a GLMM using lme4 library [36]. Error trials were excluded from the analysis (formula: `RTs ∼ Manifestation * PerceptualLoad + (1|participant)`). Overall, we computed 276.71 average amount of trials (SD = 12.84) per participant. We chose this approach because, besides considering variability across individuals, it allows us to control for many longitudinal effects during the task without transforming the data [354]. First, there are the effects of learning and fatigue [25]. Second, the response in a trial is usually heavily influenced by what happens in the preceding trial [327, 369, 602]. Sources of experimental noise are brought under statistical control using mixed-effects models.

Subjective Measures We analyzed subjective workload reports via the NASA-TLX questionnaire [238] via either two-way Analysis of Variance (Analysis of Variance (ANOVA)) or Aligned Rank Transform (Aligned Rank Transform (ART)) ANOVAs [599] depending on normality testing. We additionally collected three 7-point Likert-scale responses on UX, related to how *Distracting* was the background, how the number of objects was *Overwhelming*, and lastly, how much the background impacted the task difficulty, i.e., *Difficult*.

Participants

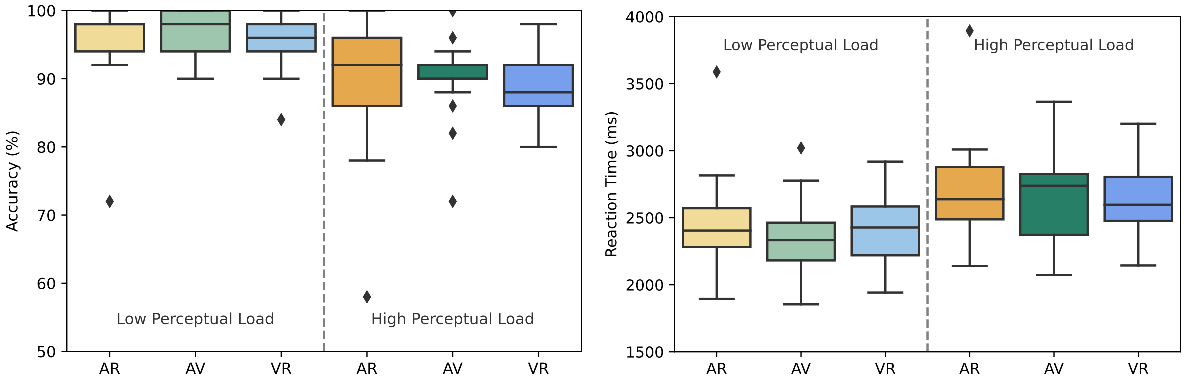
A sample of 24 participants voluntarily participated in the study recruited via institutional mailing lists, social networks, and convenience sampling. The number of participants recruited for our study aligns with recent investigations on the relationship between participants number and EEG data reliability for relatively long tasks [575] and in line with previous work in HCI and visual search domain [145, 191]. Three participants were excluded due to inadequate EEG data quality, as identified by RANSAC, which revealed that over 50% of electrodes for computing FRPs (01, 02, Oz) were classified as "bad," compromising data reliability due to a low SNR as described in [52]. This led to a total of 21 participants ($M = 25.57$, $SD = 2.64$; 8 female, 13 male, none diverse). We surveyed the familiarity of participants with AR, AV, and VR devices as in previous work [96]. All participants reported prior experience with AR ($M = 3.76$, $SD = 1.6$), with AV ($M = 2.45$, $SD = 1.07$) and VR ($M = 4.5$, $SD = 1.07$) on a scale from 1 (not at all familiar) to 7 (extremely familiar). None of the participants reported a history of neurological, psychological, or psychiatric symptoms. The study met the criteria for fast-track conditions set by the local institutional ethics board, i.e., participants were not exposed to any risks such as deception, excessive stress, or recording of sensitive information.

3.1.2 Results

In this section, we first present results on behavioral accuracy and reaction times, FRPs analysis on average peak amplitude for P3, eye tracking features, and subjective scores on perceived workload (NASA-TLX) and ad-hoc UX surveys. We employ a GLMM to investigate differences in the reaction time, FRPs, and ET feature distributions. We report the effect sizes as odds ratios for behavioral accuracy as a dichotomous variable, which reflect the magnitude of effect for each predictor in the model [288]. For continuous variables (FRPs, reaction times, eye tracking features), we employ the delta total (δ_t) as our effect size measure [244]. For subjective scores, upon the normality, using the Shapiro-Wilk test, we use two-way ANOVAs for parametric analysis and ART ANOVAs [599] for the non-parametric data. Furthermore, for post-hoc comparisons, we use either t-test or ART-C test [164] to report our results. For the ANOVA models, we computed effect sizes using Eta squared (η^2), and for ART ANOVAs, we first applied the ART procedure and then calculated Eta squared on the ranked data.

Behavioral Data

Accuracy A linear mixed-effects analysis was conducted to assess the influence of Manifestation and PerceptualLoad on the accuracy, including a random intercept for participants. Manifestation levels AV and VR did not show significant differences from AR, with $b = .00917$, $t(138) = 1.012$, $p = .313$, and $b = .00542$, $t(138) = .598$, $p = .551$, respectively. Low Perceptual Load significantly differed from the difficult level ($b = .0686$, $t(138) = 9.276$, $p < .001$). These findings suggest that while PerceptualLoad has a significant impact on accuracy, the modality of mixed reality (Manifestation) did not exert a significant effect in the given conditions. Specifically, accuracy was higher when the perceptual load was low than the high perceptual load condition. We observed the following odds ratio: in *AV*, the odds ratio was 1.01, 95% CI [0.99, 1.03], suggesting that a one-unit increase in *AV* is associated with a 1.1% increase in the odds of the outcome. Similarly, for *VR*, the



(a) Behavioral Accuracy (b) Reaction Times

Figure 3.4: Boxplot of accuracy and reaction times. a) The figure presents accuracy, showing no significant differences between AV and VR compared to AR, but a notable increase in low perceptual load conditions. b) The figure shows reaction times, with significant reductions in AV and VR compared to AR, and markedly faster in low perceptual load conditions.

odds ratio was approximately 1.00, 95% CI [.98, 1.02], indicating a marginal change in the odds with a one-unit increase. Additionally, for *Low Perceptual Load*, the odds ratio was 1.07, 95% CI [1.05, 1.09], showing a 7% increase in the odds of the outcome compared to the high perceptual load condition.

Reaction Times A linear mixed-effects analysis was conducted to assess the influence of MANIFESTATION and PERCEPTUALLOAD on reaction time (RT), including a random intercept for participants. For the MANIFESTATION, both AV and VR levels exhibited significant reductions in reaction time compared to the AR level: AV ($b = -56.20$,, $t = -2.400$, $p = .017$) and VR ($b = -70.63$, $t = -3.01$, $p = .003$). Additionally, low PERCEPTUALLOAD led to a significant decrease in reaction time compared to the difficult level ($b = -259.70$, $t = -13.55$, $p < .001$). These findings indicate that PERCEPTUALLOAD and MANIFESTATION influence reaction time. Overall, participants reacted faster in the low PERCEPTUALLOAD condition and responded more quickly in the AV and VR conditions than in AR. The model's δ_t indicated significant effects for AV ($b = -.07$, 95% CI [-.12, -.01]) and VR ($b = -.08$, 95% CI [-.14, -.03]). Moreover, PERCEPTUALLOAD in the 'Low' condition showed a substantial effect ($b = -.31$, 95% CI [-.35, -.26]), denoting its large role among the predictors.

EEG Data - Fixation Related P3

A linear mixed-effects analysis was utilized to test the impact of MANIFESTATION and PERCEPTUALLOAD on the P3 amplitude post-fixation onset (see Figure 3.5), integrating a random intercept for individual participants (formula: `P3Amplitude ~ Manifestation * PerceptualLoad + (1|participant)`). The model held significant explanatory power with a conditional R^2 of .44. Within this model, the AV and VR levels of MANIFESTATION significantly showed increased P3 amplitude, demonstrated by beta values of 2.88 ($t = 2.59$,

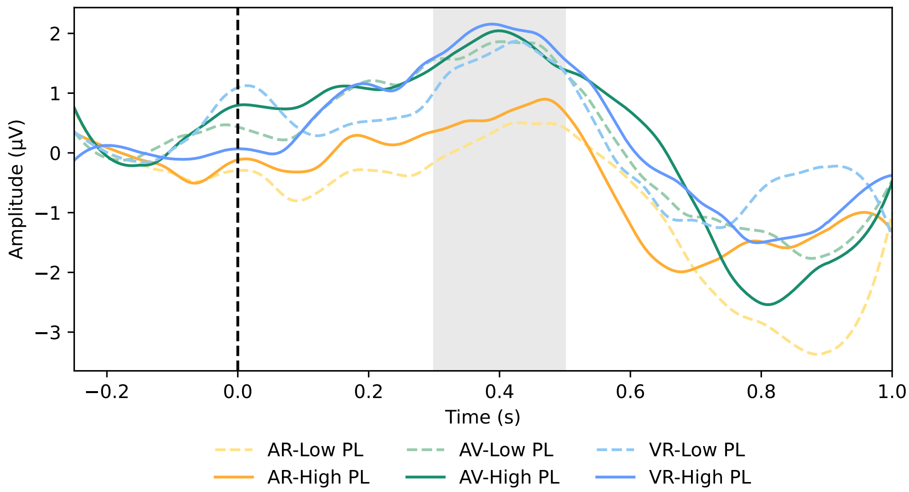


Figure 3.5: Grand Average event-locked to fixation onset. Data reflect the results obtained from parieto-occipital ROI for each MR and PERCEPTUALLOAD condition. The plot suggests a pronounced MR Step influence on P3 amplitude, with marked variations between AV and VR compared to AR.

$p = .011$) and 3.20 ($t = 2.88$, $p = .005$), respectively. In contrast, the low PERCEPTUALLOAD yields a non-significant and minor negative effect, with a beta of -.42 ($t = -.46$, $p = .645$). These findings highlight the main effect of the MANIFESTATION on allocating attentional resources during visual search tasks, as reflected in the P3 amplitude. Specifically, the AV and VR levels were associated with heightened attentional engagement, potentially facilitating more efficient visual search processes compared to the AR level. In contrast, the PERCEPTUALLOAD appeared to have a negligible impact on the resource allocation in visual search. The model's δ_t indicated significant effects for *AV* ($b = .43$, 95% CI [.10, .76]) and *VR* ($b = .48$, 95% CI [.15, .81]). Moreover, *Perceptual Load* (Low) showed a smaller effect ($b = -.06$, 95% CI [-.33, .21]), though its role among the predictors was not as pronounced.

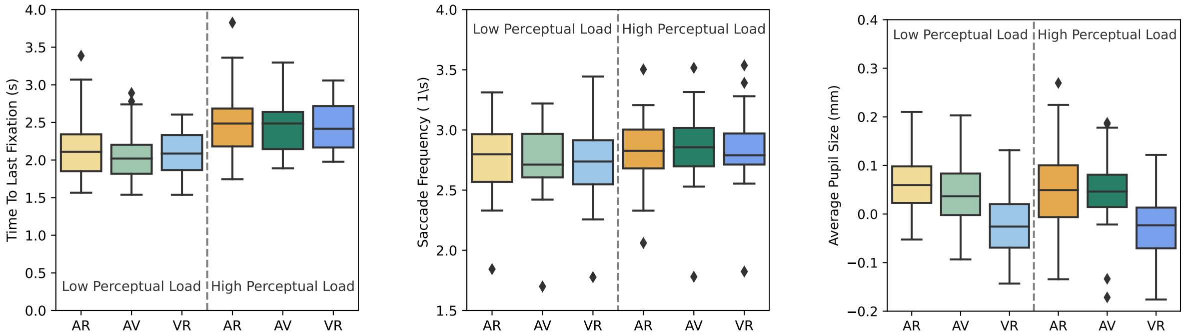
Eye Tracking Data

We used a general linear mixed-effects modeling to evaluate the effect of the independent variables, i.e., MANIFESTATION and PERCEPTUALLOAD on the set of eye tracking features (formula: `ET feature ~ PerceptualLoad * Manifestation + (1|participant)`). The analysis was conducted using the REML method, and here, we highlight the significant findings on the fixed effects and interactions. Overall, no significant interactions were detected.

Fixation Duration Analysis revealed that the effect of MANIFESTATION [AV] on fixation duration was not statistically significant (beta = -.03, 95% CI [-.07, .01], $t(7191) = -1.35$, $p = .179$), with an average fixation duration of 1.78 seconds (SD = .684) for Low and 2.06 seconds (SD = .777) for High Perceptual load conditions. However, the effect of MANIFESTATION [VR] was found to be statistically significant and negative (beta = -.04, 95% CI [-.08, -2.24e-03], $t = -2.07$, $p = .038$), where the average fixation durations were 1.77 seconds (SD = .660) for Low and 2.05 seconds (SD = .775) for High Perceptual conditions. This suggests that the VR condition led to shorter fixation durations than the AV condition. Additionally, the effect of low PERCEPTUALLOAD was statistically significant and negative (beta = -.28, 95% CI [-.31, -.25], $t = -17.51$, $p < .001$). This indicates that lower perceptual load is associated with shorter fixation durations, as evidenced by the average fixation durations of 1.77 seconds (SD = .660) for VR easy and 1.78 seconds (SD = .684) for AV easy conditions, compared to 2.05 seconds (SD = .775) for VR difficult and 2.06 seconds (SD = .777) for AV difficult conditions. In terms of standardized effect sizes, the AR condition was .22, 95% CI [.06, .38], serving as the baseline for comparison. The δ_t for AV was -.04, 95% CI [-.09, .02], indicating a slightly shorter fixation duration compared to AR. For VR, the δ_t was more pronounced at -.05, 95% CI [-011, .00], suggesting a greater reduction in fixation duration relative to AR. The effect of low PERCEPTUALLOAD was the most substantial (δ_t = -.37, 95% CI [-.42, -.33]), indicating a significant decrease in fixation duration compared to the high perceptual load.

Time to Last Fixation (Time to Last Fixation (T2LF)) Results showed that both MANIFESTATION and PERCEPTUALLOAD significantly influence T2LF. MANIFESTATION effect was significant and negative, with AV and VR conditions showing a shorter time to last fixation, reducing the response by -.057 (SE = .027), $t = -2.114$, $p = .035$, and -.072 units (SE = .027), $t = -2.643$, $p = .008$, respectively, compared to the AR condition. Regarding PERCEPTUALLOAD, the Low PERCEPTUALLOAD showed a negative significant effect, significantly decreased by -.385 units (SE = .022), $t = -17.394$, $p < .001$. The standardized effect sizes further elucidate the impact of each variable. Here, the δ_t for AV was -.05, 95% CI [-.12, .02], indicating a modest but not statistically significant reduction in T2LF compared to AR. For VR, the δ_t was -.08, 95% CI [-.16, -.01], suggesting a more pronounced and statistically significant reduction in T2LF compared to AR. The effect of Low PERCEPTUALLOAD was the most substantial, with a δ_t of -.38, 95% CI [-.46, -.31], highlighting its strong negative influence on T2LF. This suggests that conditions with lower perceptual load significantly shorten the time to last fixation.

Fixation Count In the fixed effects analysis for MANIFESTATION factor, both the AV and VR levels induced a non-significant negative effect on Fixation Count, decreasing it by -.17 units (SE = .123), $t = -1.41$, $p = .158$ and -.15 units (SE = .123), $t = -1.24$, $p = .214$, respectively. Furthermore, the low level of PERCEPTUALLOAD emerged with a significant negative impact on Fixation Count, reducing it by -1.87 units in comparison to the 'difficult' level (SE = .1), $t = -18.68$, $p < .001$. Regarding standardized effect sizes, the AR condition was at .23, 95% CI [.10, .36], serving as the baseline for comparison. The δ_t for AV was -.04, 95% CI [-.09, .01], and for VR it was -.03, 95% CI [-.09, .02], indicating slight but non-significant decreases in Fixation Count compared to AR. The δ_t for low PERCEPTUALLOAD was notably more substantial at -.41, 95% CI [-.45, -.37], demonstrating a significant reduction in Fixation Count compared to the high perceptual load.



(a) Time to Last Fixation (b) Average Saccade Frequency (c) Average Pupil Size

Figure 3.6: Boxplot on selected eye tracking features. The left figure illustrates the Time to Last Fixation, showing faster times in low perceptual load and a slight increase in the AV condition. The center figure presents the Average Saccade Frequency, with minimal visual differences between conditions. The right figure depicts Average Pupil Size, revealing a notable decrease from AR to VR, particularly under varying perceptual loads. Error bars represent the standard deviation from the mean.

Saccade Frequency For MANIFESTATION, both the AV and VR levels did not show any significant positive effects on Saccade Frequency, increasing it by .01 units (SE = .017), $t = .79$, $p = .432$ and .0044 units (SE = .017), $t = .26$, $p = .798$, respectively. In contrast, the PERCEPTUALLOAD factor demonstrated a significant and negative influence when the PERCEPTUALLOAD was low, decreasing Saccade Frequency by -.094 units compared to the high PERCEPTUALLOAD level (SE = .014), $t = -6.70$, $p < .001$. This significant negative effect indicates that participants tend to have fewer saccades when tasks are less perceptually demanding, potentially reflecting a reduced workload in these conditions. Regarding the standardized effect sizes: the AR condition, was at .06, 95% CI [-.13, .25]. The δ_t for AV was .02, 95% CI [-.03, .07], and for VR it was approximately .0066, 95% CI [-.04, .06], both showing non-significant changes in Saccade Frequency compared to AR. Conversely, the δ_t for low PERCEPTUALLOAD was -.14, 95% CI [-.18, -.10], marking a significant decrease in Saccade Frequency compared to the high perceptual load, thus aligning with the observation of fewer saccades in less demanding conditions.

Saccade Amplitude When testing MANIFESTATION factor fixed effects, the AV level presented a non-significant negative influence on saccade amplitude, reducing it by -.16 units (SE = .094), $t(7188) = -1.69$, $p = .090$. In the same direction, VR level was non-significant, $t = -.06$, $p = .951$. PERCEPTUALLOAD factor was significant and positive when the level was Low, increasing saccade amplitude by .82 units compared to the 'difficult' level (SE = .077), $t(7188) = 10.77$, $p < .001$. This substantial positive effect suggests that participants have larger saccade amplitudes when the tasks are easier, which might indicate a broader or more relaxed visual exploration strategy under lower PERCEPTUALLOAD conditions. When investigating the standardized effect sizes, AR condition was at -.10, 95% CI [-.25, .04]. The δ_t for AV was -.05, 95% CI [-.10, .01], and for VR it was approximately -.0017, 95% CI [-.05, .05], both showing minimal changes in saccade amplitude compared to AR. However, the δ_t for low

PERCEPTUALLOAD was .24, 95% CI [.19, .28], demonstrating a significant increase in saccade amplitude under less demanding conditions.

Average Pupil Size For MANIFESTATION, both AV and VR levels significantly reduced mean pupil size. The AV level led to a decrease of -.01 units, $t(7191) = -2.57$, $p = .010$, while the more pronounced VR level decreased it by -.08 units, $t(7191) = -15.64$, $p < .001$. These findings suggest varying workloads or visual engagements across the MR continuum. In contrast, the Low PERCEPTUALLOAD did not significantly affect mean pupil size, with a change of .002 units, $t = .44$, $p = .661$. We computed the standardized effect sizes, and we found that for AR condition, it was .16, 95% CI [.02, .29], indicating the baseline mean pupil size. The δ_t for AV was -.07, 95% CI [-.12, -.02], showing a moderate reduction in pupil size compared to AR. However, the effect was more substantial for VR, with a δ_t of -.42, 95% CI [-.47, -.37], indicating a significant decrease in pupil size. Meanwhile, the effect of low PERCEPTUALLOAD on pupil size was negligible (δ_t = .0096, 95% CI [-.03, .05]), reinforcing the finding of its non-significant impact on mean pupil size.

Subjective Data

Perceived Workload The raw NASA TLX scores [237] deviated from normality as per Shapiro-Wilk normality testing $W = .972$, $p < .001$. An ART ANOVA was conducted and showed a significant main effect of MANIFESTATION on perceived workload, $F(2, 100) = 10.273$, $p < .0001$. This suggests that the MANIFESTATION differentially impacts perceived workload. The main effect of PERCEPTUALLOAD was also significant, $F(1, 100) = 76.179$, $p < .0001$, indicating that workload perception varies significantly between high and low PERCEPTUALLOAD. Additionally, there was a significant interaction between MANIFESTATION and PERCEPTUALLOAD, $F(2, 100) = 5.517$, $p < .001$, suggesting that the effect of interaction modality on perceived workload is not consistent across different levels of PERCEPTUALLOAD. Post hoc contrasts for MANIFESTATION revealed a significant difference in perceived workload between AR and AV, $t(100) = 3.45$, $p < .01$, with an increased workload for AR condition. In the same direction, we found a significant difference between AR and VR, $t(100) = 4.278$, $p < .001$. The comparison between AV and VR did not show any significance$t(100) = .841$, $p > .05$. Regarding PERCEPTUALLOAD, there was a significant difference in perceived workload between the high and low PERCEPTUALLOAD conditions, with the first showing increased workload, $t(100) = 8.728$, $p < .0001$. We then compute the effect size as partial eta squared to elucidate the impact of each factor. For MANIFESTATION, the effect size was substantial ($\eta^2_{\text{partial}} = .17$, indicating that this factor accounted for approximately 17% of the variance in perceived workload. The effect size for PERCEPTUALLOAD was even more pronounced ($\eta^2_{\text{partial}} = .43$, explaining about 43% of the variance and underscoring its significant impact on workload perception. Additionally, the interaction between MANIFESTATION and PERCEPTUALLOAD had an effect size $\eta^2_{\text{partial}} = .10$, suggesting a meaningful but less pronounced impact compared to the main effects.

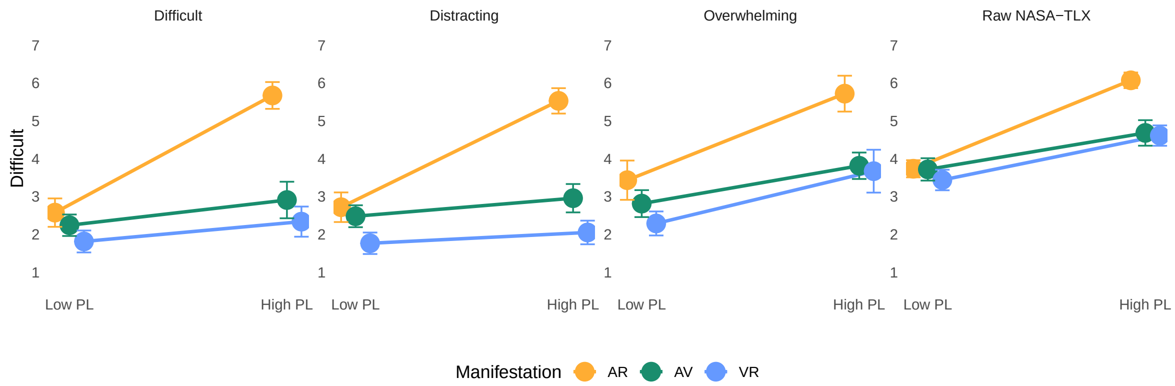


Figure 3.7: *Subjective Data.* Line plots for Raw NASA-TLX, UX Difficult, UX Distracting, UX Overwhelming 7-point Likert scales (normalized to 7-points only for graphical illustration). Here, we found the main effects of MANIFESTATION and PERCEPTUALLOAD across all subjective measures. Error bars are displayed as standard error from the mean.

User Experience

Distracting Participants' ratings on the 7-point Likert scale item "*The background made it very difficult to distinguish the target object from other objects*" deviated from normality as indicated by the Shapiro-Wilk normality test, $W = .845$, $p < .001$. ART ANOVA detected a series of significant results. We found a significant main effect of MANIFESTATION on participants' ratings, $F(2, 100) = 39.669$, $p < .001$. The main effect of PERCEPTUALLOAD was also significant, $F(1, 100) = 36.148$, $p < .001$. Finally, there was a significant interaction between MANIFESTATION and PERCEPTUALLOAD, $F(2, 100) = 5.24$, $p < .001$. Post hoc contrasts for MANIFESTATION demonstrated a significant difference between AR and AV $t(100) = 4.546$, $p < .001$. There was also a significant difference between AR and VR, with the same direction $t(100) = 8.857$, $p < .001$. Lastly, comparing AV and VR again yielded a significant difference $t(100) = 3.617$, $p = .0002$. Regarding PERCEPTUALLOAD, participants rated the difficulty in distinguishing the target object as significantly higher in the high PERCEPTUALLOAD condition than in the Low condition, $t(100) = 6.012$, $p < .001$. The effect sizes analysis revealed the substantial impact of each factor. The partial Eta squared for MANIFESTATION was .44, suggesting that this factor accounted for approximately 44% of the variance in the ratings, indicating a very strong effect. The effect size for PERCEPTUALLOAD was .27, explaining about 27% of the variance, denoting a significant influence on how participants perceived distractions. Additionally, the interaction between MANIFESTATION and PERCEPTUALLOAD had an effect size of .31, highlighting a considerable combined impact of these factors on participants' perception of distractions.

Overwhelming Participants' ratings on the number of items perceived as overwhelming deviated from normality according to Shapiro-Wilk $W = .853$, $p < .001$. An ART ANOVA was performed to analyze the aligned rank-transformed data. There was a significant main effect of MANIFESTATION on the difficulty ratings, $F(2, 100) = 17.752$, $p < .001$. Post hoc contrasts revealed that participants in the AR condition rated it significantly harder to distinguish the target object than those in the AV condition $t(100) = 3.829$, $p < .001$. Similarly, the AR condition was rated as significantly more difficult than the VR condition $t(100) = 5.868$, $p < .001$. The AV and VR conditions did not differ, with participants reporting being subjectively overwhelmed similarly $t(100) = 2.038$, $p = .13$. This suggests that the type of Mixed Reality (MR) step influenced participants' perception of difficulty, with the AR environment being perceived as the most challenging. The effect of PERCEPTUALLOAD was also significant, $F(1, 100) = 52.14$, $p < .001$. Participants rated the task as significantly more challenging in the high PERCEPTUALLOAD condition than in the low PERCEPTUALLOAD condition, with an estimated difference of 33, $t(100) = 7.22$, $p < .001$. This indicates that irrespective of the MR step, tasks with a high PERCEPTUALLOAD were perceived as inherently more difficult. Furthermore, there was a significant interaction between MANIFESTATION and PERCEPTUALLOAD, $F(2, 100) = 3.919$, $p < .05$. The partial Eta squared for MANIFESTATION was .26, indicating that approximately 26% of the variance in difficulty ratings can be attributed to the MANIFESTATION factor, suggesting a strong influence. The effect size for PERCEPTUALLOAD was even more significant at .34, accounting for about 34% of the variance and highlighting its considerable impact on perceived task challenge. The interaction effect between MANIFESTATION and PERCEPTUALLOAD had an effect size of .07, denoting a smaller but still noteworthy combined influence of these factors on participants' perception of being overwhelmed.

Difficult Scores on the perceived difficulty item due to the MR background deviated from normality as determined by the Shapiro-Wilk normality test, $W = .820$, $p < .001$. An ART ANOVA first detects a main effect of MANIFESTATION on the difficulty ratings, $F(2, 100) = 30.621$, $p < .001$. Post hoc comparisons elucidated that participants perceived the AR condition as significantly more difficult than the AV condition, $t(100) = 5.547$, $p < .001$. Similarly, distinguishing the target object in the AR environment was deemed significantly more challenging than in the VR environment $t(100) = 7.554$, $p < .001$. The AV and VR conditions showed no perceptible differences $t(100) = 2.006$, $p = .14$. This suggests that the choice of MR step distinctively influenced participants' perceptions, with the AR modality consistently emerging as the most taxing. Regarding PERCEPTUALLOAD, there was a clear distinction between the high and low PERCEPTUALLOAD conditions. Participants perceived the high PERCEPTUALLOAD tasks as considerably more challenging than their low PERCEPTUALLOAD counterparts $t(100) = 7.366$, $p < .001$. This underscores the influence of task complexity on participants' perceptions, independent of the MR environment. Lastly, the interaction between MANIFESTATION and PERCEPTUALLOAD was significant, $F(2, 100) = 19.248$, $p < .001$. This interaction implies that the effects of MR step on perceived difficulty were not uniform but contingent upon the task's inherent PERCEPTUALLOAD. The partial Eta squared for MANIFESTATION was .38, indicating that MANIFESTATION accounted for approximately 38% of the variance in perceived difficulty, marking a substantial effect. The effect size for PERCEPTUALLOAD was .35, explaining about 35% of the variance and highlighting its significant role in shaping difficulty perceptions. Additionally, the interaction effect between MANIFESTATION and PERCEPTUALLOAD had an effect size of .28, demonstrating a considerable combined influence on participants' perceptions.

3.1.3 Discussion

We evaluated the impact of different MR manifestations and task difficulty on behavioral and physiological correlates of visual search efficiency, resource allocation, and perceived workload.

Impact of Manifestation on User Performance With our first research question (**RQ1**), we wanted to investigate whether our independent variable MANIFESTATION affects users' performance in a visual search task. In particular, we were interested in objective performance measures, such as accuracy and reaction times, as well as subjective measures, such as perceived workload and individual task perception. While we did not find any differences in accuracy, we found that participants responded faster in the AV and VR conditions than in AR. Given the short duration of our search task (5s), drops in accuracy can mostly be attributed to misses rather than false positives. Thus, it is likely that we introduced a ceiling effect, and longer task durations would show differences in accuracy as well. Nonetheless, our findings of significantly different reaction times are consistently supported by our subjective measures (perceived workload and individual task perception), which depicted how challenging and demanding the AR condition was compared to the others. Participants expressed that they found the background more distracting and the overall task relatively more overwhelming in the AR environment. Compared to previous studies that explored concrete manifestations individually (e.g., in VR [145], in AR [388], or even outside of MR [473]), our work can address the gap of comparing user performance between manifestations. Given the decreased performance of the AR condition, applying search-related AR applications may be problematic, particularly in safety-critical domains such as healthcare, aviation, or emergency response [249, 285], where a slightly delayed response can have serious consequences. Considering this, designers may need to strike a delicate balance in MR applications so that augmentation enhances the user's performance without overwhelming their cognitive capacities [157]. Validating our experimental design, a low perceptual load resulted in higher accuracy rates and shorter reaction times than did a high perceptual load.

Cognitive Resource Allocation per Manifestation To reiterate **RQ2**, we were interested in resource allocations across the different manifestations. In FRPs, we found increased P3 amplitude, indicating enhanced attentional and more efficient visual search processes in both the AV and VR conditions compared to AR. First, we confirm that FRPs are elicited across MR manifestations. FRPs are grounded in natural scanning behaviors, specifically the act of fixating on relevant information, a phenomenon central to visual search tasks and routinely manifested in everyday settings. Second, we show that the environmental visual feature of MR manifestations influences task demands and can affect electrophysiological correlates of resource allocation in visual tasks. Previous ERP studies have shown late (P3) visual ERPs during high auditory working memory load compared to low auditory working memory load [452] or when users are asked to ignore distracting auditory stimuli [204]. However, in our study, participants were required to maintain central eye fixation throughout the experiment, though we only employed a fixation cross for a duration of 1750 ms.

Our findings align with Ries et al. [479], demonstrating how fixation-related P3 can discriminate task demands, extending them to MR, particularly in its AR manifestation. This reinforces the notion that perceptual load and manifestation types significantly influence cognitive resource allocation in MR environments, as evidenced by the elicited FRPs. Moreover, this outcome is more relevant as it presents a novel contribution by being the first to elucidate FRPs across different MR environments. This result is particularly significant as FRPs are ecological ERPs, offering practical insights into real-world cognitive processing. This finding has significant implications for the design of MR systems, suggesting that optimizing attentional resources should be a key consideration across the MR continuum. For example, in environments where FRPs indicate higher need for cognitive resources, designers might seek to reduce complexity or introduce features to aid information processing. These features could include directional cues, shifting the user's attention [225] or contextual highlighting and filtering, limiting the visual complexity [95].

Eye Tracking Correlates of Visual Search Efficiency To approach **RQ3**, we considered evidence of the impact of the manifestation on our eye-tracking correlates of visual search efficiency. Drawing on foundational studies in visual search, such as those by [161], we know that stimulus similarity can influence search efficiency. Their work suggests that as target and distractor similarity increases, the search becomes more difficult. Thus, one would assume that the search for virtual objects in AV is easier than searching for them in VR due to the increased perceptual consistency, which could slow the search process [92, 492, 517]. However, VR fostered more efficient visual processing than AV, mainly evidenced by shorter fixation durations. Such differences might suggest that transparency and see-through objects in AR and AV elevate the attentional load, possibly due to their decreased visual salience [603]. Further integrating the Guided Search theory, it's plausible that the VR and AV modes streamlined the preattentive processing stages, guiding attention more effectively than in the AR mode. Regarding pupil size, unlike the abundant literature that often associates higher perceptual load with increased pupil dilation [75, 417], we did not observe significant changes in the MR continuum. This could suggest a ceiling effect, in which the perceptual load under the conditions was high enough to maintain a consistent pupil size. However, even if pupil size did not differ between PerceptualLoad load levels, the Low condition was associated with faster visual processing. This was evidenced by shorter fixation durations and shorter times to the last fixation, potentially reflecting reduced workload and enhanced search efficiency. This trend of reduced cognitive demand with lower perceptual loads was further manifested in fixation count and saccade frequency metrics, where participants exhibited fewer fixations and saccades, possibly because they adopted a more relaxed visual exploration strategy.

Implications for Adaptive Systems Taking it one step further, EEG data not only holds the potential to study the electrophysiological correlates of attention but can also be used for adaptive systems [412]. FRPs have recently been employed as a feature to differentiate between relevant and non-relevant information in single-trial classification [513, 591]. Most current BCIs rely on stimulus-evoked signals for target identification, typically within the confines of fixed-location stimulus displays [412, 587]. In contrast, [187] demonstrated the potential of BCIs based on FRPs, which offer a more flexible approach. FRP-based BCIs are not constrained by stimulus-dependent tasks, allowing them to operate effectively in more complex and dynamic environments. Drawing from our results, the variation in FRP amplitude across MR manifestations indicates that users' resources and processing efficiency differ between AR, AV, and VR environments. This outcome could guide the development of adaptive interfaces by providing cues to the system on which objects in a scene are relevant to the user, allowing for modulating content or background virtuality according to the detected resource capacity. Building on this, the overarching goal is to develop transitional interfaces capable of seamlessly switching between different points on the MR continuum [21, 585]. Previous transitional interfaces found applications for visual analytics [264], collaborative learning, and manipulation tasks [486], yet their application in MR visual search represents a novel direction. Here, we envision interfaces that could dynamically switch across AR, AV, and VR by utilizing FRPamplitude variations adapting demands, task features, and environment. Here, we argue that these aspects equally influence UX in MR interactions. However, it is important to acknowledge that BCIs in MR systems can face specific challenges. A key consideration is that users might not wear MR headsets for prolonged sessions due to their still-to-be-improved usability and ergonomics [140, 380]. Moreover, the ongoing endeavor to embed EEG headsets in VR and AR largely depends on dry setups [47], with unclear outcomes on signal quality and its impact on high-accuracy classification [47], even though high-density dry setups appear promising for application and performance [184].

Limitations and Future Work In our research, we investigated the effects of the general background environment. Moving forward and examining MR interactions at a finer level, we propose investigating the dual nature of targets and distractors encompassing virtual and physical objects. Previous work has shown that real-world objects evoke more sustained, stronger ERP responses than 2D images, but it is unclear if this also applies to virtual objects or FRPs [373, 406]. Moreover, our study's representation of the MR manifestations world was relatively simplistic, predominantly involving virtual stimuli. Even if our stimuli allowed for a more controlled setting, future work should aim for a more realistic simulation by incorporating visual clutter [110, 380]. To ensure reliable and robust eye tracking results, we controlled that presented stimuli had consistent luminances and transparency. However, to design an adaptive system, we propose investigating varying luminance levels, specifically in AR, to evaluate the effect on pupil size and if such results can be applied in more ecological contexts.

Regarding transparency, we chose a video-see-through approach to avoid potential confounders associated with room luminance that are more prevalent in optical see-through MR systems. Instead, in optical see-through AR, the light from virtual surfaces is a blend of display-emitted light and environmental light, leading to semitransparent appearances distinct from real-world objects. This difference can influence the saliency of the virtual content [122] and be associated with increased temporal uncertainty in FRPs which may impact the accuracy of attention decoding in MR [560]. Moving forward, future research in optical see-through MR could explore how changes in ambient lighting affect the perception of transparency in virtual objects as compared to video see-through. This might involve studying how varying levels of environmental illumination influence the saliency of virtual content and how users adapt their visual attention in response to these changes. Such investigations could inform the development of adaptive optical see-through MR systems that adjust the rendering of virtual objects to optimize their visibility and coherence with the real-world environment.

Finally, we found a difference between the main effect on accuracy and the effects of MR Manifestation and PerceptualLoad, which we interpret as a possible change in the visual search strategy. To verify, we propose investigating the performance trade-off via cognitive modeling using the Drift Diffusion Model (DDM) [547]. DDM enables simultaneous analysis of accuracy and reaction times to extract parameters that underlie performance trade-offs. Thus, it already applies to visual search [604] and HCI [99].

3.1.4 Open Science

We encourage readers to reproduce and extend our results. Our collected dataset, MR scenarios, 3D models, and analysis scripts are open-sourced and available at this link: `https://osf.io/fncj4/`.

3.2 Study 2: Visual Search in Augmented Reality and Virtuality

Previous work in visual search showed processing virtual and physical cues simultaneously is demanding, suggesting the need for MR systems to account for these perceptual differences in their design to enhance usability [303, 416]. Thus, in MR, visual search presents varying degrees of difficulty based on their representation. Although users can currently differentiate between physical and virtual objects due to differences in fidelity, anticipated advancements in MR technology aim to merge real and virtual elements even more [21]. This convergence will likely increase the complexity of visual search tasks by making the distinction between real and virtual content less apparent, thereby placing a greater demand on users' visual processing capabilities [273]. Further, there is evidence that AR might introduce detrimental effects like split attention and visual complexity, as shown in contexts such as medical surgery [153]. Moreover, exploring Augmented Virtuality's implications in visual search is still in its infancy [529]. Despite the body of work investigating visual search within MR, understanding how users process virtual and physical target information in Augmented Reality (AR) and Augmented Virtuality (AV) remains elusive. Thus, we explore such a research gap by systematically examining visual search tasks across the Reality-Virtuality continuum. We focus on actualities, i.e., the currently experienced reality of a user on the Reality-Virtuality Continuum [21], involving a shared blend of virtual and physical information, whether distracting or target objects. We conducted a within-subjects user study with two different ACTUALITIES: *AR* and *AV*, where participants engaged in searching for two different types of TARGET that were either *Virtual* or *Physical*. Our objectives are threefold: first, to determine the impact of different actualities (AR vs. AV) and the nature of the targets (physical vs. virtual) on participants' performance and perceived workload. Second, we investigate how distractors are suppressed during visual search tasks, as indicated by the event-related potential (ERP) Distractor Positivity component [197]. Finally, we seek to evaluate visual search efficiency across the MR continuum and different target types by analyzing eye-tracking metrics such as fixations, saccades, and index of pupillary activity, which serve as indicators of visual search efficiency and cognitive load.

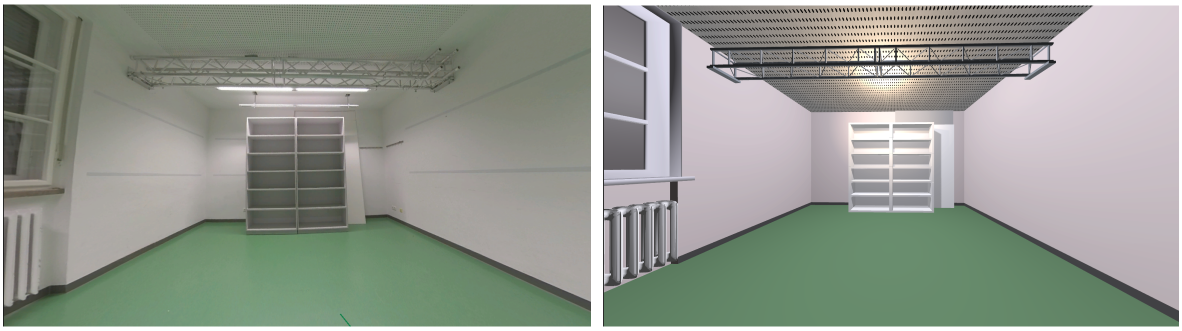
3.2.1 Mixed Reality Visual Search Environment

First, we designed an MR environment that allows users to perform a visual search task across two actualities of the MR continuum, *AR* and *AV*, where physical objects and virtual objects are presented as search targets. We employed a real-world setup to represent the physical world, objects, and their virtual counterparts to achieve this. This approach embraced an ecological methodology, drawing inspiration from David et al. [131], where participants selected objects placed on shelves, mirroring their real-world analogs. We implemented the visual search task using two different models of the same scene with different levels of virtuality, see Figure 3.8a and Figure 3.8b.

Implementation of the Real World

Physical Environment For the physical environment, we chose a room at our institution, featuring a minimalist aesthetic with white walls, a green floor, and a grey shelf is the focal point for the visual search task, see Figure 3.8a. This setting was selected to minimize visual noise and eliminate extraneous details that could detract from the task, ensuring that participants' attention was drawn primarily to the target objects. The room's simplicity also allowed for the visibility of virtual and physical objects, promoting an unambiguous and straightforward interaction in the AR simulation.

Physical Objects The stimuli object placed on the shelves set consisted of four physical objects resembling real-life counterparts: a sphere, e.g., a soccer ball; a cylinder, e.g., a soda



(a) Physical Search Environment (b) Virtual Search Environment

Figure 3.8: MR Visual Search Environments. We situated the visual search task in two environments: a physical one to display AR content and a virtual one to display the physical and virtual objects (AV condition). For the AR scenario, we chose a low-complexity room at our institution to avoid environmental distractions from the task. We modeled the AV environment as closely as possible to its' AR counterpart and controlled for luminance.

can; a cube, e.g., a rolling dice; and a pyramid, e.g., building blocks. We fabricated physical objects using extruded polystyrene with a heating wire and foam cutter with angle adjustment. Physical objects are not directly derived from virtual objects but are physical replicas of our virtual objects. We placed them on the shelf in our study room and photographed them from where participants were seated during the study, i.e., 3 m distance. As our virtual environment precisely mimics the study room, these photographs were straightforward to blend with the virtual environment. We chose photographs over a live video feed to better control the experiment by avoiding time-consuming reconfigurations between the trials and eliminating the potential for human error introduced through the condition assembly on the spot. Objects were fitted into a cube of 12 cm per dimension (2.75° visual angle) based on [363], i.e., half the shelf compartment height (24 cm) with a volume of 1,728 cm^3, ensuring that all three objects are 10 cm equidistant. Each shelf is 96 cm wide, allowing four (virtual/physical) objects to be equally spaced (12 cm distance from each other) per shelf compartment. To design the Physical Object - Visual Search trials, we took 200 pictures with 13 physical object search displays with an InstaVR 360 Pro2 (7680×4320 pixels, 120 fps) placed in randomized locations to design the trials for the physical objects that were not overlapping with the location of the virtual objects on the shelf.

Implementation of the Virtual World

Virtual Environment For the virtual world scene, as depicted in Figure 3.8b, we employed a systematic modeling approach to recreating a laboratory setting that mirrors its physical counterpart accurately, following previous work [339]. Starting with a low-fidelity model, we constructed the basic geometry to outline all principal features of the room, such as the shelf, walls, and floor, ensuring clear identification without the inclusion of details like door knobs or complex textures. Progressing to a medium-fidelity model, we refined the geometry, adding elements such as detailed window frames, accompanied by low-resolution textures to enhance visual depth. Overall, we aimed for a high-fidelity representation, with increased polygon count for all objects and the addition of all visible minor features. High-resolution textures were employed to achieve a realistic appearance, and we introduced baked lighting to incorporate static shadows. Throughout this process, we meticulously controlled for luminance to ensure consistent lighting and visual perception across different fidelity levels.

Virtual Objects The set of virtual objects was the same as for the physical objects, see paragraph 3.2.1. For the virtual object color coding, we color-picked the original color from the physical objects, resulting in the following RGB values: red `#BF1818`, blue `#377EB8`, purple `#e92053`, and yellow `#FF0000`. Virtual objects have virtual, opaque color features. To maintain spatial consistency in our visual search environment, we ensured that the virtual objects were designed to match the dimensions of the physical objects. Adhering to the specifications for the physical objects, we rendered each virtual object within a virtual cube of 12 cm per dimension, corresponding to a 2.75° visual angle and occupying a volume of 1,728 cm^3[363].

Blending the Worlds

Rendering AR For the AR condition, which integrates a physical background with physical and virtual objects simultaneously, we utilized a high-resolution image of the physical environment (captured with an Insta360 Pro 2 at 8K resolution) to present the physical items at predetermined shelf spots. Concurrently, virtual objects were superimposed onto specific, generated locations on the shelves, ensuring they did not overlap with their physical counterparts. This was accomplished by capturing images of the physical objects positioned on the shelves, maintaining a consistent distance and luminance, to serve as the backdrop for the subsequent overlay of virtual objects. Following this procedure, we ensured that the placement of physical objects in the real world was consistent, while simultaneously displaying virtual content in a controlled manner. This approach allowed for a controlled integration of physical and virtual elements within the AR environment. For displaying the virtual objects, we opted for simulating a video-see-through AR display within a sphere for its capacity to superimpose digital information directly onto the user's view of the real world [494]. The need for precision and consistency in the presentation of virtual objects drove this choice. Unlike typical AR implementations where objects might possess a degree of transparency, our approach ensured that the virtual objects were solid and visually consistent, thus avoiding confounders on object saliency due to transparency [268].

Rendering AV In the AV condition, we rendered the virtual world with virtual and physical objects on the shelves. Here, virtual objects were presented in randomized locations. To display the real objects, we created a Unity shader that renders circular sections from the sphere of the AR condition. This shader was designed to replicate the real-world appearance of the objects within the virtual environment, maintaining their texture and depth cues to preserve the natural light and depth of their physical counterparts. Each section includes an inner circle with an opacity of 1 and an outer circle with decreasing opacity as the rendered point is farther from the circle's center, resulting in a fade-out. The circles and radii were calculated using the world positions of the pixels. Using the texture position would result in ellipses due to the sphere's curvature. To insert the real objects into the virtual shelf of the AV scene, the sphere with the custom shader was positioned so that the shelf of the sphere and the virtual cabinet were exactly on top of each other. The cabinet is rendered first, then virtual objects, and finally the sphere. This ensures that the objects are not covered. The rendering order in Unity is as follows: cabinet (queue: 2000, z-buffer: on), virtual objects (queue: 2500, z-buffer: on), and sphere (queue: 3000, z-buffer: off).

(a) AR Search Display (b) AV Search Display

Figure 3.9: The visual search displays for AR and AV conditions with physical and virtual targets. In (a), participants have to search for a blue cube. The target object is physical. In (b), participants must search for a purple sphere. The target object is virtual. We did not place objects on the highest and lowest compartments to avoid object occlusion and limited visibility.

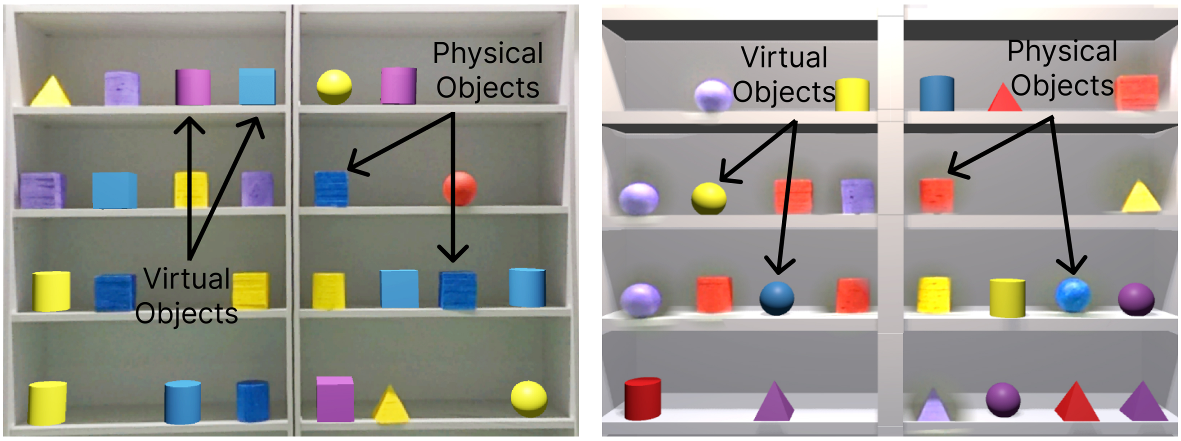
3.2.2 User Study

Our study investigates users' performance identifying physical and virtual target information in AR and AV, see Figure 3.9. We engaged participants in an adapted visual search task by Dey et al. [145] and situated the task in AR and AV while presenting physical and virtual target objects. We used a 2×2 within-participants experimental design with the independent variables ACTUALITY (two levels: *AR*, and *AV*) and TARGET (two levels: *Physical* / *Virtual*). ACTUALITY describes the envisioned the user is immersed in, and TARGET describes the type of objects, while distractors are always presented of both types. Drawing from previous work from visual search, we formulate the following research questions:

RQ1: Do different actualities impact performance and perceived workload differently?

RQ2: Do the MR actuality and target type impact and eye tracking correlates of visual search efficiency (fixations and saccades), and workload (IPA)?

RQ3: How does distractor suppression in a visual search task vary when searching for target and physical objects across the MR actualities, as indexed by Event-Related Distractor Positivity?



(a) AR (b) AV

Figure 3.10: Objects display in AR and AV conditions. We displayed virtual and physical objects across conditions over two different compartments on the shelves in rows. This ensured objects were equidistant with a constant distance of 10 cm.

Procedure

Upon arrival, we briefed participants about the study and gave informed consent. Then, we prepped the water-based EEG data acquisition. Next, we asked them to wear the HTC Vive Pro Eye headset and perform a five-point eye-tracking calibration. The main part started with participants completing a training phase and experiencing all the experimental conditions. This training phase comprised 20 visual search trials, i.e., 5 with physical targets in AR, 5 with virtual targets in AR, 5 with physical targets in AV, and 5 with virtual targets in AV. Participants needed to achieve at least 80% to proceed; failing that, they repeated the training to meet this criterion. Next, we guided participants through the four conditions containing 100 trials. To avoid learning effects, we counterbalanced the order of conditions in a balanced Latin Williams square design with four levels [582]. After each block, participants responded to the raw NASA TLX questionnaire [237]. Questionnaires were administered through the VR Questionnaires Toolkit [181]. The entire session averaged an hour in duration, which we compensated with 12 EUR.

Task

Participants carried out the visual search task in two MR environments: AR and AV. While engaged in one of the two conditions, they were presented with 25 physical and virtual objects placed on a shelf. To select the target item from 24 distractors, they used the trigger button on the VIVE controller. The target object's name was displayed laterally (left or right) in the participant's view to be capable of identifying it. This target display's location (either left or right) was randomly varied across different trials to prevent habituation effects. We chose to display the name and not the picture of the target object to ensure that participants were not biased toward recognizing either the virtual or physical version of the target. Presenting the name rather than the image ensures that both versions of the target are treated equally in the

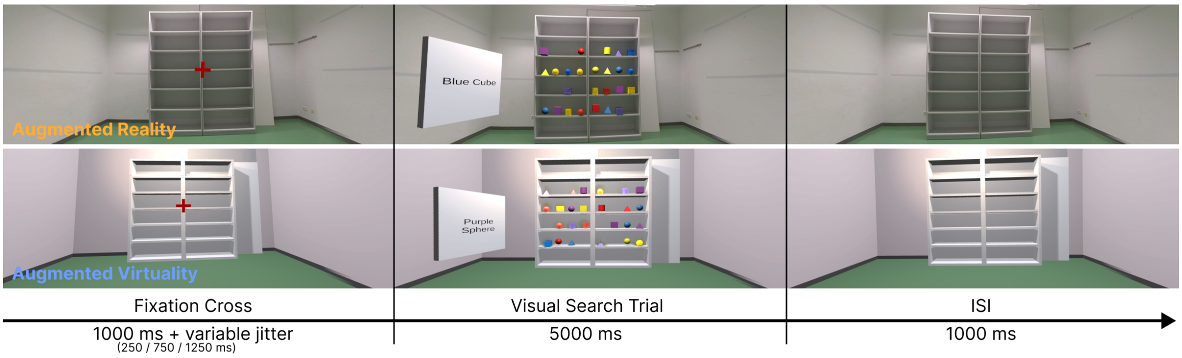


Figure 3.11: Trial Structure in the two Conditions. The visual search trial comprised three stages: Initially, a fixation cross was presented for a baseline of 1000 ms, supplemented by a random jitter of 250 ms, 750 ms, or 1250 ms, leading to a total fixation cross duration of 1250 ms to 2250 ms per trial. Following this, participants were allotted 5000 ms to discern the target from among the distractors, and this was succeeded by a 1000 ms interstimulus interval (ISI). Each participant completed 100 trials per experimental condition.

search process, as participants rely on their understanding and interpretation of the name rather than pre-existing visual features from an image. To enforce this, we did not previously inform participants of the physical or virtual nature of the target, which directly required them to identify the object that best matched the description. Participants needed to scan the MR environment visually, aiming at their chosen object using the controller's ray cast to make a selection. Once the target object was aligned with the ray cast, they pressed the VIVE controller's trigger button to confirm their choice. Participants held the controller with their dominant hand and were encouraged to respond quickly and accurately.

Trial Structure

We designed our trial based on Forschack et al. [191], with a real-world visual search task approach in mind [602]. The structure of the task, was as follows: (1) we asked participants to fixate a red fixation cross (+) with a pseudorandom duration (1250, 1500, or 1750 ms) at the center of the target display, see Figure 3.11. (2) Participants visually searched for the target object; the objects disappeared after selection. (3) After 5000ms, an inter-stimulus interval (ISI) of 1000ms was presented with no cross or objects presented to reset the neural and attentional reserve and avoid fatigue effects [20, 606]. Participants had 5000 ms after visual search display onset to select the target among distractors. For a trial visualization, refer to Figure 3.11.

Stimuli

The stimulus set included four objects, both in virtual and physical forms, designed to mirror everyday items: a sphere (resembling a soccer ball), a cylinder (similar to a soda can), a cube (akin to rolling dice), and a pyramid (comparable to building blocks). These stimuli were presented in one of four colors: red `#BF1818`, blue `#0000FF`, purple `#984EA3`, and yellow

#FF0000. Those objects were placed over a shelf of four rows in two compartments. Each compartment has sixteen possible locations. Thus, objects are spawned for 32 positions. In those 32 positions, we spawn 25 objects. Overall, 1 is the target, 12 are virtual distractors, and 12 are physical distractors. Of the 24 distractors, a third shares the shape of the target (8), a third shares the color of the target (8), and a third has a different shape and color from the target (8). In this way, we control for feature confounders contributing to the overall perceptual load of the objects that could impact visual search [487]. Objects were fitted into a cube of 12 cm per dimension(2.75 degrees of visual angle) based on [363], i.e., half the shelf compartment height (24 cm) with a volume of 1,728 cm ^{3}, ensuring that all three objects are 10 cm equidistant, see Figure 3.10b for the AV condition and Figure 3.10a for the AR condition.

Measurements

We collected a set of multi-modal variables: visual search accuracy, reaction times, missed search trials, PD peak amplitude from EEG, eye tracking features (IPA, last fixation duration, fixation count, saccade frequency), and subjective workload (raw NASA TLX [237]).

3.2.3 Apparatus

We implemented the visual search task using Unity (Version 2022.3.21f1 LTS) and presented the AR and AV conditions through an HTC VIVE Eye Pro headset, with a display resolution of 2880×1600 pixels and a 110-degree field of view. We used the MR toolkit *VRception* [224] for the implementation. The environment tracking employed two HTC Vive 2.0 lighthouses. We used a LiveAmp (BrainProducts, Germany) amplifier to record EEG signals at 500 Hz for EEG recording. We acquired eye-tracking data at 120 Hz using the HTC Vive Pro Eye headset. To integrate and stream physiological data within our Unity VR setup, we employed the Lab Streaming Layer (LSL) framework[2]. The data was then directed to our acquisition PC, which ran on Windows 10 with an Intel Core i7-11700K processor (3.60 GHz, 16GB RAM).

To validate our scenes, we measured luminance across conditions. We measured the luminance inside the headset using a lux meter sensor (LT300, Extech, USA). Using 50 measurements per condition, we found an average luminance for the AR environment of 57.98 (SD=1.41) and 58.37 (SD=1.36) lux in AV. Those values align with luminance guidelines (below 200 nits) based on eye-tracking best practices to avoid confounders for pupil size computation [85, 375].

EEG Recording & Preprocessing

We acquired EEG data (sampling rate = 500 Hz) via LiveAmp amplifier from 32 water-based electrodes from the R-Net elastic cap (Fp1, F3, F7, F9, FC5, FC1, C3, T7, CP5, CP1, Pz, P3, P7, P9, O1, Oz, O2, P10, P8, P4, CP2, CP6, T8, C4, FC2, FC6, F10, F8, F4, Fp2, Fz). We kept

[2] https://github.com/labstreaminglayer/

impedance levels below $\leq$20 kΩ. We set the reference at FCz during the recording, while FPz was used as ground. Electrodes were placed using the International 10-20 layout. For time synchronization with the MR environment, we employed the Lab Streaming Layer Framework, while for preprocessing and analysis, we used the MNE-Python Toolbox [216]. We first automatically detected bad or outliers channels via random sample consensus (RANSAC) method [52] of spherical splines for estimating scalp potential based on algorithms proposed by Perrin [436]. We then applied a notch filter (50 Hz) and band-passed the signal between (1-15 Hz) to remove high and low-frequency noise. We then re-referenced to the common average reference (CAR). We applied an Independent Component Analysis (ICA) for artifact detection and correction with an extended Infomax algorithm [342]. We automatized the labeling and rejection process of ICA components via the MNE plugin "ICLabel" [443]. Epochs that showed blinks, eye movement, muscle, or single-channel artifacts in any of the electrodes were rejected. On average, we removed $.33$ ($SD = .353$) independent components within each participant. Only trials with a correct response and last fixation on the target were used; error trials or distractor fixations were excluded from analyses.

ERP Analysis

We segmented continuous signals between 200 ms before and 1000 ms after the search display onset, removing a 200 ms baseline before stimulus onset. The Pd component was quantified as positive average peak amplitudes in the 300 – 900 ms. This window is centered upon the peak latency of each component in the grand average waveforms [484]. For ERP computation, we chose electrodes Oz based on previous work [406].

Eye Tracking Recording & Preprocessing

We acquired three-dimensional head position and orientation data from Unity and recorded eye-tracking metrics using the HTC Vive Pro Eye headset's integrated eye tracker (120 Hz). This data capture utilized the SRAnipal eye tracking SDK, which provided vectors indicating eye direction relative to both the head and the world. Blinks and related artifacts were removed. We defined blinks as missing data points from the eye tracker, with durations ranging from 75ms to 500ms. Data 200ms before and after the blinks were removed [30]. Removed data were linearly interpolated and smoothed with a 6th order Butterworth filter whose cutoff frequency was set at .15 Hz [586]. For analysis, these vectors—head direction, eye-in-world direction, and eye-in-head direction—were initially translated into two-dimensional Fick angles. This process was based on the Fick-gimbal method described by Haslwanter [239]. The transformation involved two rotational movements: one around the vertical axis and another around the horizontal axis within the former, enabling us to precisely determine the vectors' positions. We then utilized these 2D Fick angles representing the eye and head orientations as the foundation for further investigation.

Fixation-Saccade Analysis

We analyzed fixation and saccade data using pymovements [323]. For identifying fixations, we utilized pymovements' application of the ID-T algorithm [493], setting the fixation thresholds to a minimum duration of 83ms and a maximum dispersion of 1.8 degrees, in line with prior research [39, 562]. This approach enabled us to derive key metrics related to fixation, including total and average fixation duration, number of fixations, and the interval from the visual search's onset to its final fixation. In analyzing saccades, we applied the microsaccade algorithm offered by pymovements [168], which facilitated the measurement of saccade amplitude—the angular distance between the start and end points of a saccade—and saccade frequency, calculated by the total number of saccades over the trial length.

Index of Pupillary Activity Analysis

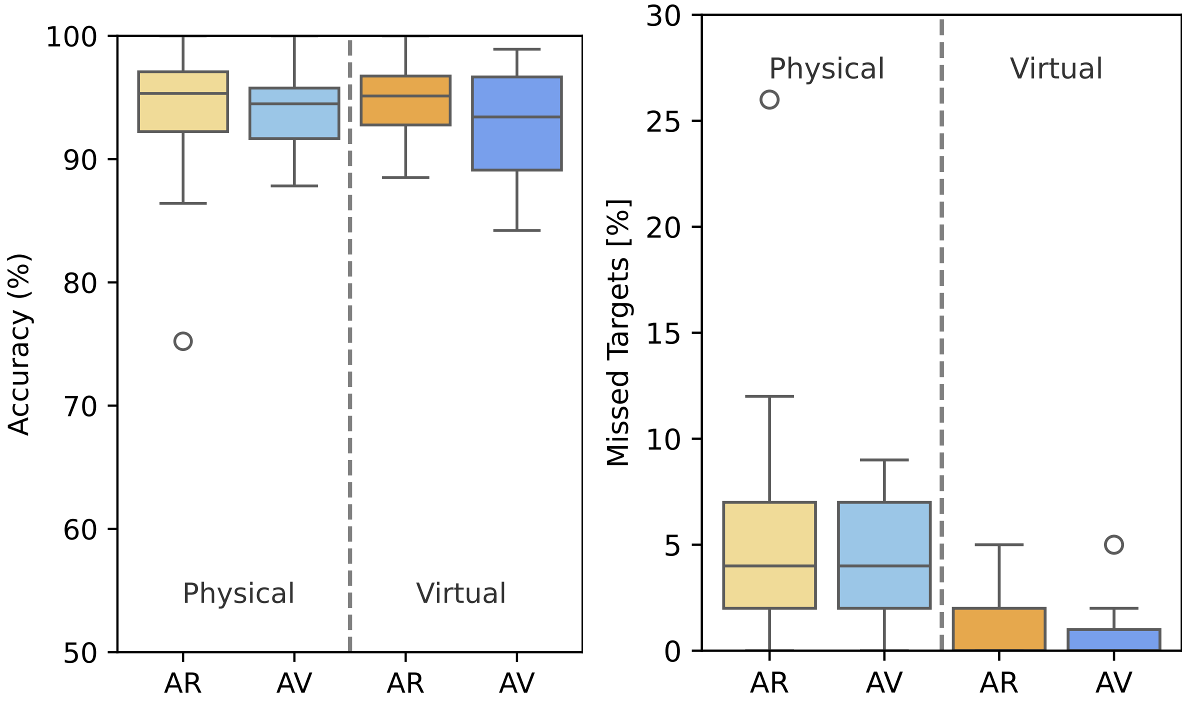
We employed the implementation by Duchowski et al. [160] for computing IPA. Thus, we utilized discrete wavelet transforms (DWT) to analyze pupil diameter signals, starting with a two-level DWT to break down the signal and explore its variability. We normalized the wavelet coefficients to ensure a uniform analysis and identified key peaks in the signal to mark significant changes in pupil diameter. We then applied a universal threshold to filter out noise.

Participants

This study engaged 20 volunteers ($M = 24.85$, $SD = 4.67$; comprising 11 females, 9 males, none diverse), recruited through institutional email lists and convenience sampling methods. Due to significant EEG data noise, identified through RANSAC algorithm showing that more than 50% of the electrode (Oz) used for ERP computation were classified "bad," thus impairing data quality due to low signal-to-noise ratio as in Bigdely et al. [52]. Overall, we did not exclude any participants. The participants' familiarity with AR, AV, and VR technologies was assessed, following previous work [96]. All participants had prior exposure to AR ($M = 2.76$, $SD = 1.56$), AV ($M = 2.88$, $SD = 1.97$), and VR ($M = 3.8$, $SD = 1.83$) technologies, rated on a familiarity scale ranging from 1 (not familiar at all) to 7 (extremely familiar). None reported any neurological, psychological, or psychiatric disorders.

3.2.4 Results

In this section, we first present the results of our multimodal evaluation. We employ a GLMM to investigate differences in the behavioral measures, ERPs, and eye-tracking features. We determined subjective workload scores using the Shapiro-Wilk test, t-tests, or paired samples of the Wilcoxon test.



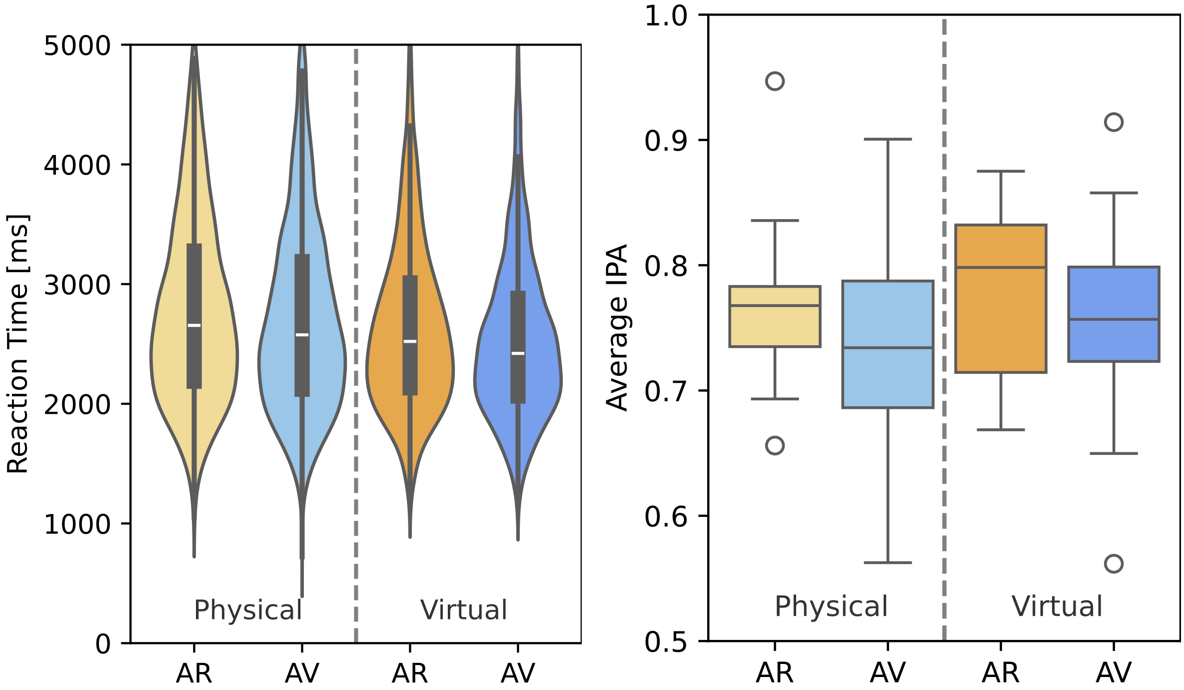
(a) Accuracy (b) Missed Targets

Figure 3.12: Accuracy and Missed Targets for the Visual Search Task. Participants performed with comparable accuracy levels across conditions. However, when inspecting the Missed Targets, i.e., when spending the entire trial duration searching for the target with no selection, we found that VIRTUAL TARGETS showed the lowest amount of Misses.

Behavioral Data

Accuracy First, we analyzed the overall accuracy, see Figure 3.12a. Within this model, the effect of ACTUALITY at the *AV* level was negative but not significant ($\beta = -.12$, 95% CI [-2.46, 2.22], $t(78) = -.10$, $p = .921$; Standardized $\beta = -.03$, 95% CI [-0.58, 0.53]). Similarly, TARGET at the *Virtual* level was positive but without statistical significance ($\beta = .85$, 95% CI [-1.49, 3.19], $t(78) = 0.72$, $p = .472$; Standardized $\beta = .20$, 95% CI [-0.35, 0.76]). Additionally, the interaction effect between ACTUALITY *AV* and TARGET *Virtual* was negative, yet not significant ($\beta = -1.84$, 95% CI [-5.14, 1.47], $t(78) = -1.11$, $p = .272$; Standardized $\beta = -.44$, 95% CI [-1.22, .35]). Participants maintained a consistent level of performance regardless of the ACTUALITY or TARGET.

Missed Targets We analyzed the targets participants missed to select within the 5 seconds of the task, see Figure 3.12b. Within this analytical framework, the effect of ACTUALITY at the *AV* level was found to be negative, but not significant ($\beta = -1.24$, 95% CI [-3.05, .58], $t(78) = -1.36$, $p = .178$; Standardized $\beta = -0.32$, 95% CI [-.80, .15]). Conversely, the effect of TARGET at the *Virtual* level on misses was both significant and negative ($\beta = -4.43$, 95% CI [-6.24, -2.61], $t(78) = -4.86$, $p < .001$; Standardized $\beta = -1.16$, 95% CI [-1.63, -.68]), indicating



(a) Reaction Time (b) IPA

Figure 3.13: Reaction times and IPA results. For reaction times, we only computed correct trials. Here, participants showed faster reaction times when searching for objects in AV searching for VIRTUAL OBJECTS. When inspecting IPA, participants showed decreased IPA, indicative of workload in the AV condition.

a reduction in misses for tasks involving virtual targets. The interaction between ACTUALITY *AV* and TARGET *Virtual* was positive but did not reach statistical significance ($\beta = 0.81$, 95% CI [-1.76, 3.38], $t(78) = 0.63$, $p = .532$; Standardized $\beta = .21$, 95% CI [-.46, .88]).

Reaction Times We analyzed the reaction time, depicted in Figure 3.13a. The model showed a significant effect of ACTUALITY at the *AV* level ($\beta = -106.11$, 95% CI [-149.34, -62.88], $t(8294) = -4.81$, $p < .001$; Standardized $\beta = -.13$, 95% CI [-.18, -.08]), indicating a reduction in reaction times. Concurrently, TARGET also showed a significant negative effect ($\beta = -239.21$, 95% CI [-284.40, -194.02], $t(8294) = -10.38$, $p < .001$; Standardized $\beta = -.29$, 95% CI [-.35, -.24]), suggesting faster reaction times when finding virtual targets. No interaction effects were found ($\beta = -7.78$, 95% CI [-72.23, 56.67], $t(8294) = -.24$, $p = .813$; Standardized $\beta = -9.45e-03$, 95% CI [-0.09, 0.07]).

ERP - Distractor Positivity

Analysis revealed that the effect of ACTUALITY at the *AV* level was significantly negative, with a β coefficient of -3.64 (95% CI [-6.33, -0.95], $t(74) = -2.70$, $p = .009$; Standardized $\beta = -0.32$, 95% CI [-0.55, -0.08]). Conversely, the effect of TARGET at the *Virtual* level did

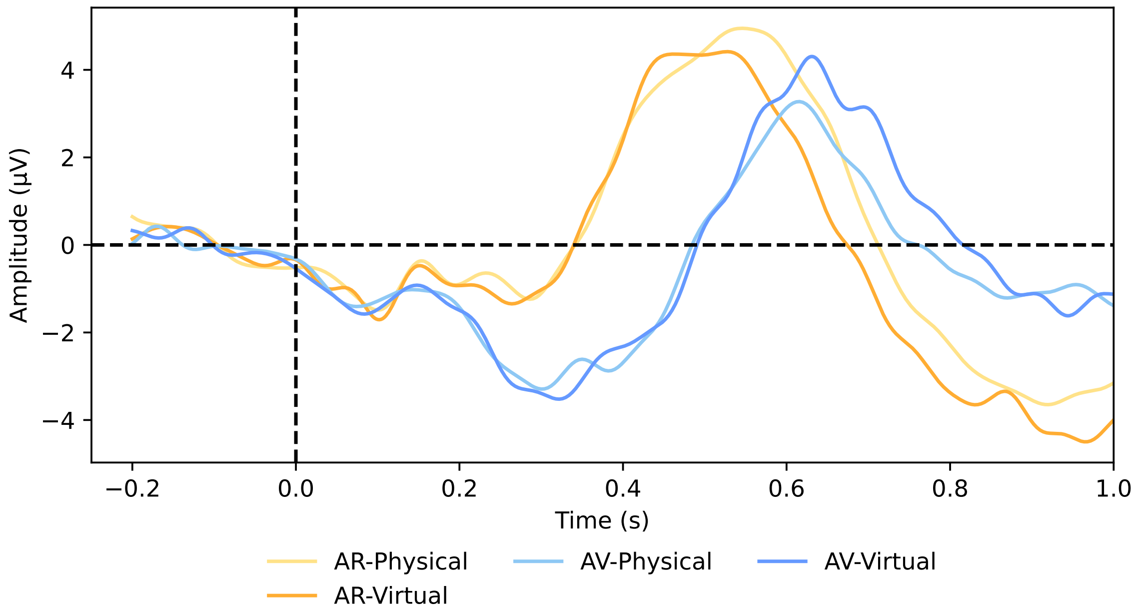


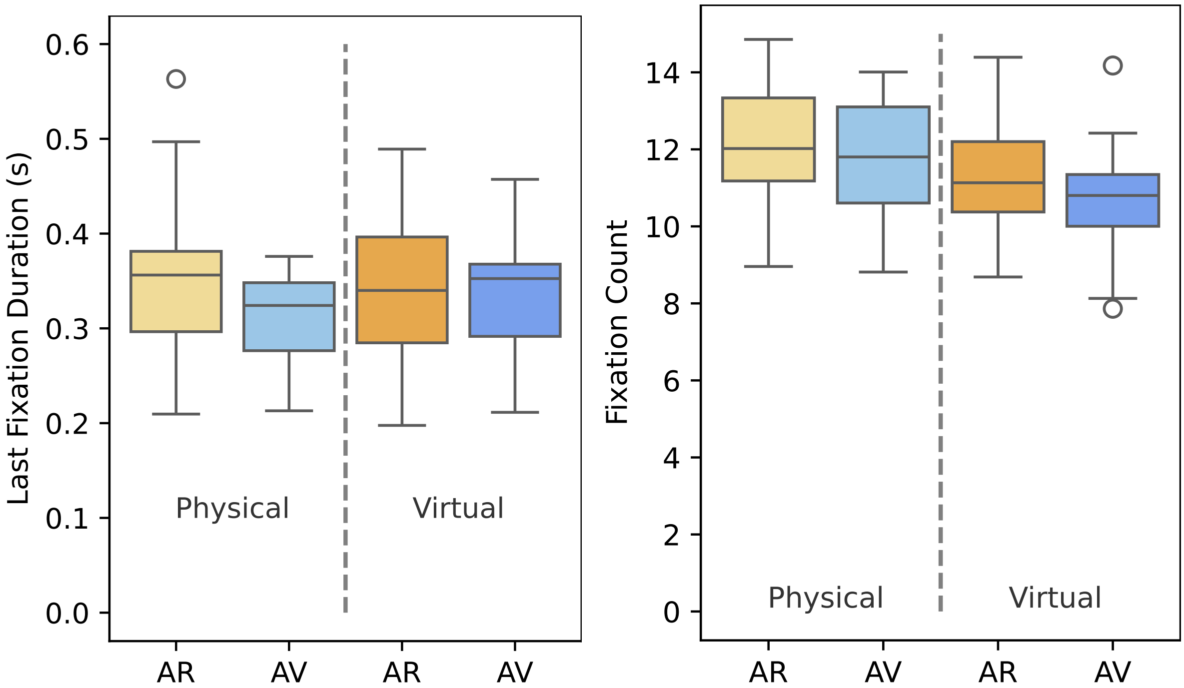
Figure 3.14: Grand Average event-locked to visual search display onset. Data reflect the results obtained from occipital ROI for each ACTUALITY and TARGET condition. The plot suggests a pronounced decrease in peak amplitude on Distractor positivity, with marked variations between AV and AR. No effects of TARGET were found.

not reach statistical significance, exhibiting a negative direction ($\beta = -0.96$, 95% CI [-3.65, 1.73], $t(74) = -0.71$, $p = .478$; Standardized $\beta = -0.08$, 95% CI [-0.32, 0.15]). Additionally, the interaction effect between ACTUALITY *AV* and TARGET *Virtual* was found to be positive but not statistically significant ($\beta = 2.05$, 95% CI [-1.76, 5.85], $t(74) = 1.07$, $p = .287$; Standardized $\beta = 0.18$, 95% CI [-.15, .51]). The grand average for each condition is visualized in Figure 3.14.

Eye Tracking Data

IPA We found a significant decrease in IPA with ACTUALITY AV ($\beta = -0.03$, $95\% CI[-0.05, -0.01]$, $t(7093) = -3.34$, $p < .001$), suggesting a reduction in cognitive load within AV environments, see Figure 3.13b. Conversely, changes associated with TARGET Virtual were not significant ($\beta = .01$, $95\% CI[-.0087, .03]$, $t(7093) = 1.09$, $p = .275$), and the interaction between ACTUALITY AV and TARGET Virtual also did not significantly affect IPA ($\beta = .008$, $95\% CI[-.02, 0.04]$, $t(7093) = .55$, $p = .582$).

Last Fixation Duration The model reported that ACTUALITY [AV] significantly reduces the duration of the last fixation ($\beta = -.04$, $95\% CI[-.05, -.03]$, $t(7093) = -6.08$, $p < .001$; Standardized $\beta = -.19$, $95\% CI[-.25, -.13]$), suggesting a shorter engagement period in AV conditions than AR, see Figure 3.15a. Conversely, the effect TARGET [Virtual] showed a non-significant reduction in fixation duration ($\beta = -.01$, $95\% CI[-.02, 2.54e{-}03]$, $t(7093) = -1.59$,



(a) Last Fixation Duration

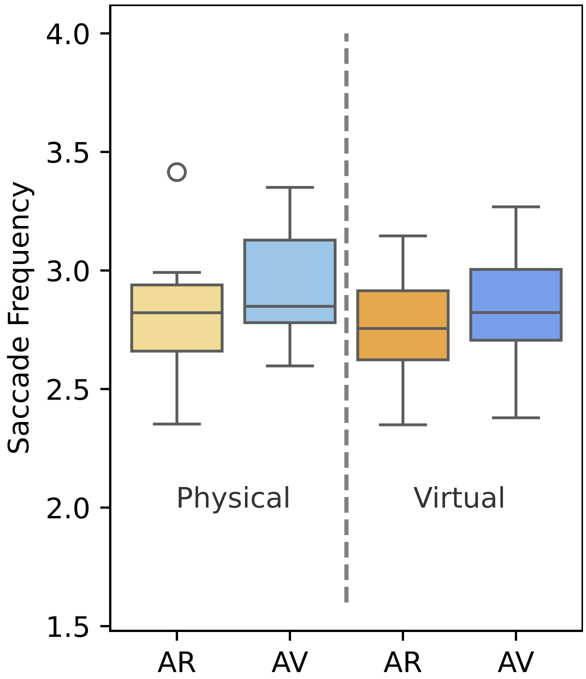
(b) Fixation Count

Figure 3.15: Last Fixation Duration and Fixation Count results. We found a negative significant main effect for AV in Last Fixation Duration (left). For Fixation Duration, we found that participants perform faster fixation in AV and when searching for VIRTUAL TARGETS.

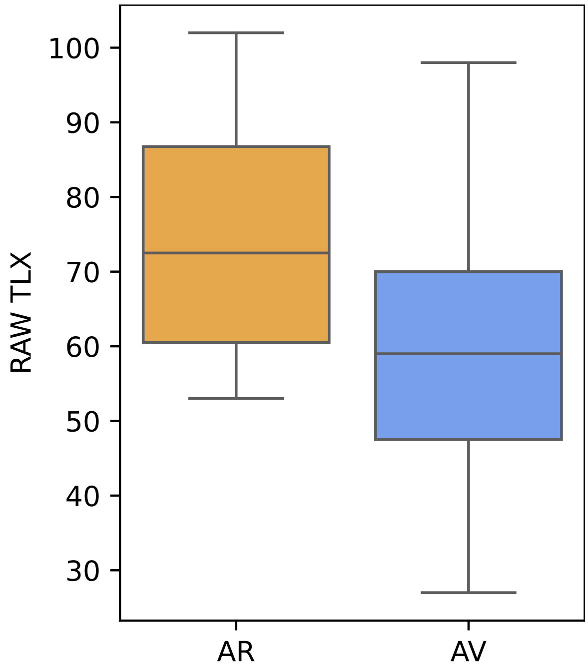
$p = .113$; Standardized $\beta = -0.05$, $95\%CI[-.11, .01]$). Moreover, a significant negative effect was observed in the interaction between ACTUALITY [AV] and TARGET [Virtual] ($\beta = .03$, $95\%CI[.01, .05]$, $t(7093) = 3.40$, $p < .001$; Standardized $\beta = .15$, $95\%CI[.07, .24]$).

Fixation Count The model output showed that ACTUALITY [AV] yielded a significant decrease in fixation count ($\beta = -.35$, $95\%CI[-.57, -.13]$, $t(7093) = -3.09$, $p = .002$). This reduction was paralleled by a significant negative impact of TARGET [Virtual] on fixation counts ($\beta = -.92$, $95\%CI[-1.15, -.69]$, $t(7093) = -7.80$, $p < .001$). However, the interaction effect between ACTUALITY [AV] and TARGET [Virtual] did not significantly alter fixation counts ($\beta = -0.22$, $95\%CI[-.56, .11]$, $t(7093) = -1.33$, $p = .183$).

Saccade Frequency Analysis revealed that ACTUALITY [AV] significantly increased saccade frequency ($\beta = 0.12$, $95\%CI[0.08, 0.15]$, $t(7093) = 6.62$, $p < .001$; Standardized $\beta = 0.20$, $95\%CI[0.14, 0.26]$) Conversely, the introduction of a TARGET [Virtual] was associated with a significant reduction in saccade frequency ($\beta = -.04$, $95\%CI[-0.07, -0.0034]$, $t(7093) = -2.15$, $p = .032$; Standardized $\beta = -0.07$, $95\%CI[-.13, -.0059]$). The interaction between ACTUALITY [AV] and TARGET [Virtual], however, did not significantly influence saccade frequency ($\beta = -.05$, $95\%CI[-.10, .005]$, $t(7093) = -1.75$, $p = .080$; Standardized $\beta = -.08$,



(a) Saccade Frequency



(b) Perceived Workload

Figure 3.16: Saccade Frequency (left) and Raw NASA TLX (right). Saccade frequencies were higher in the AV condition, reflecting a more active search process, but smaller with VIRTUAL TARGETS indexing fast target processing. Results from Raw NASA-TLX show how participants perceived the AR scenario as more demanding than the AV scenario.

$95\%CI[-.17, .009]$).

Perceived Workload

As data showed a not-normal distribution ($W = .923$, $p = .009$), a Wilcoxon signed-rank test was conducted to compare the NASA-TLX scores between AR and AV. When comparing NASA-TLX scores in AR ($M = 74.7$, $SD = 15.28$) to the scores for AV ($M = 58.75$, $SD = 17.85$), we found a significant difference in the NASA-TLX scores ($V = 164$, $p = .029$) with higher perceived workload in the AR condition, see Figure 3.16b.

3.2.5 Discussion

We evaluated the impact of different MR actualities (*AR* and *AV*), and targets (*Physical* and *Virtual*), on behavioral and physiological correlates of visual search efficiency, distractor suppression, and workload.

Impact of Actualities on Visual Search Performance

With our first research question (**RQ1**), we investigate if the ACTUALITY affected the user's performance. The overall accuracy shows a comparable outcome across conditions. On the other hand, we see that the missed targets are significantly higher in the *Physical* conditions. Combining this with our trial design in which the maximum time to search the target is 5000ms, we argue that we see ceiling effects on the overall accuracy. The higher errors in the *Physical* conditions align with the higher reaction time results in the *Physical* conditions. Here, we found a negative main effect of both ACTUALITY and TARGET, showing that participants were faster in target identification in AV and with Virtual Targets. Those results are consistently supported by our results on subjective workload, where participants reported the AV environment to be less demanding. Here, the AR environment's physical fidelity [211] emerges as a potential factor influencing user performance. The inherent visual noise in AR settings potentially distracts and overloads users' cognitive processing capabilities [462], impacting their ability to swiftly and accurately identify targets. In more ecological settings, where controlling the density and arrangement of visual elements is less feasible than in laboratory conditions, the implications of our findings become even more critical. The design of MR environments, especially those intended for real-world applications, must carefully consider how visual fidelity, object complexity, and spatial arrangements affect user performance and cognitive load.

Eye Tracking Correlates of Visual Search Efficiency

Eye tracking correlates of visual search efficiency were the basis for our **RQ2**. Here, we investigate cognitive workload, indexed by IPA, and search patterns by fixations and saccade frequency. We found that cognitive load, as indicated by a decrease in IPA, was reduced in AV suggesting that virtual surroundings facilitate a more efficient cognitive processing experience compared to AR. Furthermore, search patterns, as evidenced by fixation counts and saccade frequency, showed that AV environments and virtual targets support target identification. Specifically, we observed a significant reduction in fixation counts in AV conditions and when interacting with virtual targets, indicating a streamlined search process with fewer distractions. Conversely, saccade frequency increased in AV settings, reflecting more active visual exploration, but decreased for virtual targets, suggesting that once these targets are detected, they require less extensive scanning to process.

Interpreting our results within Guided Search Theory [601], which suggests both top-down and bottom-up mechanisms drive visual search, we can derive that AV may facilitate these mechanisms more effectively than AR. As indicated by IPA, a decrease in cognitive load in AV suggests less demanding attention resource management [281]. Additionally, fewer fixations and adjusted saccade rates in AV and virtual objects point towards a more efficient attention allocation, likely due to clearer cues or more distinct targets enhancing search efficiency [196]. The lowered saccade frequency for virtual targets highlights their ability to quickly draw and keep attention, a sign of effective bottom-up processing [272]. This efficiency could result from the features of virtual elements, which are simple and less noisy than those in

AR, mirroring Guided Search Theory's focus on how stimulus traits direct attention. This suggests AV optimizes visual attributes to engage users better and streamline search tasks.

The findings suggest a trade-off between physical fidelity and optimizing for visual search efficiency. More simple objects, which are easier to identify due to reduced visual noise, align with the principles of effective bottom-up processing by enhancing the saliency of targets. However, the simplification of object design can lead to spatial constraints, such as making objects perceptually closer, potentially complicating the visual search task as the perceptual load increases. This interpretation highlights the importance of carefully balancing the design of MR environments and objects, considering both physical fidelity and the spatial arrangement of stimuli. Thus, this result has implications for the design of MR tasks where visual search efficiency is crucial, such as in training simulations [266], educational content [201], or navigational aids [475]. In scenarios where quick identification and interaction with virtual elements are essential, designing environments and tasks with lower fidelity and diminished visual noise, such as in our AV, could enhance performance. For example, in educational MR applications, presenting virtual objects that are salient and easily distinguishable from the surrounding environment could facilitate learning and information retention. Similarly, in navigational aids, ensuring that virtual indicators or paths are designed to stand out against the real-world backdrop could support wayfinding.

Distractor Suppression in MR Visual Search

In **RQ3** we investigated if distractor suppression varies across ACTUALITIES and for TARGETS. Thus, we focused on Distractor Positivity, an ERP component that reflects suppressive process towards distracting information after visuals search display onset [197]. Our results showed that in AV settings, the effect of actuality on PD was significantly negative, suggesting an efficient allocation of cognitive resources leading to diminished distractor processing. This aligns with the functional significance of PD in visual search: a decreased amplitude usually indicates efficient distractor suppression, resulting in improved behavioral performance through improved target focus [246]. On the other side, in AR, we found an increased Pd positivity as which points towards an increased distractor saliency distractors [500].

Integrating ERP results with the eye tracking ones, where lower IPA denoted higher cognitive efficiency, supports the notion that participants in AV environments experienced a streamlined visual search process. The smaller amplitude of PD, suggesting increased processing of distractors, aligns with the observation of fewer fixations and reduced IPA, indicating an early and effective suppression of irrelevant stimuli facilitated by top-down control mechanisms. The early suppression of distractors facilitates improved attention allocation to targets, thereby resulting in faster reaction times, as reflected in our findings within the AV condition and consistent with previous work [246].

Towards Adaptive Mixed Reality

Our multimodal evaluation allows us to better understand visual search efficiency in MR and consider our metrics as input for an adaptive MR system that can be aware of users' context [351]. Adaptive MR systems, informed by eye-tracking features and Distractor Positivity, can be the foundation for hybrid Brain-Computer Interfaces (Brain Computer Interface (BCI)) or adaptive systems responsive to workload and attention fluctuations. These systems can dynamically modulate the visual nature of stimuli or introduce virtual aids to augment user performance in visual search tasks [96, 352]. By inputting gaze features and ERP components indicative of cognitive effort towards distractors, MR adaptive interfaces can infer interaction intent and future actions in real-time, i.e., identify exploratory visual search behavior or when their attention is diverted from target information.

Here, we envision interfaces that could dynamically vary the saliency of distracting single elements by dimming, blurring, or otherwise de-emphasizing non-essential visual elements [95]. Conversely, targets or necessary information can be highlighted through increased visual saliency or contextual highlighting, thus making them easier to identify and process. This approach aligns with Cheng et al. [96] exploration into leveraging virtual-physical semantic connections to optimize MR layout designs, where the virtual content's placement and appearance are adapted based on contextual relevance and user workload.

Limitations and Future Work

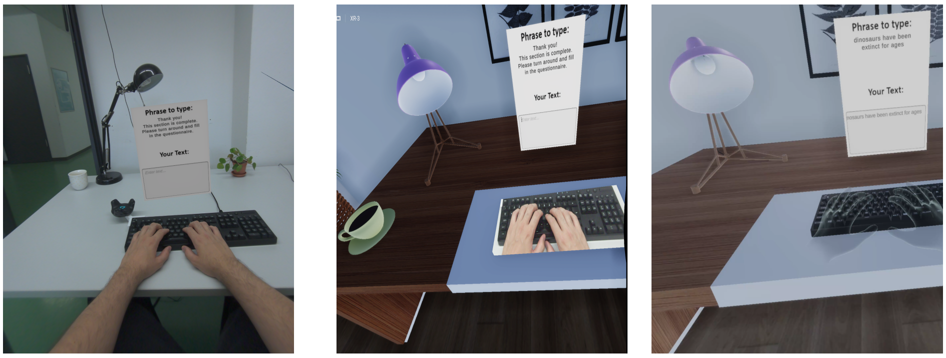
While we manipulated MR actualities and target objects in our study, we acknowledge the limitation of utilizing simplified objects and environments. We made this decision to allow for a controlled experimental setup with stimuli resembling the shape of simple objects. However, our stimuli set and environments do not fully capture the complexity of real-world scenarios. Recognizing this, we propose a replication study to determine if the observed effects persist with real-world objects and their virtual counterparts.

Moreover, to ensure reliable and robust eye-tracking results, we controlled the presented stimuli and environments to ensure consistent luminance. However, with an ecological scenario in mind, we propose investigating varying luminance levels, specifically in AR, to evaluate the effect on Index of Pupillary Activity (IPA) and if such results can be applied in adaptive MR systems.

3.2.6 Open Science

We encourage readers to reproduce and extend our results and analysis methods. Our experimental setup, collected datasets, and analysis scripts are available on the Open Science Framework (OSF)[3].

[3] https://osf.io/zdpty/



Augmented Reality | Augmented Virtuality | Virtual Reality

Figure 3.17: In a user study we investigate physiological correlates for cognitive states and performance during a typing task in three mixed reality environments, thereby varying the degree of immersion.

3.3 Study 3: Engaging in Productive Tasks in Mixed Reality

In our third research question (**RQ3**), we investigate a typing task on a physical keyboard at three different manifestations of the reality-virtuality continuum [390, 520], namely: *AR*, *AV*, *VR*. We lay the groundwork for a new class of transitional interfaces driven by implicit physiological responses rather than explicit user commands. Our investigation into the physiological correlates of attention and engagement across different MR environments provides insights for developing adaptive interfaces that seamlessly transition users through the reality-virtuality continuum based on their cognitive state and support task performance. For this, we conducted a within-subjects lab study ($N = 18$) asking participants to type phrases for 6 minutes per condition. During the typing tasks, we recorded three physiological signals: ECG, EEG, and eye tracking (ET), thereby analyzing physiological responses associated with engagement, attention allocation, and workload levels. We are chiefly interested in exploring distinctions of physiological states between different MR environments and identifying physiological patterns for individual cognitive states during the interaction.

3.3.1 User Study

The primary objective of our study is to systematically examine how different manifestations of the reality-virtuality continuum affect typing performance and if we can observe variation in workload and engagement states as shown by different physiological signals (ECG, EEG, eye tracking). Building on this, we aim to understand how the manifestations influence the physiological markers associated with workload and task engagement. Thus, we conducted a within-subjects experimental design with one independent variable MANIFESTATION on the reality-virtuality continuum (three levels: *AR*, *AV*, and *VR*), see Figure 3.17.

In the *AR* condition, participants typed on a physical keyboard with a view of their surrounding physical environment with typing content displayed over a digital layout. In the *AV*

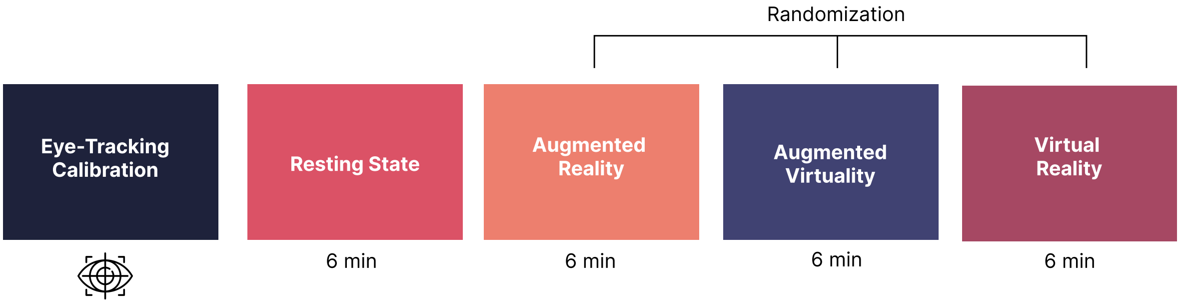


Figure 3.18: *Experiment Procedure.* The experiment encompassed five blocks. In the first block, participants performed a calibrated eye tracker and, in the Resting state block, participants relaxed in the neutral VR environment without distracting elements. Finally, the experimental blocks started, manipulating MANIFESTATION in a randomized order. In between blocks, participants filled in NASA-TLX reporting their perceived workload.

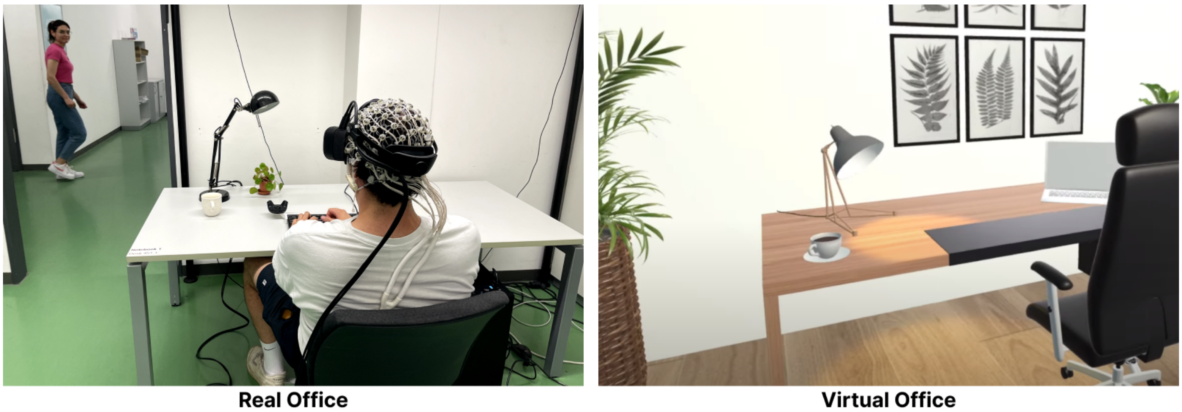


Figure 3.19: During the study, participants were seated at a desk in a simple and clean office space (left). In the VR and AV conditions, a corresponding virtual office was presented (right).

condition, participants typed on a view of the keyboard, and their hands were blended into the virtual environment. And finally, in the VR condition, participants typed on a high-fidelity haptic VR keyboard with virtual hands. To enhance the ecological validity of the study, we conducted the experiment in a real office. We seated participants at a standard office desk facing the door – a typical setup in our institution. As such, the participants will be confronted with real-world visual distractions in the *AR* condition, e.g., bypasses Figure 3.17 left. On the other hand, auditory instructions from the environment are noticeable in all conditions.

Apparatus

We designed the virtual environment for the study in Unity (Version 2021.3.19f1) and presented it via a Varjo XR-3 MR headset with a resolution of $1920 \times 1920px$ per eye for the focus area and $2880 \times 2720px$ per eye for the peripheral area. As per the manufacturer's recommendation, we used three HTC Vive lighthouses 2.0 for environment tracking. For hand-tracking, we

Figure 3.20: Frontal and side view of the office model used for the VR and AV conditions.

employed the Ultraleap SDK[4] for Varjo with a tracking frequency of 200Hz.

In line with prior work such as McGill et al. [380], we used a Logitech G810 Orion with a UK English layout as a physical keyboard. In VR, we rendered an identical VR copy of the keyboard using a model provided by Logitech. We used the Logitech G BridgeSDK[5] to render the key movements (up and down) in VR. Moreover, we tracked the physical keyboard constantly with a VIVE tracker to align and move both keyboards, virtual and physical, simultaneously, see Figure 3.21 left. Finally, we recorded environmental noise from the experimental room via a professional microphone (NT-USB+, RØDE, Australia, 48kHz). We recorded the environmental noise to control it in our statistical modeling, see **??**. Overall, in the AR condition, the average sound level was 64.6 ($SD = 5.16$) dB on average, for the AV condition, the average sound level measured 69.1 ($SD = 8.43$) dB and, in the VR condition, the average sound level recorded was 60.1 dB ($SD = 2.86$).

We acquired four physiological measurements: ECG via PolarH10 chest strap (Polar, Finland, 130 Hz), and EEG signal (LiveAmp Amplifier, BrainProducts, Germany, 500 Hz), and for eye tracking, we employed the built-in system in the Varjo XR-3 (Varjo, Finland, 200 Hz) employing the Tobii XR SDK [6]. Physiological data were streamed within the Unity VR environment within the Lab Streaming Layer (LSL) framework[7] to the acquisition PC (Windows 10, HP Z1 Entry Tower G6, i9, 3,8 GHz, 32GB RAM).

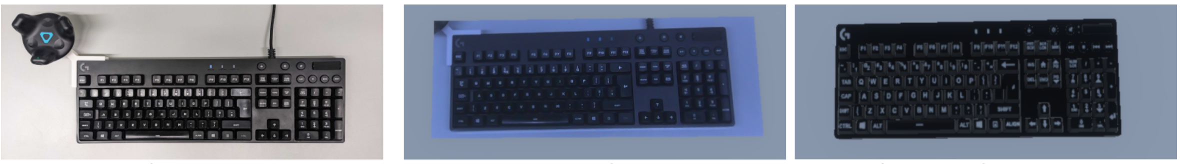
We designed the VR and AV environments for the experiment based on a virtual office; see Figure 3.20. The virtual office design prioritized a realistic office rendering while not evoking negative cognitive and psychological responses [45, 611] and controlling for saliency and its effect on attention allocation and task performance [482]. The virtual office was a room in an apartment. The desk in the room corresponded to the desk in the real world, with biophilic and realistic elements, i.e., plants and office stationery. On the left side of the table

[4]https://docs.ultraleap.com/varjo/

[5]https://github.com/Logitech/logi_bridge_sdk/tree/master

[6]https://developer.tobii.com/xr/

[7]https://github.com/labstreaminglayer/



Physical Keyboard AR Condition | **Augmented Reality Keyboard** AV Condition | **Virtual Reality Keyboard** VR Condition

Figure 3.21: Top view of the keyboards as seen by study participants. We used a UK keyboard layout.

was a floor-to-ceiling window. We maintained the luminance consistent across conditions, following eye-tracking best practices [85, 375] with 180 nits as a comfortable overall fit.

Measures

We evaluated four aspects of the MR-typing interaction: (i) typing performance based [528], (ii) accuracy and time to first key press to gauge how effectively participants could locate the keyboard, (iii) overall perceived workload (raw NASA-TLX [238]), (iv) physiological correlates of task-engagement extracted from EDA (SCL and NSSCR), EEG (HR and HRV), EEG (alpha and theta powers), and Eye-tracking (IPA and saccade frequency). For EEG, we selected one more EEG feature related to workload. We chose the workload index, defined as the ratio between alpha (Pz, P3, P4) [44, 364], and theta, for which we chose frontal channels (F3, F4,Fz, Cz, F7, F8) [319, 364]. Research seems to indicate that task load manipulations are followed by an increase of theta band activity in frontal brain regions, followed by a decrease in alpha power in the parietal areas [409, 468, 498].

Physiological Recording and Processing

In this section, we report the physiological measures collected and the preprocessing stages that allowed us to compute physiological correlates of workload and engagement. We acquired four physiological measures, i.e., EEG, ET, ECG, and EDA. However, due to strong motion artifacts and missing data, the EDA signal delivered by BITalino biomedical toolkit (1000 Hz) [37] was not usable, and thus, we do not report them in this work. For EEG Alpha and Theta and HR we used the Resting state condition for normalization [279].

EEG Recording and Preprocessing EEG data were recorded from 32 Ag-AgCl pin-type passive electrodes mounted over a water-based EEG cap (R-Net, BrainProducts GmbH, Germany) at the following electrode locations: Fp1, Fz, F3, F7, F9, FC5, FC1, C3, T7, CP5, CP1, Pz, P3, P7, P9, O1, Oz, O2, P10, P8, P4, CP2, CP6, T8, C4, Cz, FC2, FC6, F10, F8, F4, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, C1, C5, TP7, CP3, P1, P5, PO7, PO3, Iz, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, F2, AF4, AF8 according to the 10–20 system. One LiveAmp amplifier acquired EEG signals with a sampling rate of 500 Hz. All electrode impedances were kept below ≤ 20 kΩ. We used FCz as an online reference and Fpz as ground. For offline preprocessing, we used MNE Python [216]. We first notch-filtered at 50 Hz to remove power line interference, followed by a band-pass filter between 1-50 Hz to eliminate noise at high and low frequencies. Next, we re-referenced the signal to the common average reference (CAR) and applied the Infomax extended algorithm for Independent Component Analysis (ICA) [342]. We utilized the "ICLabel" MNE plugin [443] for automatic artifact detection and correction.

Eye-Tracking Data Preprocessing Before analysis, we perform outlier removal and filtering of all the collected gaze data. Our data cleaning consisted of several steps: 1) data points with invalid data flags from the eye tracker for either eye were discarded, 2) any data point with an eye velocity over $1000°/s$ or acceleration over $100,000°/s^2$ for either eye was disregarded as the movement is considered physically impossible [255], 3) blinks and accompanying artifacts were removed. Blinks were defined as missing data from the eye tracker, which have durations between 75ms and 500ms. Data 200ms before and after the blinks were removed [30]. 4) the disregarded data was linearly interpolated, 5) data was lastly smoothed with a 6th order Butterworth filter whose cutoff frequency was set at 0.15 Hz [586].

ECG Recording and Preprocessing We collected ECG data at a sampling rate of 130 Hz using a Polar H10 chest strap (Polar, Finland). Before data collection, the ECG electrodes were moistened with lukewarm water and positioned over the xiphoid process of the sternum, just below the chest muscles. Our analysis of the ECG data focused on evaluating HR and HRV in the time domain. For processing the ECG data, we utilized the Neurokit Python Toolbox [367]. Initially, we applied a (Finite Impulse Response (FIR)) band-pass filter with a range of 3 to 45 Hz and a 3rd order to preprocess the ECG signal. Subsequently, we employed Hamilton's method [230] to segment the signal and identify the QRS complexes.

Procedure

Upon the participants' arrival, we explained the study procedure, answered any open questions, and asked participants to give informed consent. The experimenter proceeded to configure the EEG and ECG recording setup. Then, we asked participants to wear the Varjo XR-3 headset and completed a one-point eye-tracking calibration. Next, the experimenter calibrated the hand tracking by aligning the actual hands of the participant with the tracked rendered hands in the AR scene, ensuring that the fingers were aligned with the keyboard. Then, we started the experimental procedure. First, the participants observed a 6-minute resting state, where they sat comfortably in a neutral VR, i.e., Unity skybox, without distracting elements, keeping their hands on their thighs without moving. The resting state is a basal condition for normalizing the experimental conditions. After the Resting State, participants moved to the experimental phase, which consisted of three randomized experimental conditions lasting six minutes each.

The typing task was inspired by previously established work with physical [146, 326] and virtual keyboards in different MR environments [223, 380]. Participants were instructed to type phrases for six minutes during each MR condition, presented in randomized order, see Figure 3.18. At the beginning of every condition, one training phrase was provided to familiarize with the keyboard layout, as in McGill et al. [380]. Drawing from previous work [326, 380], we employed the MacKenzie 500 phrase set. Phrases were randomly chosen and displayed above the text entry field, see Figure 3.17. Participants were instructed to type as accurately as possible. Thus, they had the option but were not obligated, to correct any errors by using the backspace key, aligning with error correction practices outlined in prior work [380, 531]. Once participants completed typing a single sentence, they pressed the "Enter" key to move to the next sentence. Participants typed an average of 41.929 (SD = 19.048) Words per Minute (Words per Minute (WPM)) in the AR condition, 42.08 (SD = 18.226) WPM in the AV condition, and 43.568 (SD=20.834) WPM in VR condition. In between conditions, participants fill out the raw NASA TLX [238] to evaluate the perceived workload. Overall, the experiment lasted 45 minutes.

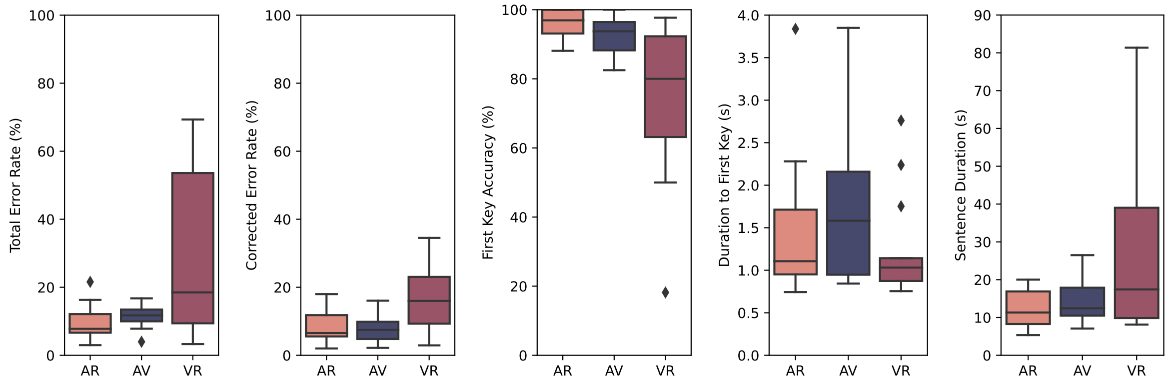
Participants

A sample of 18 participants voluntarily participated in the study. We excluded four participants due to inadequate EEG data quality, as identified by RANSAC, which revealed that over 50% of electrodes of interest for computing alpha and theta frequencies were classified as "bad," compromising data reliability due to a low signal-to-noise ration (Signal-To-Noise-Ratio (SNR)) as described in Bigdely et al. [52]. Thus, we employ a final sample size of 14 participants ($M = 27.9$, $SD = 4.1$; 4 female, 6 male, and 4 diverse). Participants provided written informed consent before their participation. None of the participants reported a history of neurological, psychological, or psychiatric symptoms and all had normal or corrected-to-normal vision. On average, participants reported spending 113 ± 23 hours per week using computers. The average self-assessed expertise level, on a scale from 1 (*novice*) to 10 (*expert*), was 7.2 ± 1.2, in line with previous work [381]. All participants reported prior experience with AR ($M = 4.12$, $SD = 1.11$), AV ($M = 1.33$, $SD = 1.23$), and VR ($M = 5.4$, $SD = 2.1$), rated on a scale from 1 (not at all familiar) to 7 (extremely familiar) as in previous work [96]. The study met the criteria for fast-track conditions set by the local institutional ethics board.

3.3.2 Results

In this section, we first present results on typing metrics and reaction times, HR and EEG correlates of attention and task engagement, i.e., Alpha and Theta powers, EEG index of cognitive workload, alpha-to-theta ratio, and, lastly, eye tracking measures of workload and engagement. We employ a GLMM[8] for all measures on single sentence trials to counteract effects of learning and fatigue [25], as the response in a trial is usually heavily influenced by what happened in the preceding trial. By using mixed-effects models, sources of experimental noise are brought under statistical control. We selected the formula: `measure ~ Manifestation + (1 | participant) + (1 | sound)` for the GLMM. Finally, for

[8]We used a Restricted maximum likelihood (REML) estimation method and Satterthwaite's approximation for degrees of freedom



(a) TER (b) CER (c) First Key Accuracy (d) Duration to First Key (e) Sentence Duration

Figure 3.22: Comparison of typing performance measures across conditions including Total Error Rate (TER), Corrected Error Rate (CER), First Key Accuracy (FKA), Duration to First Key, and Sentence Duration.

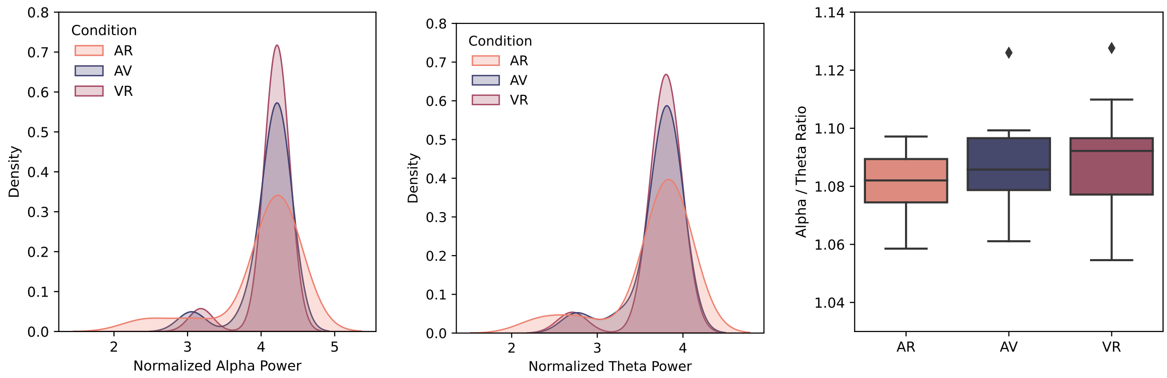
perceived workload evaluated through NASA-TLX, we use one-way repeated measures ANOVA or Friedman's test depending on normality assumption violation as per the Shapiro-Wilk test. We excluded the last sentence typed in each condition as this abruptly ended when the condition duration expired.

Typing Metrics

Total Error Rate (TER) For the Total Error Rate (TER), see Figure 3.22a, we chose the AR condition as an intercept for the GLMM. Compared to this, the AV condition showed a non-significant increase in TER ($beta = .02$, $p = .18$), while in contrast, the VR condition significantly increased the TER ($beta = .08$, $p < .001$). This indicates that participants made more errors in the VR condition compared to the AR and AV conditions. Standardized beta values further emphasize the impact of the VR condition on increasing TER.

Corrected Error Rate (CER) The base CER, see Figure 3.22b, for the AR condition was 0.08 ($p < 0.001$). Comparatively, the AV condition did not significantly alter the CER ($beta = -.002$, $p = .792$), indicating no notable difference from the AR condition. On the other hand, the VR condition significantly increased the CER ($beta = 0.0463, p < 0.001$), highlighting a higher corrected error rate during VR conditions.

First-Key Accuracy When inspecting the accuracy of typing a first key correctly (see Figure 3.22c), the linear mixed model showed a significant main effect. We observe significantly decreased accuracy in the AV and VR conditions (AV: $beta = -0.04, p = .024$; VR: $beta = -0.11$, $p < 0.001$, respectively), indicating lower first key correctness in these conditions compared to AR.



(a) Alpha (b) Theta (c) Alpha Theta Ratio

Figure 3.23: Analysis of participants' EEG data across MRconditions as an indicator of attention allocation, task engagement, and workload. We present alpha and theta powers in a distribution plot while alpha-theta ratio as a boxplot. We report increased alpha and Theta in AV, signaling higher internal attention and task-engagement. A lower alpha theta ratio signals increased workload, which is detected in AR condition (c).

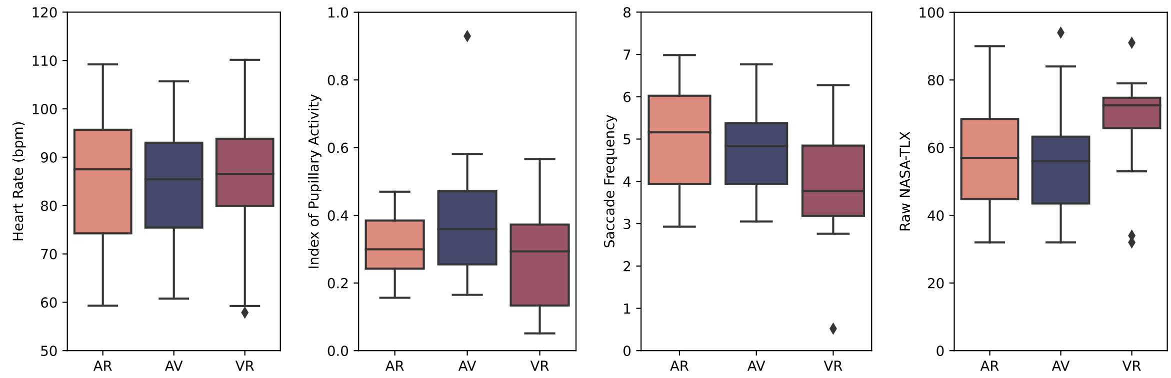
Duration To First Key For the time needed between the start of the trial and the first keystroke (see Figure 3.22d), we found no significant differences across conditions. The AV and VR conditions demonstrated non-significant effects, with beta values of 0.16 and -0.10, respectively, and correspondingly $p = .213$, $p = .473$). This suggests that the MR manifestations (AV and VR) did not significantly alter the participants' readiness or reaction time in initiating the typing task.

Sentence Duration Examining the effects of individual MR conditions, the AV condition showed a non-significant positive effect with a beta value of 1.50 ($p = 0.162$), while the VR condition demonstrated a significant positive effect, with a beta value of 6.08 ($p < .001$). This reveals that participants took longer to type a complete sentence in VR.

Electroencephalogram (EEG)

Alpha power Within the GLMM, the effect of the AV condition was statistically significant and positive, with a beta value of 0.08 ($95\% CI = [.04, .12], p < .001$), see Figure 3.23a. Additionally, the effect of the VR condition was statistically significant and positive, evidenced by a beta value of .07, $p = .003$. The significant positive effects in alpha power for both AV and VR conditions suggest that participants experienced heightened internal attention and engagement in these MR environments.

Theta power Regarding fixed effects, the AV condition was found to have a positive and statistically significant influence on theta power, denoted by a beta value of 0.04, (95% CI [6.05e-03, .08], $p = .022$), see Figure 3.23b. VR condition showed a positive, albeit not statistically significant, with a beta of 0.03, (95% CI [-.01, .07], $p = .202$). These results suggest



(a) Heart Rate (b) IPA (c) Saccade Frequency (d) Perceived Workload

Figure 3.24: Physiological measures from ECG (Heart Rate) and eye tracking (Index of Pupillary Activity - IPA and Saccade Frequency), and Raw NASA-TLX scores across conditions as indicators for workload.

increased engagement in the AV manifestation. Although the VR condition was positive but not significant, it could imply that while the VR environment may have some influence on engagement, it is not as pronounced or consistent as in the AV environment.

Alpha-to-Theta Ratio as a Workload Index When investigating the effects on the EEG workload index, both the AV and VR conditions significantly positively affected the workload index, see Figure 3.23c. Specifically, the AV condition exhibited a beta value of .01, (95% CI [2.07e-03, 0.02], $p = .018$). Similarly, the VR condition demonstrated a beta of .01, (95% CI [1.42e-03, 0.02], $p = .027$). These findings suggest that AV and VR conditions are associated with a decrease in cognitive workload, as indicated by the positive beta values in the EEG workload index, as compared to the AR condition.

Heart Rate (HR) Analyzing the different conditions revealed that both AV and VR conditions had statistically significant negative effects on the heart rate (see Figure 3.24a). Specifically, the AV condition showed a beta value of -1.28 and $p < .001$. Similarly, the VR condition indicated a $\beta = -1.27$ with $p < .001$. This decrease in heart rate could imply that participants were more relaxed or less stressed in AV and VR conditions compared to the AR condition.

Index of Pupillary Activity (IPA) Within the generalized linear mixed model, the AV condition significantly positively affected IPA ($\beta = 0.14$, 95% CI [.05, .23], $p = .002$), as shown in Figure 3.24b. The VR condition, on the other hand, showed a significantly decreased IPA with such significant influence ($\beta = -.04$, 95% CI [-0.07, -0.01] $p = .002$). This indicates that the AV environment may have led to higher cognitive engagement and mental effort, as evidenced by the increased IPA. In contrast, the VR environment showed a significantly decreased IPA compared to the AV condition, suggesting that participants experienced reduced cognitive engagement and mental effort.

Saccade Frequency The analysis of saccade frequency showed a significant main effect of condition (see Figure 3.24c). Specifically, the AV condition was associated with a significant decrease in saccade frequency, with a coefficient of $\beta = -0.26$, (95% CI [-0.38, -0.13], $p < .001$). Conversely, the VR condition demonstrated a pronounced negative effect on saccade frequency, evidenced by $\beta = -0.51$ (95% CI [-0.65, -0.37], $p < .001$). This indicates that both AV and VR conditions led to reduced saccade frequency, with a more pronounced effect in VR, suggesting higher visual focus in these environments compared to the baseline AR condition.

Perceived Workload Shapiro-Wilk testing showed a normal distribution of Raw TLX scores ($W = 0.962, p = .182$). Thus, we conducted a repeated-measures ANOVA to investigate the influence of MR manifestations on perceived workload. The analysis revealed a significant effect of the MR condition ($F(2, 26) = 4.03, p = .029, \eta^2 = .065$). We applied post-hoc pairwise comparisons utilizing Bonferroni-adjusted paired t-tests to discern the differences between conditions. The outcomes indicate no significant distinction between the AR condition and both others (AR AV: $p = 1.0$; AR VR: $p = .220$). However, a significant difference was observed between the AV and VR conditions ($p = .032$), as illustrated in Figure 3.24d. Those results suggest that while there was no significant difference in perceived workload between the AR condition and either the AV or VR conditions, participants perceived a significantly higher workload in the VR condition compared to the AV condition.

3.3.3 Discussion

We conducted a study to investigate the influence of MR environments on typing performance, subjective workload, and physiological correlates of workload and task engagement. We did this to first replicate the results of the seminal work by McGill et al. [380], and extend it by integrating physiological correlates of workload and task engagement varying across MR manifestations. We first summarize our results, then relate our findings and replication of previous work and its relation with physiological workload and task engagement. We then conclude with insights for adaptive transitional interfaces and highlight future fieldwork.

Replication of Typing Performance Across MR Manifestations

McGill et al. [380] reported how augmenting a virtual environment with a real keyboard reduced the error rate and allowed users to better identify overall keyboard position and single keystrokes compared to VR. However, low stereoscopic resolution and latency in VR led to negligible gains in AV. From the typing performance performance, we come to a similar conclusion, as AV showed a decreased error rate while still having lower first-key accuracy. Our results further contribute to the comparison with the AR reality condition. In our setup, the AR condition mimicked real-world MR productivity settings, with augmented display and environmental noise, while still being kept under statistical control. Here, we did not find any significant difference between the AV and AR conditions regarding performance accuracy measures (TER and CER) and the time needed to type a single sentence. This result is confirmed by the reports on perceived workload, where no significant differences were found in the AR-AV contrast but more interesting also in the AR-VR contrast. We argue that here, subjective measures point towards not only an increased workload for VR but also no difference when comparing a condition where the real environment is displayed (AR) and a scenario where only task-relevant objects, i.e., the keyboard and the hands, are displayed (AV).

Physiological Features Highlight Improved Engagement Without Task Overload in AV

Next to replicating previous work, our contribution complements and integrates behavioral results with implicit measures of task engagement, attention allocation, and workload. Physiological correlates of engagement and workload discriminated across conditions, pointing towards an improved task engagement in AV while showing increased workload and external attention allocation in AR. Alpha power increased in the AV condition, indexing either a decreased workload [454] and an internal attention state [44]. Here, first, the AV condition shielded users from environmental distractors while engaging them in a virtual environment that allowed them to focus on the task, i.e., an internal attention state. Internal attention proves beneficial in typing tasks as it facilitates a focused and uninterrupted workflow. This interpretation is supported by the not significant measures in typing measures, where AV promoted comparable accuracy and speed to AR, minimizing the distractions and cognitive load associated with processing external stimuli. Similarly, EEG correlates of engagement, i.e., increased theta power, supporting the increased internal state in the AV condition [120]. On the other side, an opposite direction between alpha and theta is used to compute the Workload index [318, 468]. Here, the AR condition showed increased physiological workload as compared to AV.

The eye-tracking features complement results on physiological engagement and workload derived from EEG. Here, IPA showed higher AV values than AR, while VR did not show significant variations from AR. IPA can be interpreted as a measure of mental effort, and sensitive to mental fatigue; the more fatigue is experienced, the smaller is the IPA [11]. Moreover, recent work proved how constant or increasing IPA changes if participants were positively engaged in solving a task [424] and correlating with EEG-related measures [151]. Moreover, we found reduced saccade frequency in AR compared to AV and VR. Here, reduced saccade frequency in AV and VR can reflect decreased visual exploration, implying that users found it easier to focus on relevant information, thus possibly reducing cognitive load and enhancing task engagement [158, 321]. Conversely, the higher saccade frequency observed in the AR condition might indicate a heightened level of visual exploration and cognitive workload, as users might be trying to integrate and process a higher volume of both virtual and real-world stimuli, which can potentially increase the mental effort and fatigue experienced during the task [527].

Finally, inspecting peripheral physiological stress and arousal measures, HR was increased in the AR condition. AR integrated a more complex blend of real and virtual elements, leading to an increased mental effort and ultimately increasing physiological stress. At the same time, AV allowed for optimal information presentation, and decreasing the visual load did not majorly impact the participants' physiological arousal.

Towards Physiologically-Adaptive Mixed Reality

We ground our work not only in the experimental setting but also in the control theoretic perspective envisioned by McGill et al. [380]. Here, the MR control loop encompasses inputs from reality and feedback from the virtual environment, forming a high-level system analogous to on-screen feedback observed during typing tasks. As users engage more intensively with real interactive elements like keyboards or individuals, these elements can be progressively integrated into the VR view. Our work follows this approach by integrating the physiological computing perspective, which similarly encompasses physiological features processing (i), translation into a system response (ii), and shaping the future or predicted psychophysiological reaction from the user (iii). Our work is an offline study that represents the first necessary building brick for designing physiologically adaptive MR systems to promote task engagement and user productivity.

Our results show that when typing on a physical, real-world keyboard in AV, users have comparable performance but improved engagement as compared to AR. These results can be employed then to promote transitions within the continuum and allow for slowing fading and diminishing reality [95]. Such transition needs to be context-aware and consider multivariate patterns in the user. Consequently, we adopt a sensor fusion strategy to gauge cognitive load and engagement to offer reliable real-time results. The features we computed have already been employed in previous adaptive systems in AR for IPA [352], ECG in VR [291], and EEG [145] using real-time estimation of different states. We propose to investigate this in the future and expand our framework with sensor fusion approaches and multimodal user state analysis.

Limitation and Future Work

Our study, though examining the interactions between cognitive load, and engagement, across behavioral subjective, and physiological variables in MR environments, recognizes certain limitations and foresees promising future work in the field of MR for productivity.

While ecologically valid, the typing task utilized in this study might not fully include the many variables of primary office tasks. Future investigations could thus encompass more complex tasks, such as creating slide presentations or coding, to better simulate real-world office environments. This also applies to interaction with other agents, either being human in MR [380] or virtual avatars in VR and AR [108, 261]. Further, we evaluate merely one of several possible text-input techniques, namely typing, as this remains the predominant modality in day-to-day office work. However, MR technology may well bring about a paradigm change for text-input, fueled by further research on alternative modalities (e.g., eye-tracking or voice-based).

Second, while we controlled the visual load and embedded the task at a real office level to maintain ecological validity, VR and AV environments might allow larger opportunities to isolate participants better from potential distractions. This invites further research to evaluate the effects of manipulating sound environments and their congruence with the visual presentation of AV and VR environments.

Finally, in order to develop and evaluate a working MR adaptive system, we need to consider how transitions are accepted by users and their effect on attention and workload. In previous work in VR, users seemed to prefer simple and short transitions [182]. However, it is not clear if the same either applies across MR manifestations and if a sudden change in sensory information, either by enriching or diminishing the environment, might impact physiological arousal and engagement. Thus, we propose to evaluate different design choices for transitions in MR and which "direction" across the continuum might better support users' goals in the individual task.

3.3.4 Open Science

We encourage readers to review, reproduce, and extend our results and analysis methods. To achieve this goal, we make available our collected datasets, sentence stimuli, MR environments, and experiment setup and analysis scripts at this link `https://osf.io/juk9x/`.

3.4 Summary

In this first chapter, we explored attention and engagement states across the MR continuum. Our studies collectively address how users interact and identify virtual and physical information within AR, AV, and VR environments.

In section 3.1, we investigate the attentional demands of processing virtual and physical cues in AR and AV, highlighting the increased cognitive load and visual complexity encountered in AR settings. Participants demonstrated more efficient object identification and reduced perceived workload in AV, suggesting that the fidelity of the virtual environment significantly influences user performance and cognitive strain. Here, our findings carry implications for understanding how users allocate attentional resources in an MR environment, suggesting that optimizing attentional resources should be a key consideration across the MR continuum.

In section 3.2, we investigated the visual search efficiency across the MR continuum by focusing on suppressing distracting information and processing target information in AR and AV environments. Our findings indicate that distractor suppression is more efficient in AV, with eye-tracking results revealing less scattered and cognitively demanding visual search patterns than AR. This research identifies the interplay between virtuality and information physicality must be considered when designing interactive MR systems.

Lastly, in section 3.3, we emphasize how attention allocation and task engagement vary across the Mixed Reality (MR) continuum during a typing task, as their physiological correlates show. We found that the best typing performance was verified in AR and AV, with physiological engagement peaking in AV, where the workload also decreased. This suggests that AV offers a balanced integration of real-world elements (such as a physical keyboard for typing) with the benefits of MR environments (such as reduced real-world distractions). These findings imply that AV can effectively support typing performance while enhancing user engagement without the cognitive strain typically associated with high-workload tasks. These insights collectively highlight the critical role of virtual content moderation in MR systems. The amount of virtual information presented can significantly influence user attention and engagement. The balance of virtual and physical elements within an MR setting is, therefore, a driving factor for maintaining optimal engagement and cognitive performance.

4

PHYSIOLOGICAL COMPUTING FOR ADAPTIVE VIRTUAL REALITY

> *"All fixed set patterns are incapable of adaptability or pliability. The truth is outside of all fixed patterns."*
>
> **– Bruce Lee.**

In the chapter, we investigate the feasibility and implications of leveraging physiological signals—specifically, correlates of engagement and attention allocation—to inform the design of adaptive VR systems. This inquiry is articulated through a series of focused studies, each addressing a specific aspect of the overarching question: Can we support user performance by considering physiological correlates of engagement and attention allocation as input for a VR adaptive system?

We first explore secondary task difficulty adjustments based on physiological arousal markers [110], i.e. EDA. Here, we find that when grounded in users' engagement levels, such adaptations can significantly enhance the VR experience by diminishing workload. The system's dynamic response to engagement maintains user engagement and balances dual-task demands.

Furthering this exploration, we shift our attention to the effects of these difficulty adjustments on attention allocation mechanisms [103]. Here, we adopt a multimodal analysis approach by investigating EEG, ECG, and EDA correlates of adaptive task difficulty and their effect on physiological markers of cognitive load, engagement, and attention allocation. These findings suggest specific EEG frequencies are sensitive to such adaptations and can be utilized to balance user attention and cognitive engagement.

We then shift our focus from task-relevant adaptations to manipulating environmental visual complexity on task performance [108]. Here, we employ EDA but enhance the signal processing by not focusing on the raw signal but employing tonic features associated with task engagement. Our approach aims to balance providing rich, immersive environments and not overwhelming users with excessive visual stimuli. By tuning visual complexity based on engagement metrics, VR experiences can be tailored to support optimal performance.

Again, we adopt a multimodal approach by investigating the effect of adapting environmental visual complexity on physiological correlates on attention allocation, engagement, and workload [104]. This analysis highlights the critical role of visual elements in shaping cognitive load and attentional focus within VR settings. It points towards the potential of adaptive systems to enhance user interaction by intelligently modifying environmental cues based on attentional feedback using EEG features.

Finally, the chapter presents the design and evaluation of an adaptive system that considers EEG correlates of attention allocation to support task performance [111]. This innovative approach marks a significant step towards creating VR environments that respond to and anticipate user attentional focus, facilitating a more personalized and effective interaction.

This chapter is based on the following publications.

Francesco Chiossi, Robin Welsch, Steeven Villa, Lewis L. Chuang, and Sven Mayer. 2022. Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience. In *Big Data and Cognitive Computing*, 6(2), 55.
https://doi.org/10.3390/bdcc6020055

Francesco Chiossi, Changkun Ou, and Sven Mayer. 2023. Exploring Physiological Correlates of Visual Complexity Adaptation: Insights from EDA, ECG, and EEG Data for Adaptation Evaluation in VR Adaptive Systems. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA.
https://doi.org/10.1145/3544549.3585624

Francesco Chiossi, Yagiz Turgut, Robin Welsch, and Sven Mayer. 2023. Adapting Visual Complexity Based on Electrodermal Activity Improves Working Memory Performance in Virtual Reality. In *Proceedings of the ACM on Human-Computer Interaction*, 7(MHCI), 1-23. ACM, New York, NY, USA.
https://doi.org/10.1145/3604243

Francesco Chiossi, Changkun Ou, and Sven Mayer. 2024. Optimizing Visual Complexity for Physiologically-Adaptive VR Systems: Evaluating a Multimodal Dataset Using EDA, ECG, and EEG Features. In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces*. ACM, New York, NY, USA.
https://doi.org/10.1145/3656650.3656657

Francesco Chiossi, Changkun Ou, Carolina Gerhardt, Felix Putze, and Sven Mayer. 2024. Designing and Evaluating an Adaptive Virtual Reality System Using EEG Frequencies to Balance Internal and External Attention States. In *International Journal of Human-Computer Studies*, 178, 103433.
https://doi.org/10.1016/j.ijhcs.2024.103433

4.1 Study 4: Adapting Secondary Task Difficulty based on Physiological Arousal

4.1.1 User Study

We conducted a study to evaluate whether our physiologically adaptive system can support users' comfort and usability in the virtual environment. To do so, participants performed a

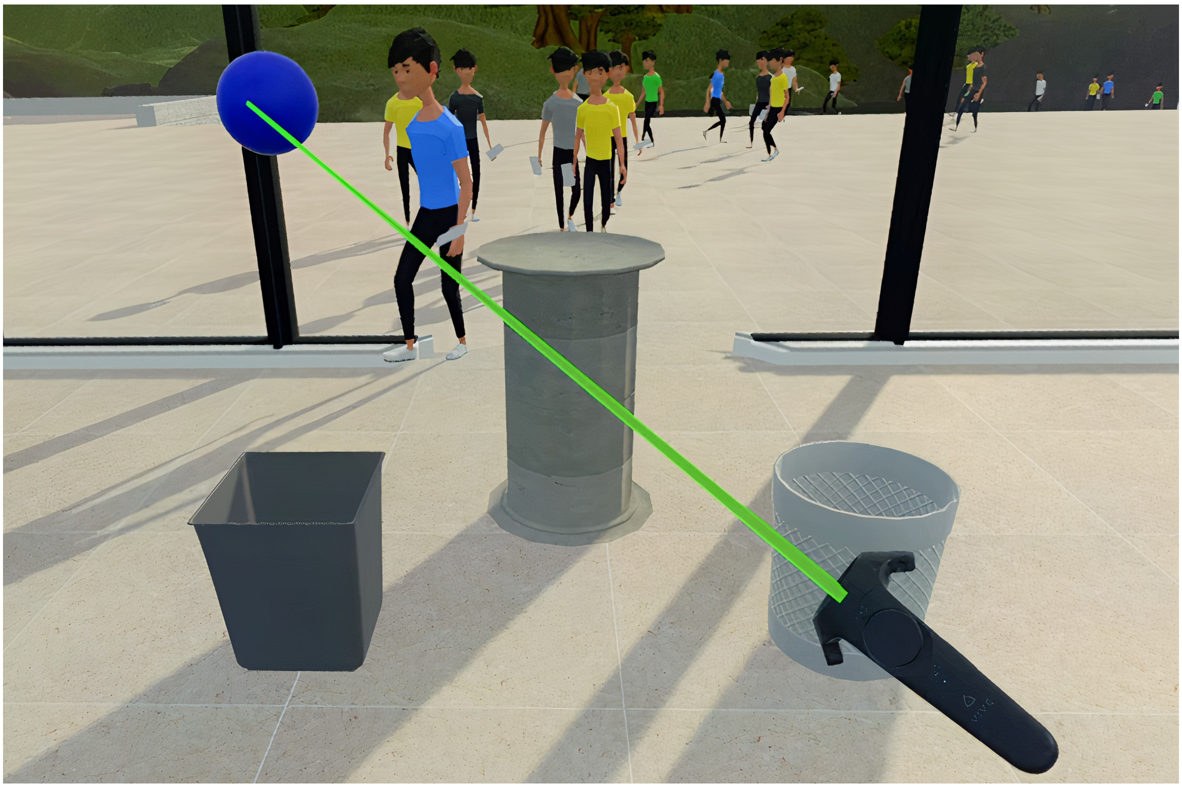


Figure 4.1: Game view capture of a single trial of the VR n-back ($n = 1$) and the visual detection tasks. Participants were required to place a sphere into the corresponding bucket. If the sphere matched the color of the previous sphere one step before, participants placed it into the right bucket. If not, the sphere had to be placed into the left bucket. The visual detection task required participants to monitor if visitors of a museum either possessed a ticket to enter the building or not. To signal a missing ticket after detection, the participant had to select the NPC.

visual working memory n-back task (primary task) [308] and a visual detection task (secondary task). Both tasks were simulated as real-world activities in a virtual environment. While the primary task simulates a high workload task, the secondary task simulates situations where the user is distracted by, for instance, other players [345], notifications [490], or graphical artifacts [511].

In the n-back (i.e., 1-back) task, participants determined if the color of a presented sphere matched the color of the last presented sphere, see Figure 4.1. Participants were presented with colored spheres that they had to place in either a left or right bucket depending on their color (mis)match with the previous sphere. Spheres were either green, red, blue, or black ([382, 467]). A sound cued the appearance of each sphere. Then, the participant had to pick up the sphere within 4 s; otherwise, it would count as an error. New spheres would appear either after 4 s or when the current sphere was placed into one of the two buckets. We counted missing a sphere as an error.

In the secondary tasks, we asked participants to inspect the museum tickets of a STREAM of virtual NPCs. In detail, participants were required to distinguish museum visitors who had a ticket either in their left or right hand from those without a ticket. Museum visitors were represented by virtual NPCs that were walking past the participant on either the left or right side. We set the percentage of virtual NPCs who have no ticket to be 15%, which we found to be a suitable value in the informal pilot tests for this task. These had to be identified by clicking on them with the right input controller. All NPCs had randomized visual characteristics (i.e., hair, shirt color), and their shirt color turned red when clicked on. The NPCs' distance from the user was kept at a minimum of 2.5 m to avoid proxemics confounds with the level of arousal of participants [353].

The five non-adaptive conditions had fixed STREAMS of 7, 22, 37, 52, or 67 NPCs per minute entering. In the adaptive condition, the system adapted the number of NPCs and added more NPCs when the physiological activity of the user did not indicate physiological arousal or removed NPCs. We randomized the six test conditions. Participants were not aware whether they would experience the adaptive condition or one of the non-adaptive conditions.

Participants

Eighteen volunteers ($M_{range} = 23 - 31$; $M_{age} = 27.9$, $SD_{age} = 2.9$; $male = 9$, $female = 9$) participated in our study, of which we had to exclude three from the analysis as the EDA electrodes lost contact. We recruited the participants using our institutional mailing lists and social networks and using convenient sampling. Exclusion criteria required participants to not experience intense physical activity or consume any caffeine or nicotine in the 3-hour pre-study period [26]. None of the participants reported a history of neurological, psychological, or psychiatric symptoms. They were also required to submit a negative test for SARS-CoV-2 within 48 hours prior to participation.

4.1.2 Online Physiologically-Adaptive System

Here, we propose an online adaptation of virtual environments based on EDA. More specifically, we developed a real-time physiologically adaptive system based on the user's arousal as indexed by EDA. This system determines if its user can handle more visual complexity or reduce complexity to minimize cognitive load. Consistent with prior literature [533], the online signal processing pipeline derives a mean EDA value by applying a 20-s median moving-average filter that replaced each data point with the average of its 20 neighboring data points.

To enable the user-dependent adaptation of the system, the physiologically-adaptive system is initialized with a baseline recording $b_0 \cdots b_i$. For later baseline comparison, we calculate Δb as follows[1]:

$$\Delta b = \overline{b_{j-i} \cdots b_i} - \overline{b_0 \cdots b_i}, \tag{4.1}$$

[1]We denoted the mean value as $\overline{x}$

where i is *sampleRate* $*$ *windowBaseline* and j is *sampleRate* $*$ *timeBaseline*. The baseline period is used to compute the baseline slope. In our implementation, this is the slope between the averaged values of the first and last 20 s over the three minutes using the aforementioned signal processing pipeline.

$$\Delta s = \overline{s_{j-i} \cdots s_i} - \overline{s_0 \cdots s_i}, \tag{4.2}$$

where i is *sampleRate* $*$ *windowOnline* and j is $-$*sampleRate* $*$ *window*. Thus, this determines the current slope of physiological arousal. Due to the other external influences, such as baseline arousal of the virtual environment, this alone is not sufficient for adaptation.

In the final step to determine how to adapt the virtual environment, we compare the baseline measurement Δb to the online measurement Δs. The difference then determines how to adapt the environment, which we define as follows:

$$adaptation(\Delta s) = \begin{cases} \texttt{increase of Stream} & \text{if } \Delta s \leq \Delta b - t \\ \texttt{decrease of Stream} & \text{if } \Delta s \geq \Delta b + t \end{cases} \tag{4.3}$$

where t is the threshold parameter that enables the physiological-adaptation system to work with a delay, counteracting high-frequency changes in EDA and, thus, preventing rapid and unstable adaptations. Moreover, for a stable adaptation, Equation (4.3) is only used for adaptation every 20 s.

Apparatus

We designed the virtual environment for the study in Unity (Version 2019.4.24f1), see Figure 4.1, and presented it via an HTC VIVE VR headset with a display resolution of 2160×1200 pixels, refresh rate of 90Hz, and an average field of view of 110°. The environment used for the experiment is a replica of the Neue Nationalgalerie in Berlin, Germany. The 1-back task takes place in the entrance of the building, see Figure 4.1 and, for the detection task, agents with and without tickets in their hands entered the entrance.

We performed the EDA data recording according to the guidelines by Babaei et al. [26]. We placed Ag/AgCl electrodes (7 mm surface diameter) on the participant's non-dominant hand (inner distal phalanges of the index and middle fingers). An electrolyte solution (0.5% NaCl) was applied to the acquisition site to ensure proper hydration and to minimize the effect of individual differences. We attached the electrodes 10 min later, with double-sided adhesive collars. We exploited the exosomatic recording principle with a direct current (DC) of a constant 0.5 V voltage.

We used Equation (4.3) to adjust the number of NPCs in the adaptive condition. In detail, when the EDA slope computed in the 20-s window was greater than the baseline slope added to the threshold slope, two NPCs were removed from the scene. On the contrary, if the EDA slope was lower, four NPCs were additionally spawned. A fixed range of 7–67 NPCs was maintained regardless of the participants' measured physiological activity. These settings were determined empirically during a number of test sessions.

Measurements

We recorded three physiological measurements: (1) EDA signal via the LiveAmp amplifier (BrainProducts GmbH, Germany), using a 250Hz sampling rate; (2) ECG (Polar H10 chest strap, Finland) at 130 Hz; and (3) EEG signal (DSI-VR 300, Wearable Sensing, San Diego, CA, USA) at 250 Hz. Physiological data were streamed and recorded within the Lab Streaming Layer framework (LSL). We only used the EDA signal as an indicator for physiological arousal. In addition, performance accuracy metrics were computed for both tasks. For the 1-back task, errors were represented by the proportion of times the sphere was placed in the wrong bucket. For the visual detection task, errors were represented by missing an Not-Player Character (NPC) or selecting an NPC with a ticket.

Responses on three standardized questionnaires and two custom Likert items were collected to evaluate user experience and workload. First, raw subjective workload measures (NASA-TLX [237, 238]). Second, perceived gamefulness of system use (Game Experience Questionnaire (in-Game Core Module) [270]). Here, we only recorded the subscales on Competence, Immersion, and Positive Affection allowing for content validity [336]. Third, the Fast Motion Sickness scale (FMS) to control for motion sickness [304]. Finally, participants rated two general usability statements on a 5-point Likert scale (strongly disagree - strongly agree); questions: "I would like to use the system in the future," and "The flow of the people was appropriate."

Procedure

Upon arrival, we briefed the participants on the study procedure, and we answered any open questions, which were followed by signing the informed consent form. Next, they performed a Snellen visual acuity test. Finally, the EDA sensor, EEG-VR headset, and ECG chest strap were attached to the participants.

The study began with a trial phase to allow participants to familiarize themselves with the VR environment. The VR trial phase started with participants practicing the 1-back task until they reached an accuracy level of at least 95% within a sequence of 80 spheres. Next, three minutes of EDA recording was performed, during which the participant was asked to ignore the NPCs (STREAM: 37 agents/min) and to perform the 1-back task. This provided the baseline measurement for the physiologically-adaptive system.

The testing phase consisted of six blocks for each test condition. Each block lasted for six minutes. Visual feedback for the 1-back task was provided, and participants were instructed to maintain 90% performance. In between blocks, participants filled in questionnaires (FMS, NASA-TLX, In-game Game Experience Questionnaire (Game Experience Questionnaire (GEQ)) subscales, ad-hoc questionnaires) on the previous block and rested for 2 min to stabilize their

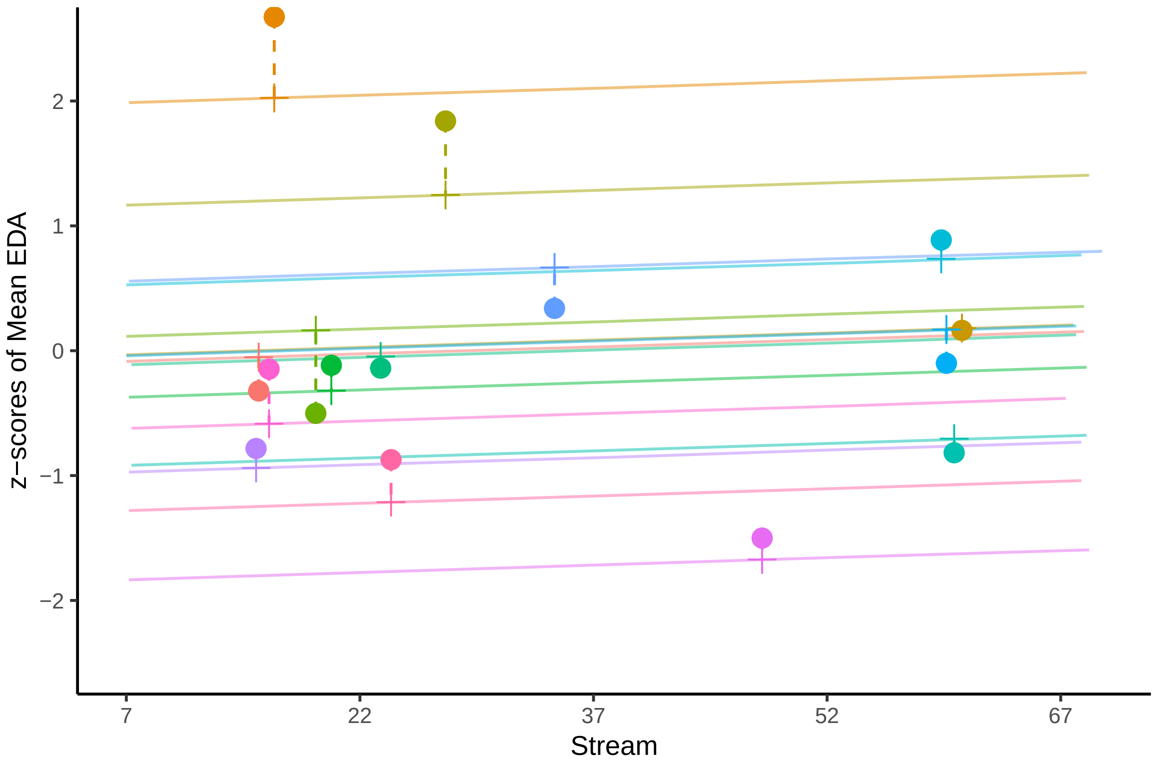


Figure 4.2: Individual predicted standardized mean EDA from the optimal STREAM for the non-adaptive condition (crosses) with individual regression lines, as well as the actual mean EDA (points) at local maxima of adaptation. Each line represents a participant.

physiological state. The entire experiment lasted approximately one hour. We compensated participants with 10 Euros for their participation.

4.1.3 Results

We evaluated the usability of a physiologically adaptive system in a Social VR scenario. In the following section, we report the results of our study. As no participant aborted due to simulator sickness, however, electrodes were not attached correctly for three participants; the following analysis is based on the data of 15 participants. Results of the analyses can be found online on the DaRUS Open Data Platform, at [112].

We analyzed indicators of physiological arousal, performance, and subjective experience across the adaptive and non-adaptive conditions in R [129, 460, 593]. First, we present the results for the five non-physiologically adaptive conditions and present a comparative analysis of the adaptive condition to determine if the physiologically adaptive condition produces: (1) superior performance; (2) a more enjoyable game experience; and (3) lower levels of perceived workload. We fit a linear mixed model (estimated using a restricted maximum likelihood approach and nloptwrap optimizer) [329, 366] to predict our outcomes as a function

Table 4.1: Means across non-adaptive conditions with the slope of GLMM and their *t*-values estimated by Wald approximation, as well as their respective *p*-value.

	Stream 7		**Stream 22**		**Stream 37**		**Stream 52**		**Stream 67**		**LMM**		
	M	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***	***b***	***t***	***p***
n-back Acc. [%].	97.423	1.607	95.381	2.660	94.401	2.430	92.133	3.182	91.055	4.162	−.107	−9.396	<.001
Visual Det. Acc. [%]	97.208	3.073	95.579	4.647	94.834	2.670	93.067	3.116	9.949	2.613	−.100	−7.083	<.001
Raw NASA TLX	5.267	2.520	7.644	2.268	9.500	3.616	9.133	3.282	11.833	3.262	.097	8.101	<.001
EDA Mean (std)	−.091	1.031	−.083	1.001	−.034	.896	.003	1.083	.166	1.056	.004	2.536	.014
SCL	9.078	2.974	9.102	2.888	9.242	2.585	9.348	3.127	9.819	3.047	.012	2.535	.014
GEQ—Competence	2.667	.939	2.600	1.039	2.200	.862	2.600	.687	1.967	.834	−.009	−2.149	.036
GEQ—Pos. Affec.	2.400	.806	2.500	.655	2.467	.550	2.433	.729	2.100	.761	−.004	−1.318	.193
GEQ—Immersion	.833	.724	1.200	.841	1.133	.876	1.067	.821	1.167	.939	.004	.906	.369
Stream Appropriate	1.400	.632	1.600	.737	2.000	.756	1.933	.799	1.467	.743	.003	.754	.454
Desire To Use	1.467	.640	1.800	.775	1.800	.941	2.267	.799	1.333	.617	.001	.309	.759

of STREAM. We computed the *p*-values using the Wald-Approximation for the calculation of degrees of freedom. To account for the repeated-measures structure in our data, we added a random intercept for every participant to our model. We use Welsh-corrected *t*-tests for comparing means or Wilcoxon-rank tests for ranks.

Non-Adaptive Conditions

Electrodermal Activity For mean EDA, the model intercept is at −0.16 (95% CI [−0.67, 0.36], $t(71) = -0.60$, $p = 0.551$). Within this model, the effect of STREAM is statistically significant and positive (beta = 0.004, 95% CI [0.00, 0.007], $t(71) = 2.54$, $p < 0.01$. In other words, every additional NPC increases mean EDA by about 0.004 standardized µS units 4.2. This is consistent with the raw NASA-TLX score, see Table 4.1. Increasing the number of NPCs in the simulation elevated the NASA-TLX score by about 0.1 per NPC.

In contrast, STREAM affects performance accuracy negatively in both tasks, see again Table 4.1. Increasing STREAM reduces accuracy up to 10% in both the n-Back and the visual detection task. Together with this, we supplement our analysis with a SCL computation as a measure of tonic activity. To compute SCL, EDA data were filtered with a 3Hz, high-pass, fourth-order Butterworth filter to remove high-frequency noise and decomposed into tonic and phasic components by means of a non-negative deconvolution analysis [43]. For SCL, the model intercept is at 8.89 (95% CI[7.40, 10.39], $t(71) = 11.86$, p < 0.001). The effect of STREAM is statistically significant and positive (beta = 0.01, 95% CI [0.00, 0.02], $t(71) = 2.54$, $p < 0.013$. This result implies that the addition of every NPC increases SCL by about 0.01 standardized µS units. We found that the subjective experience scales (GEQ Immersion, GEQ Competence, GEQ Positive affect, desire to use, and Appropriateness of Stream) are not affected by our manipulation for all $p > 0.05$, see Table 4.1.

Adaptive Condition

On average, participants adapted to having a STREAM of 32.89 (*SD* $= 18.6$) NPCs in their environment. This represents a medium level of STREAM, relative to the fixed range (i.e.,

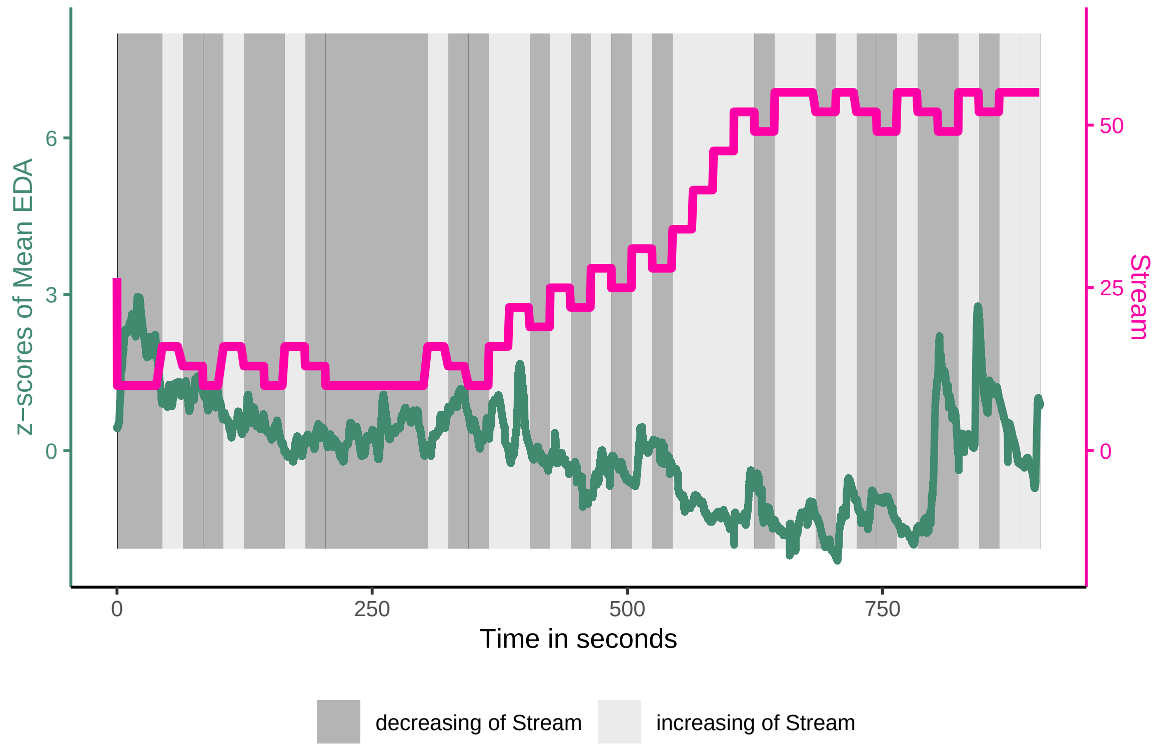


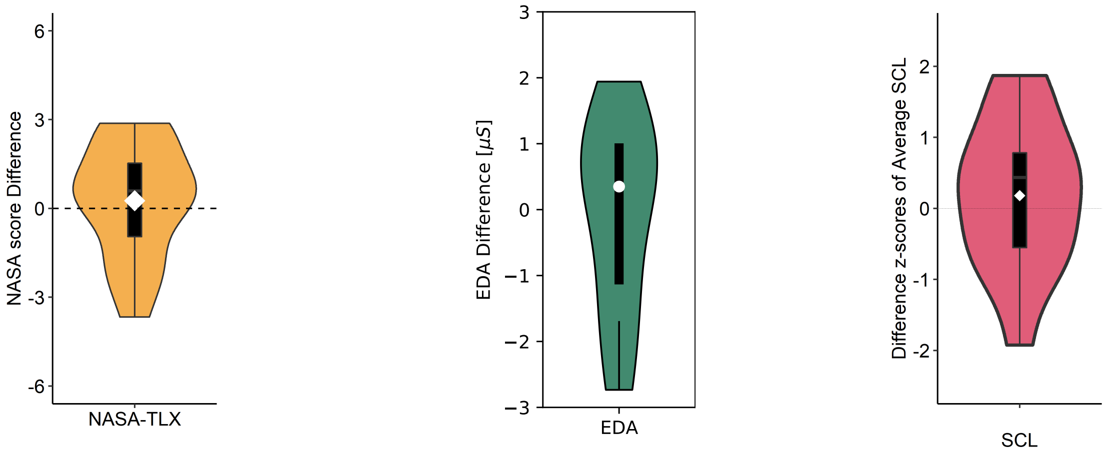
Figure 4.3: Adaptation across time for one participant. The pink line indicates STREAM, the green line indicates the z-scored mean EDA signal that was used for adaptation. Grey areas indicate whether the algorithm chose to increase (light grey) or decrease (dark grey) the STREAM in a time window of 20 s.

7-67) that was adopted for the non-adaptive conditions. Figure 4.3 illustrates how STREAM adapted over time for one participant.

The recorded measures from the non-adaptive conditions cannot be directly compared to their counterpart in the physiologically adaptive condition. This is because STREAM was a fixed value in the former and an adapted value in the latter. Thus, we computed individual expected values from the five non-adaptive conditions using linear mixed models with STREAM as a predictor and calculated the expected value as predicted from our statistical model of the non-adaptive conditions, see Table 4.2 in the *Prediction* column and at the local maxima in STREAM for the adaptive condition.

In other words, any aspects of the user experience that are different in the physiologically adaptive condition compared to the non-adaptive conditions would be revealed if the actual measurement in the condition is significantly higher or lower than the corresponding value that would be predicted from a regression of the measurements of the non-adaptive conditions.

Electrodermal Activity Measures The physiologically adaptive condition did not significantly increase physiological arousal in either the EDA mean or SCL according to the linear mixed model. Individual regression predicted a mean EDA of -0.023 µS ($SD = 0.957$), while in



(a) NASA (b) EDA (c) SCL

Figure 4.4: The relative difference for (**a**) raw NASA-TLX score difference, (**b**) standardized mean EDA, and (**c**) averaged SCL scores.

Table 4.2: Mean predicted and actual means across measures with Welsh-corrected *t*-test or Wilcoxon-signed-rank test depending on the Shapiro test and Cohen's d. We compare the predicted value from the LMM at an optimized Stream with the actual value at this rate of Stream.

	Actual		Prediction		Diff		*t*-Test/Wilcoxon *			Shapiro	
	M	*SD*	*M*	*SD*	*M*	*SD*	d	*t*/*Z*	*p*	*W*	*p*
n-back Acc. [%]	94.892	2.662	94.471	2.007	.422	1.801	.234	93.000	.064	.825	.008
Visual Det. task Acc. [%]	95.478	3.837	94.696	2.766	.781	1.706	.458	1.774	.098	.944	.440
Raw NASA	8.578	3.205	8.317	2.250	.261	1.952	.134	.518	.613	.938	.363
EDA Mean (std)	.040	1.074	−.023	.957	.063	.364	.172	.668	.515	.976	.937
SCL	9.456	3.100	9.275	2.761	.181	1.051	−.172	.666	.516	.976	.939
—Competence	2.767	.623	2.441	.374	.326	.436	.746	2.890	.012	.933	.298
—Positive Affection	2.300	.978	2.396	.262	−.096	.866	−.111	52.000	.679	.864	.027
—Immersion	1.500	.707	1.067	.334	.433	.574	.755	2.924	.011	.903	.105
Stream Appropriate	3.467	.743	1.669	.058	1.798	.729	2.465	9.548	<.001	.902	.104
Desire To Use	3.533	.640	1.728	.101	1.805	.591	3.053	12.000	<.001	.827	.008

the adaptive condition, the actual mean EDA was 0.040 µS ($SD = 1.074$) after a stable Stream was achieved, see Figure 4.4b. Similarly, the adaptive algorithm did not significantly increase tonic arousal as indexed by SCL. Actual SCL was 9.456 µS ($SD = 3.1$), reflecting predicted values of 9.275 µS ($SD = 2.761$), see Figure 4.4c. SCL and EDA results are summarized in Table 4.2.

Workload and Performance The physiologically adaptive condition did not increase subjective workload. Individual regressions predicted a mean raw NASA-TLX score of 8.31 ($SD = 2.25$), while workload was 8.57 ($SD = 3.20$) in the adaptive condition after a stable Stream was achieved, see Figure 4.4a and Table 4.2. This was mirrored for both the performances in the detection task and the n-back task. The adaptive algorithm regulated performance based on the individual linear fit to a local optimum and, thus, we found no sig-

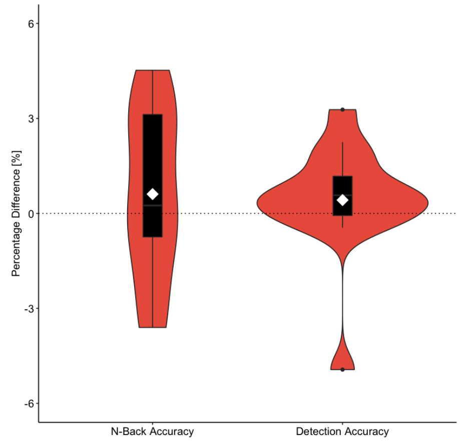


Figure 4.5: The relative difference for overall task accuracies in the n-Back and visual detection tasks.

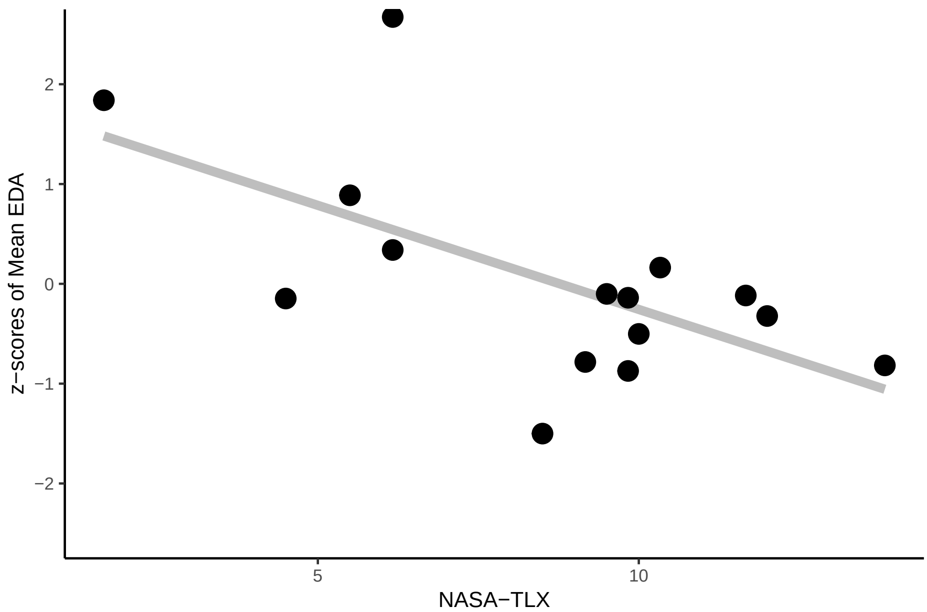
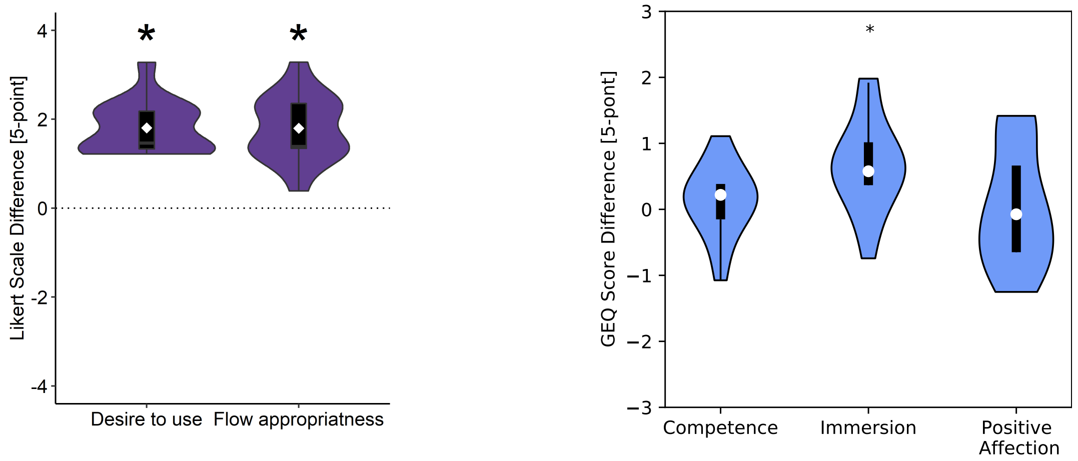


Figure 4.6: Standardized mean EDA at local maxima of adaptation as a function of raw NASA-TLX for the adaptive condition. There is a significant negative correlation between EDA and workload, $r(13) = -0.62$, $p = 0.013$.

nificant differences between the expected and actual performance for the adaptive condition as seen in Figure 4.5 and Table 4.2.

Figure 4.6 illustrates the negative relationship between workload and mean EDA. The algorithm allowed participants to perform tasks at their individual optimal levels of physiological arousal. Participants with relatively higher mean EDAs experience lower subjective workload $r(13) = -0.62, p = 0.013$. This shows the successful adaptation of the algorithm and that participants approached an individual local maxima in perceived workload.



(a) UX (b) GEQ

Figure 4.7: The relative difference for (**a**) usability questions measured on a 5-point Likert scale and (**b**) GEQ subscales (Competence, Positive Affection, and Immersion). * indicates that measurements are significantly different from the no-adaptation baseline. Outliers were defined as data points with a value greater than 2 SDs on the log-scale from its participant-mean. Outliers are represented as bold dots.

Subjective Evaluation The subjective feeling of being immersed, measured by the GEQ-immersion scale, was increased in the adaptive condition ($M = 1.556$, $SD = 0.705$) as compared to its expected value ($M = 1.032$, $SD = 0.323$), as predicted from the non-physiologically adaptive conditions. STREAM was also considered to be more appropriate in the adaptive condition, compared to the predicted value derived from the actual measures of the non-adaptive conditions. This converges with a desire to use the system as it increased when compared to the actual and the expected desire to use the system, see Figure 4.7a. In summary, participants favored the physiologically adaptive condition in terms of STREAM appropriateness, the desire to use, gaming immersion (GEQ), and their feelings of competence (GEQ) and positive affection (GEQ). Those results are graphically summarized in Figure 4.7b.

4.1.4 Discussion

We presented a design of an online physiologically adaptive VR system. In detail, our system can adapt to the complexity of a VR environment based on the user's arousal as measured using EDA. We tested this system in a VR environment where the adaptive system optimizes the stream of NPCs. In our evaluation, participants completed a primary task (n-back) and a secondary task (visual detection) while either being presented with five static levels of STREAMS of NPCs or an adaptive STREAM of NPCs.

First, we investigated the effectiveness of the STREAM modulation. Here, we lend credibility to earlier results on crowdedness [141, 353] and visual complexity [271, 396]. In detail, a higher STREAM of NPCs impacts EDA and the task performance, albeit in both the primary and the secondary task. Thus, the desire to use the system is negatively affected by increasing STREAM. Thus, we found that an online adaptation of the STREAM is feasible for modulating workload. Therefore, we investigated the effectiveness of such a physiologically adaptive VR system. Here, we compared the baseline from the non-adaptive conditions to the adaptive condition using a prediction model. We found significant differences between adaptive and non-adaptive conditions for the *desire to use the system*, *GEQ competence*, and *GEQ immersion*. Interestingly, participants also rated the flow significantly more appropriate than the baseline, while we kept the range of possible flows within the same bounds. Thus, we conclude that our physiologically adaptive VR systems that adapt the virtual environment's task complexity can improve VR experiences, especially in complex environments such as this simulated Social VR. This is a promising step towards online adaptive VR environments based on physiological sensor data. Our study adds to other findings in physiological computing, namely that sympathetic arousal, here indexed by EDA, can depict task engagement [81, 317, 432]. Additionally, our qualitative results are in line with studies that have adapted VR environments based on physiological arousal. Thus, our results are in line with prior work in terms of perceived workload [175] and task engagement [402] when interacting in a physiologically adaptive VR system. These results are promising for the implementation of EDA measures in a biocybernetic loop for adaptation, confirming subjective results obtained with other measures i.e., EEG and ECG.

At first, it is counterintuitive that EDA was not lower when using the adaptive system, as the system optimizes in order to not overload the user. However, we argue that lowering the EDA was not expected nor intended. In our implementation, we used a $-2/+4$ adaptation i.e., if NPCs need to be removed, we removed two, while we added four in the `more` adaption. Thus, this pushed the user to always take a bit more at once while reducing it only gradually. This prevents the user from just tuning out but also keeps up the pressure while not overloading. Moreover, in a real-world scenario, keeping as many characters in the environment as possible is important to not lose an understanding of the Social VR scenario one engages in, but also to prevent lowering the level of immersion and presence. We further see this manifested in the significantly higher rating of immersion. Here, we argue participants were less likely to be bored or stressed, as we optimized for optimal arousal according to the MIM [607] where flow-experience and task engagement correlate with rising demands [433]. As a result, they focused on the environment and, thus, achieved a higher level of immersion.

Behavioral results indicated that STREAM adaptation did not significantly impact n-Back and detection accuracies. This might point toward an insufficient increase in task complexity. However, it is not corroborated by either the EDA processed by our algorithm or from the non-negative deconvolution analysis. Second, as distractor processing is influenced by increased executive load [132], future studies should investigate task complexity adaptation under increased working memory load conditions. This conclusion is shared with the work of Dey et al. [145] and with Ewing et al. [175] when considering increased task demands.

Our findings have a broader impact on collaborative environments; virtual agents can assist or perform tasks within a collaboration VR environment in future VR applications. Such systems could exploit physiologically adaptive systems to modulate the VR complexity (e.g., amount of virtual coworkers or concurrent visually displayed information). This allows supporting concurrent goals without the need to prioritize performance over comfort while preventing cognitive overload. We argue that our results are not limited to Social VR per se. For instance, physiological adaptation can be used to transition on the reality-virtuality continuum [391]. Here, the VR complexity can overload the user with the virtual component, and, thus, it could gradually fade out the digital overlay. Thus, varying the blend of realities according to users' physiological arousal can be the next step toward a physiologically adaptive mixed reality system.

EDA data acquisition is low-cost and unobtrusive; skin-interfaced wearable systems are easy to implement in real or virtual scenarios, i.e., biofuel cell-based self-powered wearable sensors [203] or conductive fabric gloves [343]. Consequently, we argue that the next step is to embed physiological sensing into VR controllers, allowing practitioners to use them without a tedious setup process. This is consistent with recent developments in entertainment computing, where a player's emotional state as inferred from physiological correlates was used to predict actions in a game [333] or dynamically adapt the game narrative to induce states of arousal [208].

Limitations

EDA activity has a one-to-one relationship with the state of physiological arousal [81], which can serve as a workload indicator. However, the adaptation pipeline used in this study, although with a low computational cost, could benefit from more advanced and standardized methods of EDA component deconvolution [43, 220]. So far, this has not yet been used for online adaptation.

Tonic changes occur with a slow frequency and, therefore, may not be adequate for adaptive interactions that occur at a faster pace. Here, other physiological measures that are experimentally more demanding, but have better time-resolution, such as EEG, could be more suitable for adaptive systems [175]. Nevertheless, for adapting visual complexity in our Social VR scenario, adaptation in a third of a minute was suitable with regard to fostering immersion; here, faster adaptation could potentially break immersion as non-player characters would appear and disappear seemingly randomly.

We targeted the complexity exclusively in the visual domain. Still, evaluating multimodal physiologically adaptive systems could better support users' comfort as multimodal integration facilitates task-relevant information selection [150, 544]. For example, in a social VR scenario, the visual representation could disappear, and voices from afar could be toned down to support the user in focusing on the primary task.

According to Braithwaite et al. [70], about 10% of the general population do not exert a strong electrodermal response attuned to sympathetic arousal. This is especially relevant when considering commercial applications of EDA sensing, e.g., for adaptive VR games where potentially 10% would not benefit from a solely EDA-based system. Thus, supplementing EDA with other arousal measurements such as the heart rate or eye-gaze could make measurements more robust and make the technology more accessible to the general population. Together with this, we had to exclude three participants via medical-grade instrumentation for high data quality due to a lack of data quality. This is consistent with EDA guidelines [70] and could be attributed to either technical failures and to individual variations in EDA. While we controlled for other factors such as physical stress [243] or caffeine consumption [34], using EDA in applied scenarios can not be shielded from such influences.

Finally, we acknowledge that only the data of 15 participants were used in the analysis. While on the one hand, this is a low number, we could still show large effect sizes for the appropriateness of the stream, and most importantly for the desire to use the adaptation. As such, larger sample sizes will be needed only for the measures with small and medium effect sizes.

4.1.5 Open Science

The data presented in this study are openly available in DaRUS Open Data Platform at `https://darus.uni-stuttgart.de/dataset.xhtml?persistentId=doi:10.18419/darus-2820`.

4.2 Study 5: Investigating Physiological Correlates of Secondary Task Difficulty Adaptations

Physiological computing is an emerging field investigating how physiological correlates of human affective and cognitive states can be used as an input in adaptive systems to achieve specific goals [178]. VR, in particular, is a fertile ground for such physiologically-adaptive systems as it allows for online manipulation and adaptation of visualizations, virtual content, and interactions [113] that would otherwise be impossible in physical reality. Adaptive VR systems can optimize a set of behavioral, physiological, and subjective measures by dynamically adjusting the system's current task parameters to improve performance and support users to maximize their amount of productive work or task engagement. Physiologically-adaptive systems are now deployed in various VR scenarios such as social VR, exergaming, and cognitive training. Chiossi et al. [110] adapted the visual complexity of the secondary task in the form of virtual agents based on EDA. Campbell et al. [84] made use of variations in HR for adapting the physical load of a VR exergame [84], while Dey et al. [145] adjusted the amount and properties of distractors in a visual search task based EEG alpha oscillations.

Central and peripheral physiological measures showed to be able to quantify and predict workload [88, 546] across various VR applications, such as learning [119], balance training [148] or executive tasks [430]. HR increased with mental workload [74] and EDA discriminated across workload levels [385] and showed better test-retest reliability than other physiological measures [395]. Furthermore, different EEG features behave differently upon the involvement of specific attentional processes, i.e., external or internal attention [120, 365] or working memory [420], which might be differentially allocated in complex tasks with different degrees of workload [27, 117]. Based on the adaptive alpha response to task demands, a decrease/increase in alpha oscillations has been associated with cortical excitation/inhibition in Working Memory (WM) [80, 282] and visual detection tasks [170]. Moreover, an increase in frontal-midline theta EEG oscillations was reported when cognitive demands for updating, organizing and retrieving information were recruited [258, 497]. Finally, beta oscillations have been shown to discriminate between different task complexity levels [94, 180] and correlate with physiological arousal [242, 393]. However, there is no universal physiological measure or method to index mental workload, as different physiological measures have been shown to discriminate between different features of task load [88]. Certain measures are more sensitive to task demands, and others are more sensitive to task complexity.

Those challenges are shared with the ones of physiological computing [178], such as psychophysiological inference, i.e., mapping a physiological signal to a specific cognitive state of the user. Therefore, combining different physiological measurements can allow for a hybrid online evaluation of the adaptive system. Essentially, a second signal would be used to evaluate the success of the adaptation instead of just measuring the final effect of an unsuccessful adaptation, i.e., a decrease in task performance. This solution has been proposed to increase the reliability, proficiency, and utility of BCI systems, otherwise known as multimodal or hybrid BCIs [33, 227]. A first attempt was performed by Labonte et al. [330], which employed automatic facial expression analysis as a second signal to evaluate dynamic difficulty adjustments of Tetris based on alpha and theta EEG oscillations. However, they reported that the hybrid adaptive system did not improve the participants' experience as the hybrid system showed more negative affect than the control condition.

Therefore, more work is needed to link behavioral performance, workload, and physiological measures to investigate how to align them for user's personalization and adaptation effectively [463]. This is specifically relevant for adaptive systems, as multimodal input has been relatively overlooked [113] or mainly focused on alternative channels for adaptation, i.e., speech and gesture recognition [276]. Thus, the goal of this work is to investigate which are the relationships between a range of measures, extracted from physiological signals such as EDA, ECG, and EEG, and evaluate the effect of VR system adaptation on such measures. This is especially relevant considering that different physiological signals might need different time windows to react to adaptations, and therefore some might be more suitable for faster paces of adaptations, while other might need slower paces [205, 539]. To achieve this, we analyzed a dataset encompassing multiple physiological signal recordings and physiological adaptation. Specifically, we chose the dataset of Chiossi et al. [110], which featured an adaptation of secondary task difficulty based on EDA feature and co-registration of ECG and EEG data.

In this work, we shifted our focus from evaluating the adaptive system to a detailed analysis of the relationship between the various physiological measures, and the effect of VR system adaptation on them. By measuring the impact of these changes over different physiological measures, we intend to evaluate the user's reaction to the adaptation in real-time. Based on the logic of the physiologically-adaptive VR system and previous work, we hypothesize that:

H1 When the adaptive system adjusts for increased secondary task difficulty, this should result in increased workload, resulting in increased participant's physiological arousal. We hypothesis that to an increase in secondary task difficulty should correspond an increase in physiological correlates of arousal, i.e., beta oscillations, SCL and average amplitude of non-specific skin conductance responses (nsSCRs).

H2 An increased secondary task difficulty could increase the working memory load of N-Back task, as indexed by increased theta oscillations [497].

H3 An increased secondary task difficulty could increase the visual load in the visual detection task, resulting in increased external attention and therefore decreased occipital alpha oscillations [365].

H4 As HR and HRV increase are related to increasing task [134] and visual attention demands [68], we might expect increased HR and HRV when the system adjusted for increased secondary difficulty.

4.2.1 Dataset Processing

We utilized the dataset from Chiossi et al. [110] containing behavioral, physiological (EEG, ECG, and EDA), and subjective data. We refer to their paper for a detailed description of the task implementation and data collection. The dataset included 18 participants ($M_{range} = 23 - 31$; $M_{age} = 27.9$, $SD_{age} = 2.9$; *male* $= 9$, *female* $= 9$), but only 15 are included as three participants were removed due to technical issues. They recorded behavioral and physiological (EEG, ECG, and EDA) during the task. They recorded EDA (at 250Hz) via the GSR module by BrainProducts GmbH, Germany and ECG (at 130 Hz) via a Polar H10 chest strap (Polar, Finland). EEG data recording was performed at 250 Hz with a 7-channel dry electrode cap embedded into the HTC VIVE headset from Wearable Sensing (DSI-VR 300, San Diego, CA, USA) using the electrode positions: FCz, Pz, P3, P4, PO7, PO8, Oz of the 10-20 system.

EDA Data

We processed EDA data via the Neurokit toolbox [367]. Preprocessing pipeline for EDA data encompassed first via a third-order Butterworth filter with a 3Hz high-pass cutoff. Then, we applied a nonnegative deconvolution analysis [43] to extract tonic and phasic components. Specifically, we computed the average amplitude of Non-Specific Skin Conductance Responses (nsSCRs) and the average tonic SCL. We identified nsSCRs peaks using a $.05\mu S$ threshold value, upon guidelines (cf. [453]).

ECG Data

We evaluated ECG activity in the time domain, focusing on HR and HR. As for EDA data, we used the Neurokit Python Toolbox [367]. We first filtered the ECG signal by the Finite Impulse Response (FIR) band-pass filter (3–45 Hz, 3rd order), and then segmented by Hamilton's method [230] to identify the QRS complexes and extract mean HR and HRV, defined as the root mean square of the successive differences (RMSSD).

EEG Data

We processed the EEG raw data via the MNE Toolbox [216]. EEG data were recorded with a sampling frequency of 250 Hz from dry electrodes placed on Fz, P3, Pz,P4, PO7, Oz, PO8 locations (10/20 system), with reference set at linked earlobes. We notch-filtered the signal at the power frequency of 50 Hz and then band-passed between 1 and 70 Hz to remove high and low-frequency drifts. We then performed a visual inspection to identify and exclude corrupted channels. Finally, we referenced the data to the CAR. Next, we computed an ICA using MNE [40] to identify and correct artefactual components automatically with the ICLabel plugin [443]. We then analyzed the preprocessed EEG data in three frequency bands: Theta (4–8 Hz), Alpha (8–12 Hz), and Beta (12-30 Hz), using Welch's method [590]. We computed alpha oscillation for posterior sites, i.e., PO8, PO7, and Oz electrodes and extracted Theta and Beta oscillations from midline sites, i.e. Fz and Pz. Then, we log transformed them to achieve normal distributions [454]. Moreover, we computed the ratio of midline theta activity's absolute power to posterior alpha activity's absolute power as an implicit measure of workload [76, 202].

4.2.2 Experimental Task

Participants started the experiment upon informed consent signing. The experiment required participants to be engaged in a dual-task paradigm, encompassing a VR N-Back task (N=1) (primary) and a visual detection (VD) task (secondary), where they had to select Non-Playable Characters (NPCs) without a ticket. NPCs entered and walked past the participants' surroundings with or without a ticket. Participants had to select NPCs without a ticket and click on them with the trigger of the HTC VIVE controller. Specifically, participants experienced six conditions; five non-adaptive conditions had fixed STREAM of 7, 22, 37, 52, or 67 NPCs per minute entering the scene and one adaptive condition. In the adaptive condition, the STREAM was adapted based on user's arousal as measured using EDA. In more detail, Chiossi et al. [110] acquired a baseline EDA baseline recording and computed the baseline EDA tonic component. Then, in the adaptive condition, they adapted the secondary task difficulty based on the variation of the tonic EDA component every 20 seconds. Therefore, secondary task difficulty adjustments were performed in the 20s-window by either (I) adding 4 NPCs to the scene, i.e., increasing the visual complexity of the VD task, if they detected that the participants showed a decreased EDA tonic component as compared to the baseline or

(II) removing 2 NPCs from the scene, i.e., decreasing the visual complexity of the VD task, if the online EDA tonic component was increased when compared to the baseline, reflecting a state of higher arousal.

4.2.3 Data Analysis

Our analyses examined physiological indicators of cognitive workload and arousal collected while participants were jointly engaged in a visual WM task and in a visual detection task. To evaluate the effect of visual complexity adaptations, we focused our analysis on the adaptive condition, segmenting EDA, ECG, and EEG signals into 20 seconds epochs based on when the STREAM of NPCs was adapted. Specifically, two variations in the STREAM, based on the adaptation algorithm : (I) *Increase*: based on decreased arousal state compared to baseline; (II) *Decrease* based on increased arousal state as compared to baseline. The VR-physiologically adaptive system performed an average of $M = 4.89$ *Increase* ($SD = 1.691$) while the STREAM was decreased on average of $M = 5.06$ ($SD = 1.89$). Depending on the normality testing via Shapiro-Wilk test, we performed paired t-test for normally distributed distributions and Wilcoxon signed-rank test for not-parametric distributions. Second, we compute Pearson correlation to investigate relationships between the extracted physiological features.

4.2.4 Results

We present quantitative findings based on the physiological and behavioral data from the dataset. We investigated differences in the signal when participants were either exposed to a *Increase* or a *Decrease* of STREAM. Therefore, evaluate the effects of visual complexity adaptation over the extracted dependent variables.

EDA Results

Results of SCL, and the average amplitude of nsSCRs are depicted in Figure 4.9a. **Skin Conductance Level (SCL):** Given a violation of normality ($W = .83, p < .001$), a Wilcoxon signed-rank test did not detect any significant differences in the SCL when participants experienced a *Increase* compared to a *Decrease* ($W = 359, p > .05$). **Non-specific Skin Conductance Responses (nsSCRs):** Similarly to the SCL, also the average amplitude of the nsSCRs was not normally distributed ($W = .87, p < .05$). Therefore, a Wilcoxon signed-rank test showed that nsSCRs showed an increased amplitude when participants were exposed to a *Increase* when compared to a *Decrease* ($W = 616, p < .001$).

ECG Results

We display the results on ECG measures, i.e., Heart Rate and Heart Rate Variability, in Figure 4.9b. **Heart Rate:** HR showed a normal distribution ($W = 0.97$, $p > .05$), a paired t-test did not reveal any effect of STREAM on HR. ($t = -.62$, $p > .05$). **Heart Rate Variability:** As HR,

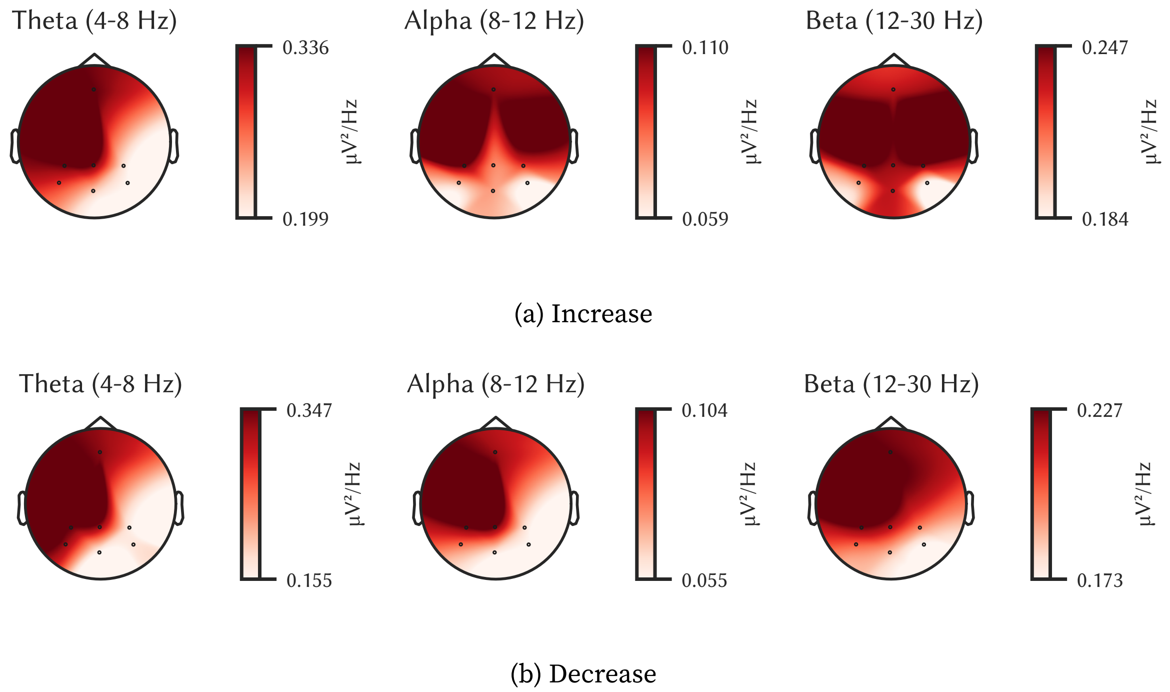
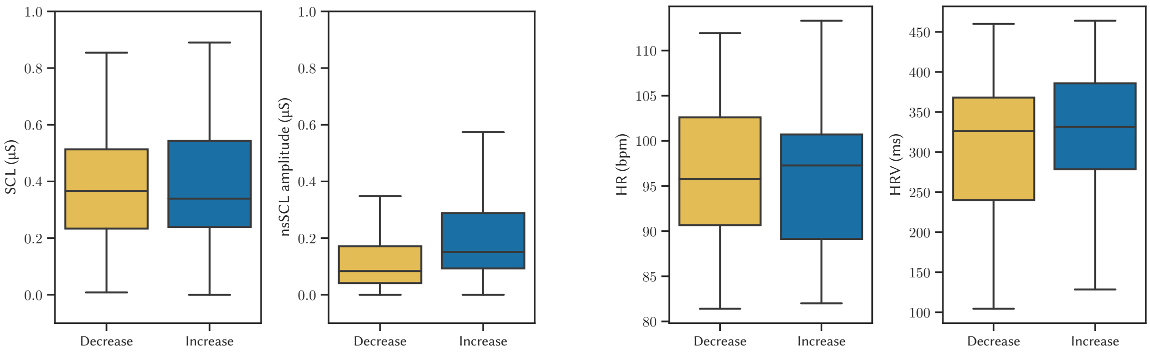


Figure 4.8: Topographic maps for the increase and decrease variations of secondary task difficulty for the EEG oscillations of interest. View is top looking down with nose at top.

also HRV was normally distributed ($W = 0.92$, $p > .05$) and not influnced either by *Increase* or *Decrease* ($t = .812$, $p > .05$).

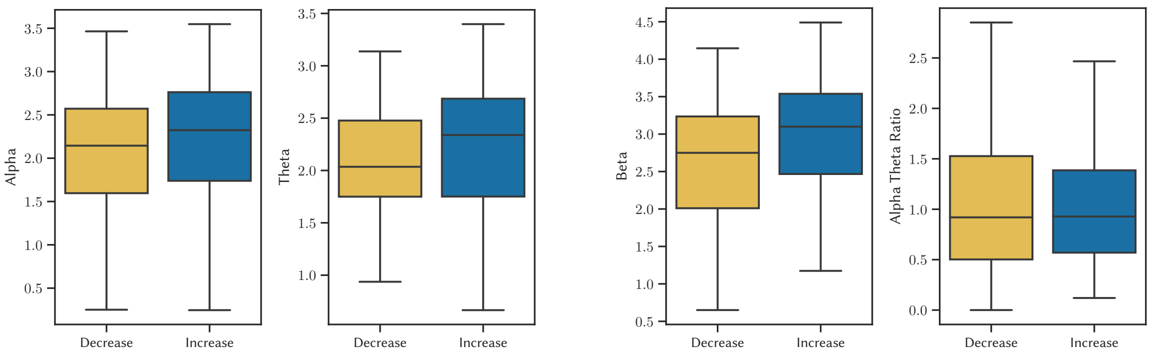
EEG Results

The results from the comparison between the variations in STREAM on EEG features, i.e., Alpha, Theta, and Beta oscillations, are shown in Figure 4.33 and topographic distribution in Figure 4.8. Here, we supplement our results with the ratio of Alpha and Theta oscillations. **Alpha Band:** The alpha power was normally distributed ($W = 0.96$, $p > .05$). However, a paired sample t-test did not reveal a statistically significant difference between *Increase* and *Decrease* of STREAM ($t = 1.1$, $p > .05$). **Theta Band:** Average Theta power showed a normal distribution ($W = 0.98$, $p > .05$) and showed significantly increased power in the *Increase* as compared to the *Decrease* ($t = 2.06$, $p < .05$). **Beta Band:** Beta power distribution was not normally distributed ($W = .92, p < .05$). Therefore, we submitted Beta scores to a Wilcoxon signed-rank test, which showed a significantly increased Beta power for a *Increase* ($W = 440$, $p < .05$). **Alpha-Theta Ratio:** The ratio between averaged Alpha and Theta powers lead to a not-normal score distribution ($W = 0.91, p < .05$), a Wilcoxon signed-rank test did not detect any difference between the two STREAM variation ($W = 337$, $p > .05$).



(a) EDA (b) ECG

Figure 4.9: EDA and ECG results. On the left, we depict the results for SCL and average nsSCRs amplitude. On the right, we show results for HR and HRV. The only significant difference is detected in the nsSCRs, which are increased in the INCREASE.



(a) Alpha/Theta (b) Beta/Alpha-Theta-Ratio

Figure 4.10: EEG oscillations results. In order, we present the differences in alpha, theta and, beta oscillations in the two variations of secondary task difficulty. In the last plot, we plot the A/T Ratio. Beta and Theta oscillations are significantly increased in the INCREASE.

Behavioral Performance Results

Here, we report the results of the variation in performance when the STREAM was either increased or decreased for both primary (N-Back) and secondary (Visual Detection) tasks. Results are shown in Figure 4.11. **Primary Task Performance:** Accuracy scores did not show a normal distribution ($W = .925$, $p < .05$). Here, a Wilcoxon signed-rank test did not show any further significance ($W = 103.5$, $p > .05$). **Secondary Task Performance:** Secondary task accuracy was normally distributed ($W = .982$, $p < .05$). However, as primary task performance, change in STREAM did not show any significant change ($W = 301$, $p > .05$).

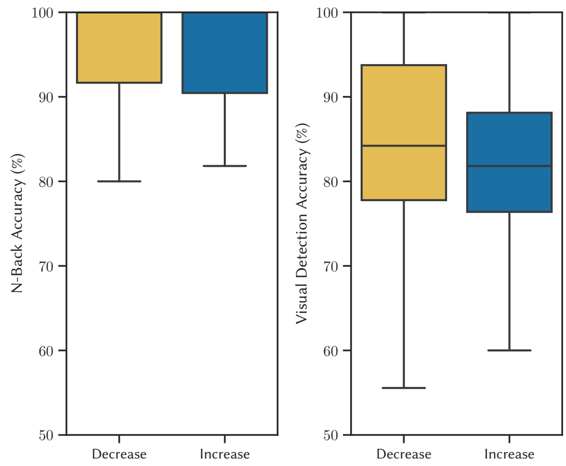


Figure 4.11: Behavioral performance results. Boxplots for the accuracy of the primary task (N-Back) and secondary task (Visual detection task). No significant differences are detected in the two measures.

Correlation Analysis

The correlation matrix depicted in Figure 4.12 displays the Pearson coefficients between EDA, ECG, and EEG measures and behavioral performance in the N-Back (primary) and VD (secondary) tasks together. Results suggest that 21 out of 44 correlations were statistically significant and were greater or equal to $r = .35$, $p < .05$. Correlation between Beta and Theta oscillations was reported to be strongly positive ($r = .96$, $p < .001$), while Alpha oscillations strongly significantly correlated with Theta ($r = .7$, $p < .005$) and Beta ($r = .71$, $p < .005$). Regarding EDA features, SCL, and nsSCRs amplitude were significantly positively correlated ($r = .63$, $p < .001$), and similarly SCL strongly correlated with Alpha oscillations ($r = .75$, $p < .005$). Finally, HR and HRV showed a significant negative correlation ($r = -.683$, $p < .001$). HR showed low positive correlations with Theta ($r = -.417$, $p < .05$), with A/ T Ratio ($r = .352$, $p < .001$) and negative with Beta oscillations ($r = -.403$, $p < .001$). Finally, HRV showed a moderate positive correlation with an EEG arousal correlate, i.e., Beta ($r = .544$, $p < .001$) and low positive correlations with Theta ($r = .493$, $p < .05$) and A/T Ratio ($r = .352$, $p < .05$).

4.2.5 Discussion

In this work, we evaluated the effect of secondary task difficulty adaptation over different physiological measures coregistered during interaction with a VR physiologically-adaptive system.

We first hypothesized (**H1**) that INCREASE in task difficulty for the VD task would have increased the arousal response in the form of beta oscillations and EDA measures, i.e., SCL and nsSCRs. We partially verified this hypothesis by finding increased arousal for beta and nsSCRs, but not for SCL. This difference in the EDA can be explained by the fact that nsSCRs are a phasic component and therefore show faster responses to environmental stimuli than its tonic

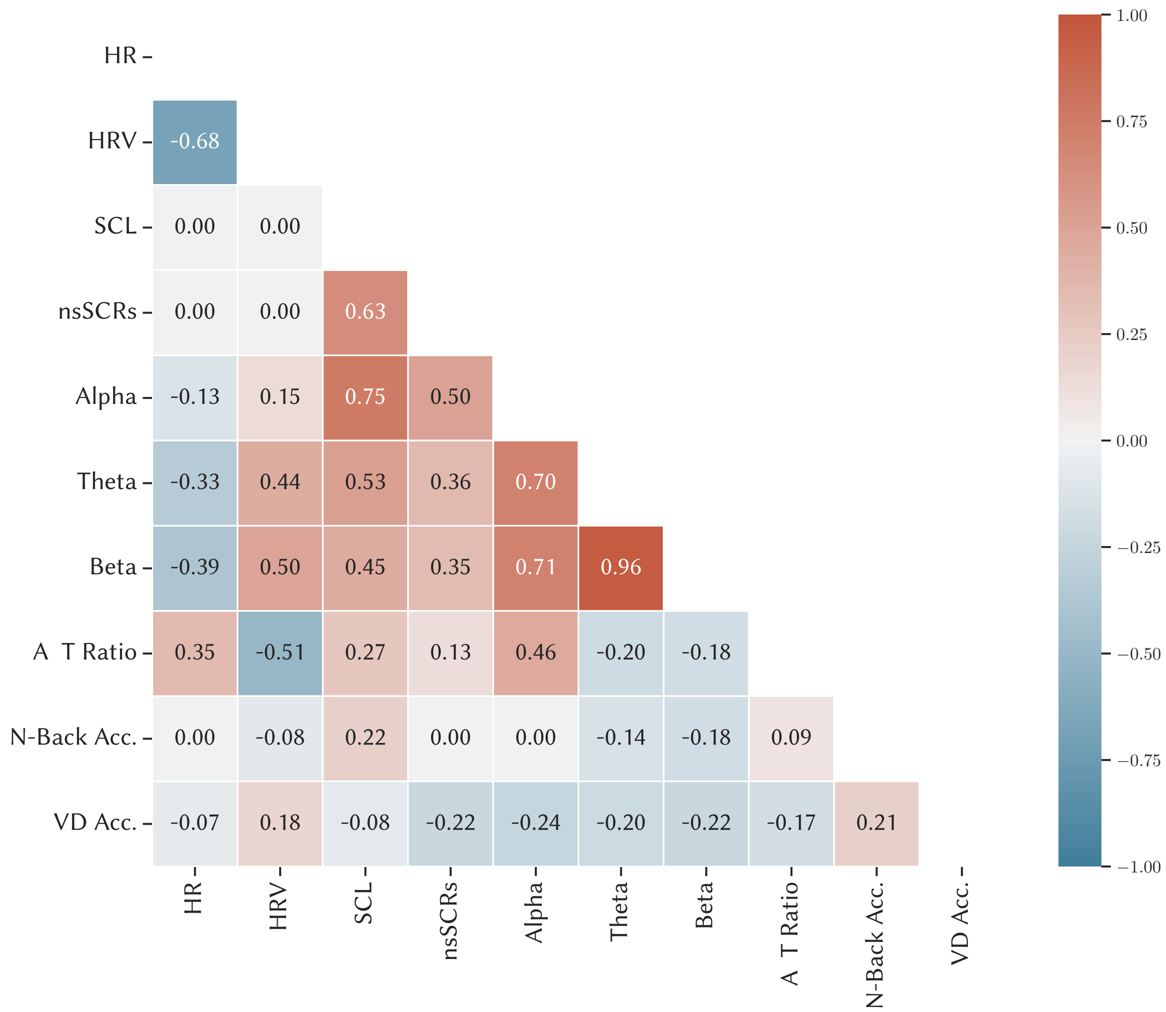


Figure 4.12: *Correlation Matrix*. Colors indicate the strength of Pearson correlation coefficients. HR = heart rate; HRV = Heart Rate Variability; SCL = Skin Conductance Level; nsSCRs = average amplitude of the nonspecific skin conductance responses; Alpha = average Alpha power; Theta = average Theta Power; Beta = average Beta power; A T Ratio = ratio between average Alpha and average Theta power; VD Acc. = average performance in the visual detection task; N-Back Acc. = average performance in the N-Back task.

counterpart and similar time dynamics as brain patterns [450]. We can therefore argue that for evaluation purposes, a phasic component might be a better candidate than tonic components to evaluate physiological reactivity to adaptations that occur at a fast pace ($\sim 20s$). Similarly, we can draft a similar conclusion regarding **H4**, where ECG-related measures, i.e., HR and HRV, failed to discriminate between the two levels of secondary task difficulty adaptation. HR and HRV are able to discriminate to different task demands but only at the overall task level. In fact, HR discriminated between highest and lowest task load but in two different conditions [532], or between tasks but not for task difficulty [564].

Regarding results from EEG oscillations, we verified **H1** and **H2**, but not **H3**. We replicate results from Deiber et al. [139], which found increased beta and theta oscillations with increased WM load. Thus, an increased VD task difficulty impacted the WM load over the N-Back task as participants had to allocate more attentional resources to the VD task and put more effort into updating the WM information. Theta oscillations are one of the most robust indicators of WM engagement and cognitive control [283, 478, 497]. This result is also supported by the strongest correlation we reported between beta and theta in our analysis. Alpha oscillations, instead, were not significantly affected by the difficulty adjustment. We can explain this as an insufficient increase in the visual load given the adaptation parameters chosen by Chiossi et al. [110] (-2 / +4 variation in visual complexity). Alpha increase has been classically related to increasing visual and attentional load [311, 439], even in adaptive systems [145]. Hence, future adaptive systems that adjust visual complexity and employ alpha oscillations should optimize adaptation parameters with larger differences. Although prior research has been exploring optimizing system behavior based on human inputs [421], our results are related to the evaluation of difficulty adjustments; such outcomes should motivate future researchers to investigate optimizing physiological features [149] by adapting system parameters, which could be a more reliable measure that complements subjective functions [422] using Bayesian optimization.

Finally, correlations provided two specific interesting results, showing high correlations between Theta and Beta oscillations, SCL and Alpha oscillations and moderate between SCL and nsSCRs amplitude. Even though our results are related to the evaluation of difficulty adjustments, such outcomes should motivate future researchers to investigate these features as input and to improve adaptation algorithms in hybrid systems. Furthermore, our results show how multimodal evaluation of physiological reactions to adjustments in adaptive systems is feasible and can be a promising side tool to improve adaptation algorithms in hybrid systems.

4.3 Study 6: Adapting Environmental Visual Complexity based on Physiological Arousal

4.3.1 Architecture of the Physiologically-Adaptive VR System

The closest work to our system is by Chiossi et al. [110]; they based their adaptation on raw EDA data for task difficulty adaptation. Specifically, they adapted the secondary task difficulty of a visual detection task. Therefore, they computed the EDA slope of a 180sec baseline recording before the study. With this, they adapted the difficulty based on the difference between the EDA baseline slope and the online EDA slope computed every 20sec. A major drawback of the work by Chiossi [110] is that they tend to compare to an old baseline, e.g., after one hour of online adaptation, the baseline recording is also one hour old. Thus, they are not accounting for effects like EDA saturation and drifts [221, 309]. In this work, we overcome this limitation by eliminating the baseline recording and, thus, only including the latest EDA data, i.e., we adjust the amount of NPCs comparing the EDA signal to the averaged EDA signal acquired in the previous 20sec. This choice allows for increased reliability of the signal on which the adaptation is based on and better deployment outside of the lab. The architecture of adaptive and control adaptive conditions is depicted in Figure 4.22.

First, we preprocessed the raw EDA signal by removing low-frequency noise with a fourth-order Butterworth filter with a 3 Hz high-pass cutoff. Second, we used non-negative deconvolution analysis to decompose the signal and extract its tonic component, i.e., SCL [43].

Second, we apply a rolling window approach for the adaptation. We use a long-term window for low-frequency changes in SCL of window length w_1 (in our study, w_1 is 180sec long). Additionally, we use a short window for high-frequency changes with window length w_2 (in our study, w_2 is 30sec long). Resulting in the three points t_0, t_{-1}, and t_{-2}, see Figure 4.22. We calculate the SCL level at all three points t by averaging the SCL values from t to $t - \epsilon$. Here, ϵ allows averaging over some data before the points to stabilize the value. Thus, in the following, the SCL level at t_x is the mean value of $t_x - \epsilon$ until t_x, defined as $SCL(t_x)$.

Thirds, we compute the slopes of the change in SCL in the two windows, i.e., s_1, and s_2. We calculate s_1 from the average tonic value of t_{-2} and t_0, $s_1 = (SCL(t_0) - SCL(t_{-2}))/w_1$. Moreover, we calculate s_2 the same way but using t_{-1}, t_0, and w_2.

Finally, we compare the low-frequency slope s_1 to the high-frequency slope s_2. We compare their difference and drive the adaptation as the following:

$$adaptation(s_1, s_2) = \begin{cases} \texttt{increase} & \text{if } s_1 \leq s_2 - \theta \\ \texttt{decrease} & \text{if } s_1 \geq s_2 + \theta \end{cases} \tag{4.4}$$

Here, the threshold parameter θ allows the adaptation to occur within a certain variance, preventing quick and unstable adaptations. Lastly, Equation 4.5 is only used for adaptation every 20s, as depicted in Figure 4.22.

4.3.2 User Study

We state the following research questions, informed by related work:

RQ1: Can online dynamic adjustments of visual complexity based on a peripheral physiological measure of arousal support task performance?

RQ2: Does adaptation based on the motivational intensity model support task performance?

The present experiment was designed and conducted to evaluate if a physiologically-adaptive system, based on SCL, supports the user's behavioral performance as compared to not adaptive ones and an adaptive control condition. As a primary task, we chose the established N-Back task [275] as adapted from [110] in an immersive VR environment. This task required a

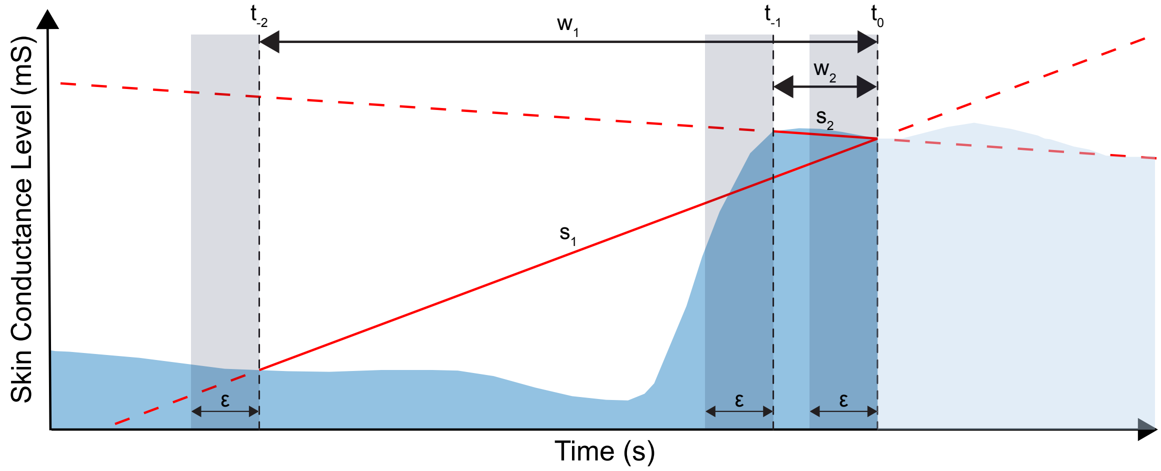


Figure 4.13: An overview of how the adaptation is computed. In this case s_2 shows an increased slope as compared to s_1. This result would lead to a decrease in STREAM in the adaptive condition or an increase of STREAM in the control adaptive condition.

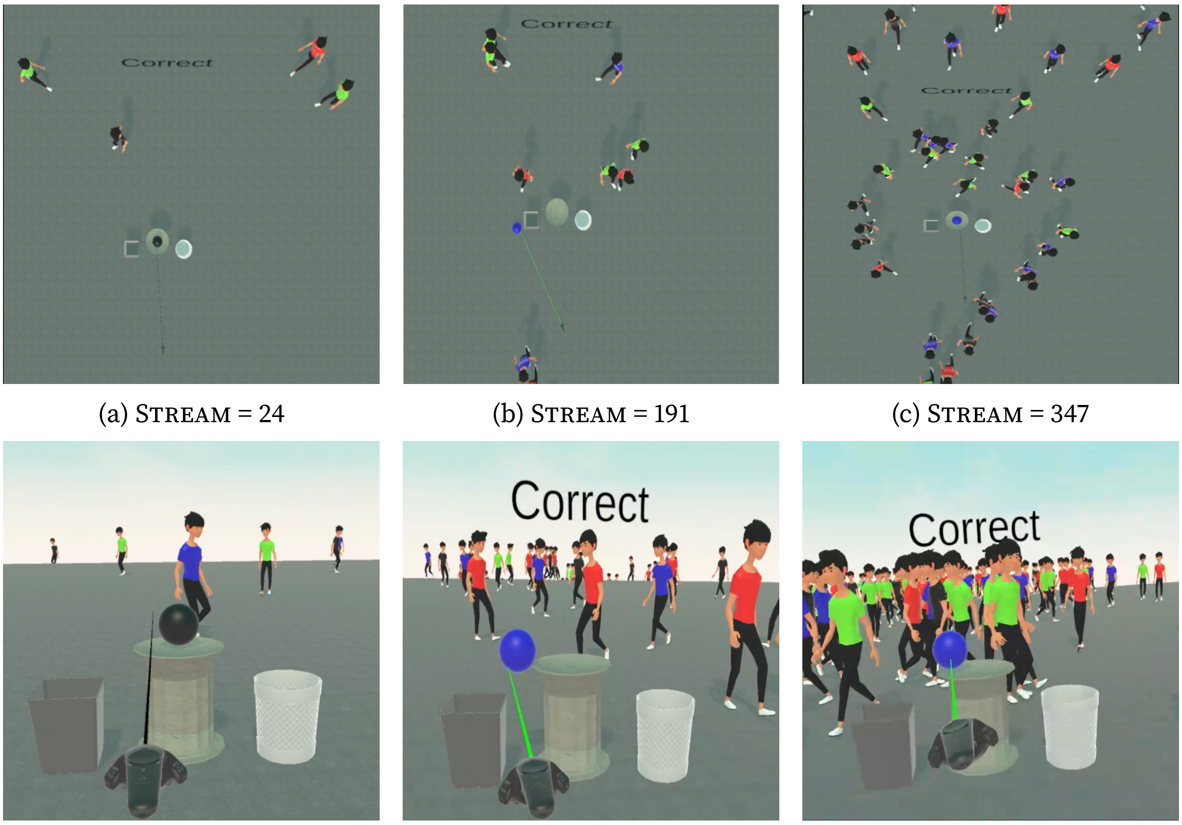
constant information update stored in working memory for each trial and constant attention to spheres presented, and maintenance of previously presented information.

Design

To examine differences in performance, perceived workload, and UX, we performed a within-subjects study for the system's adaptability factor (Physiologically-Adaptive system vs. Reverse adaptive system vs. Not Adaptive system). The order of all conditions was randomized across participants. The experiment encompasses seven blocks, of which five have a stable STREAM, and two have an adaptive STREAM. Here, STREAM is defined as the amount of NPCs entering the VR scene per minute, see Figure 4.14. The "adaptive" experimental condition involved the manipulation of the visual complexity (i.e., the STREAM of NPCs) through changes in the participant's tonic EDA level as measured by Skin Conductance Level (SCL) [453]. We evaluated four aspects of the system: (i) N-Back task performance, (ii) average SCL, (iii) overall perceived workload, subjective engagement (in-game Game Experience Questionnaire), and UX (ad-hoc survey) across conditions. Specifically, as a control condition to investigate RQ2, we implemented a reverse adaptation that follows a reversed algorithm based on the motivational intensity model [607] that increases visual complexity and physiological arousal. In the remaining five non-adaptive conditions, participants were stimulated with fixed STREAM of 24, 110, 191, 270, and 347 of NPCs entering the scene per minute. Participants were not aware whether they would experience the adaptive condition, control adaptive condition, or one of the non-adaptive conditions to avoid biases in performance and subjective ratings.

Physiological Data Recording

We followed the recent guidelines and framework for the human-computer interaction [26] for EDA data recording. EDA was recorded using standard Ag/AgCl electrodes (7 mm surface



(a) Stream = 24 (b) Stream = 191 (c) Stream = 347

(d) Stream = 24 (e) Stream = 191 (f) Stream = 347

Figure 4.14: VR capture of a single trial of the VR n-back from a birds-eye perspective in the first row, and first-person view in the second row. In (a) is depicted the condition with low visual complexity with Stream = 24. In (b) is depicted the condition of moderate visual complexity with Stream = 191 and lastly in (c), the highest visual complexity with Stream = 347.

diameter) placed on the distal surfaces of the middle phalanges of the index and middle fingers of the participant's non-dominant hand. Data collection started with an electrolyte solution application (0.5% $CaCl_2$) over acquisition sites to ensure proper hydration and minimize individual differences' effects. After the application, the participant waited 10 minutes before the electrodes were attached to the participant's phalanx utilizing double-sided adhesive collars. For data acquisition and amplification, we used a LiveAmp from BrainProducts GmbH amplifier combined with a Sensor and Trigger Extension for a GSR-Module, which exploited the exosomatic recording principle with direct current (DC) with a constant voltage of 0.5 V. Sampling rate was set at 250 Hz. For offline data preprocessing, we followed the same pipeline as for the adaptive and adaptive control conditions via the Neurokit toolbox [367]. After non-negative deconvolution analysis, we derived two metrics of physiological arousal: the average amplitude of Non-Specific Skin Conductance Responses (nsSCRs) and the average tonic SCL. We identified nsSCRs peaks using a .05 mS threshold value, following the recommendations by the Society for Psychophysiological Research [453].

Apparatus

We implemented the VR environment and tasks using Unity 3D (Version 2020.1.8f1). We acquired three physiological measurements: EDA signal using a GSR Module (BrainProducts GmbH, Germany, 250 Hz), ECG via PolarH10 chest strap (Polar, Finland, 130 Hz), and EEG signal (DSI-VR 300, Wearable Sensing, San Diego, CA, 300 Hz). For the scope of this paper, we will only report and analyze results on EDA. Physiological data were streamed within the Unity VR environment within the Lab Streaming Layer (LSL) framework[2] to the acquisition PC (Intel Core i7 with 3.00GHz, 32GB RAM).

The number of NPCs in the adaptive condition was adjusted using Equation 4.5. After multiple testing runs, we set w_1 to be 180sec long and w_2 30sec long. Further, we determined ϵ to be 20sec. In detail, sixteen NPCs were added to the VR scene if the s_2 slope was smaller than the s_1 slope. On the other hand, eight NPCs were removed if s_2 slope was greater than s_1 slope, which is indexing increased arousal state. Those adjustment values represent 10% and 5% of the starting STREAM value for adaptation, respectively. Several test sessions were employed to determine these settings empirically. Regarding the control adaptive condition, we used the opposite algorithm. Therefore, if s_2 was smaller than s_1 we increased the STREAM of 16 NPCs, while if it was bigger than s_1 we decreased of 8 NPCs, see Figure 4.22. This design is inspired by the previous work of Chiossi et al. [110], which employed a +4/-2 for task-relevance elements adaptation. Here, we employed a +16/-8 design to account for task-irrelevance of the distractors based on previous work [103]. Regardless of the adaptive or control adaptive condition, or the physiological activity of the participants, NPCs were spawned and removed from the VR environment within a range of 24-347.

To extract the tonic component via non-negative deconvolution analysis, we streamed EDA raw data via a Transmission Control Protocol (TCP) /Internet Protocol (IP) client to the TCP/IP server developed by Python network programming. This implementation allowed us to pass forward and backward data between LSL and the VR Unity environment. To ensure that the time of the VR Unity scene is synchronized with the time of the bridge server's operating system, we synchronized both systems with a Network Time Protocol (NTP) service. We preprocessed online EDA raw data using the Neurokit Python Toolbox [367] within the Python Client-Server bridge service.

4.3.3 Task

The task was adapted from the N-Back in the study by Chiossi [110]. Participants were immersed in a neutral VR environment and presented with a marble-like pillar and two buckets placed on the left and the right, respectively. Over the pillar, spheres were spawned in four possible colors (green, red, blue, and black), following [382]. The color sequence was randomly generated. Participants had to grab spheres with an HTC VIVE controller and drop them into the correct buckets. If the sphere matched the color of the sphere

[2] `https://github.com/labstreaminglayer/`

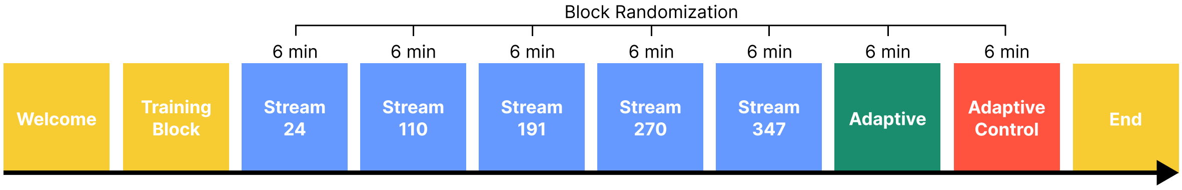


Figure 4.15: The experiment encompassed seven different blocks for data recording (blue, green, and red). Participants filled in NASA-TLX [237], GEQ subscales, and ad-hoc UX surveys between these blocks. Participants started the experiment with a training phase and then experienced the seven blocks in randomized order.

presented two steps before, the sphere has to be placed in the right bucket. If the sphere's color did not match the sphere's color two steps before, the sphere had to be put on the left bucket. Each sphere's appearance was signaled by a tone (800 Hz). Then, in order to avoid making an error, the participant had to pick up the sphere within 4sec. New spheres would appear when the current sphere was placed into one of the two buckets or after 4sec. Here, participants received accuracy feedback every 20 spheres and were instructed to maintain 90% performance.

Procedure

Upon arrival, we informed participants of the study procedure and answered any open questions, followed by signing the informed consent form. Next, we attached the participants' EDA sensor, EEG-VR headset, and ECG chest strap. We collected ECG data positioned over the xiphoid process of the sternum beneath the chest muscles. Participants performed a VR visual WM task, i.e., the N-Back task (N=2). The experimental procedure encompassed a training phase and the execution of the seven experimental blocks in randomized order. The experimental procedure is depicted in Figure 4.15.

The study began with a *training phase* to familiarize with the VR environment. Here, participants practiced with the 2-back task until they reached an accuracy level of at least 90% within a sequence of 80 spheres. The *experimental phase* consisted of seven experimental blocks, lasting six minutes each. Between the blocks, we asked participants to complete the raw NASA-TLX [237], the Game Experience Questionnaire (in-Game Core Module) [270], and two system's UX ad-hoc surveys on a 7-point Likert Scale 1) "I would like to use the system in the future," 2) "The flow of the Not-Playable Characters was appropriate." The experiment lasted for around one hour, and the participant received monetary compensation of 15 Euros.

Participants Twenty participants ($M_{age} = 26.05$, $SD_{age} = 3.62$; $male = 15$, $female = 5$) took part in our study. We recruited the participants using our institutional mailing lists and social networks and using convenient sampling. We did not recruit participants that experienced intense physical activity or consume any caffeine or nicotine in the 3-hour pre-study period, according to Babaei et al. [26]. None of the participants reported a history of neurological, psychological, or psychiatric symptoms. However, due to technical difficulties,

Table 4.3: Results of repeated-measures ANOVA or Friedman test on Behavioral Performance, subjective questionnaires and physiological arousal measures across conditions of stable STREAM.

	Stream 24		Stream 110		Stream 191		Stream 270		Stream 347		ANOVA / Friedman	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*p*	F/χ^2
N-Back Accuracy [%]	88.14	6.22	87.12	87.53	88.21	5.90	87.33	5.52	83.51	9.55	.64	2.49
Raw NASA TLX	42.94	13.49	48.87	17.64	46.57	17.60	50.0	16.63	55.59	17.98	<.01	15.6
SCL	−.177	.32	.14	.18	−.002	.25	−.17	.322	−.02	.32	<.01	15.24
nsSCRs	−.002	.003	−.001	.002	.00	.004	−.002	.003	−.002	.004	.07	8.37
GEQ—Competence	2.12	.86	2.38	.93	2.47	.7	2.0	0.90	2.18	.88	<.05	.07
GEQ—Pos. Affection	1.50	1.03	1.24	1.17	1.41	1.15	1.29	1.19	1.65	1.34	.81	1.54
GEQ—Immersion	.94	.92	.62	.88	.85	.98	1.13	1.29	1.27	1.65	.46	3.59
Stream Appropriate	3.53	2.07	3.12	2.15	2.94	1.91	3.35	3.35	3.53	2.18	.44	.44
Desire To Use	4.18	2.07	3.82	2.27	3.77	2.11	4.06	2.05	4.12	2.37	.57	2.91

after EDA signal inspection, we removed 3 participants, leading to N = 17. According to the fast-track conditions of the local institutional ethics board, the study qualified as fast-track, i.e., participants are not subject to any risk (e.g., deception, stress beyond normal levels, recording of sensitive information).

Results

We analyzed indicators of physiological arousal, performance, and subjective experience across the adaptive and non-adaptive conditions. We first briefly present the results for our system's non-adaptive levels of STREAM and then investigate whether the adaptive systems can support performance, increase user experience, and relatively result in a lower level of perceived workload. Normality of residuals for ANOVA models was checked via D'Agostino normality test [128] as Shapiro-Wilk test is too sensitive for n >50. Upon normality testing, we use one-way repeated measures ANOVAs (RM ANOVAs) for parametric analysis or a Friedman's rank sum test for not-normally distributed data. Furthermore, for post-hoc comparisons, we use Conover's tests [121]. In this section, we perform statistical testing across all conditions i.e., stable STREAM and adaptive STREAM conditions.

Stable Stream Conditions

Behavioral Performance We conducted a Friedman test, as the D'Agostino normality test showed that performance data are not normally distributed ($\chi = 9.65$, $p < .05$). We computed as accuracy the proportion of times the sphere was placed in the correct bucket overall. This analysis revealed that the accuracy in the N-Back task was not significantly influenced by STREAM ($\chi = 2.49$, $p = .64$), see Table 4.3.

Electrodermal Activity Results **Skin Conductance Level** Due to a D'Agostino tested normality assumption satisfaction ($\chi = 2.93$, $p = .23$), a RM ANOVA showed that average SCL was significantly influenced by STREAM ($\chi = 2.49$, $p < .05$). Pairwise comparisons via a Conover test with Holm's correction showed that the highest condition of STREAM (347) significantly decreased SCL as compared to the lowest condition of STREAM (24), while STREAM 110 showed increased SCL as compared to the condition where STREAM was equal to 270 NPCs entering the scene per minute (all $p =< .005$).

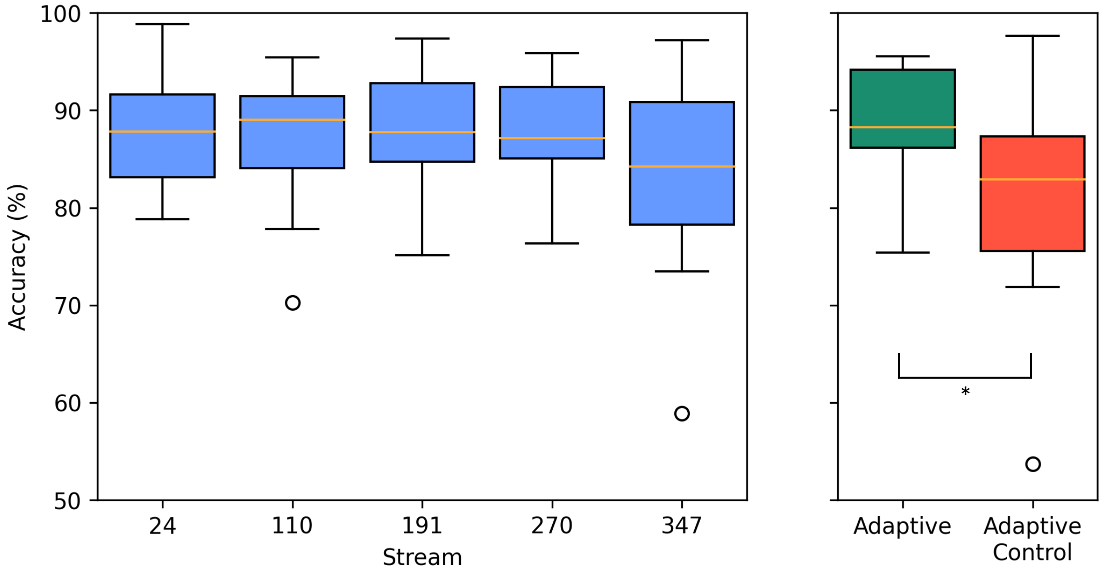


Figure 4.16: Box-plots for accuracy in the N-back task as a function of STREAM in the stable and for adaptive conditions, respectively depicted on the right and on the left . Here, we did not find any significant differences across Stable STREAM conditions. In the adaptive conditions, the adaptive control showed significantly decreased accuracy on an average of 7.59% less than the adaptive condition.

nsSCRs Since the normality assumption was violated ($\chi = 15.78$, $p < .05$), we conducted a Friedman test indicating that the average amplitude of nsSCRs is significantly influenced by STREAM ($\chi = 8.37, p = .07$). A Conover's post-hoc test did not show any significant differences between pairs.

Subjective Results We summarized the subjective for results NASA, GEQ and UX in Table 4.3.

Perceived Workload The raw NASA-TLX scores did not meet the normality assumption (D'Agostino: $\chi = 64.74, p < .001$). We, thus, performed a Friedman test, reporting a significant effect ($\chi = 15.65$, $p < .001$). The Conover's post-hoc test revealed that the stable condition with the highest STREAM (347) showed significantly increased workload as compared to the lowest condition of STREAM (24) and to STREAM (190) (all $p =< .005$). Results are depicted in Figure 4.19.

GEQ Competence. Since GEQ-Competence subscale ratings were not normally distributed (D'Agostino: $\chi = 50.28$, $p < .001$) we performed a Friedman test that detected a main effect of STREAM ($\chi = 11.39$, $p < .05$). Conover's post-hoc revealead only one significant difference in the pairwise comparison between. Specifically, participant felt less competent in condition STREAM (270) as compared to condition of decreased STREAM (191). For the the descriptive statistics (M & SD) see Table 4.3 and Figure 4.20. **Positive Affection.** GEQ-Positive Affection ratings were not normally distributed ($\chi = 14.46$, $p < .001$). Thus, we used a Friedman which did not show any effect of STREAM ($\chi = 1.53$, $p = .81$). **Immersion.** Similarly, the

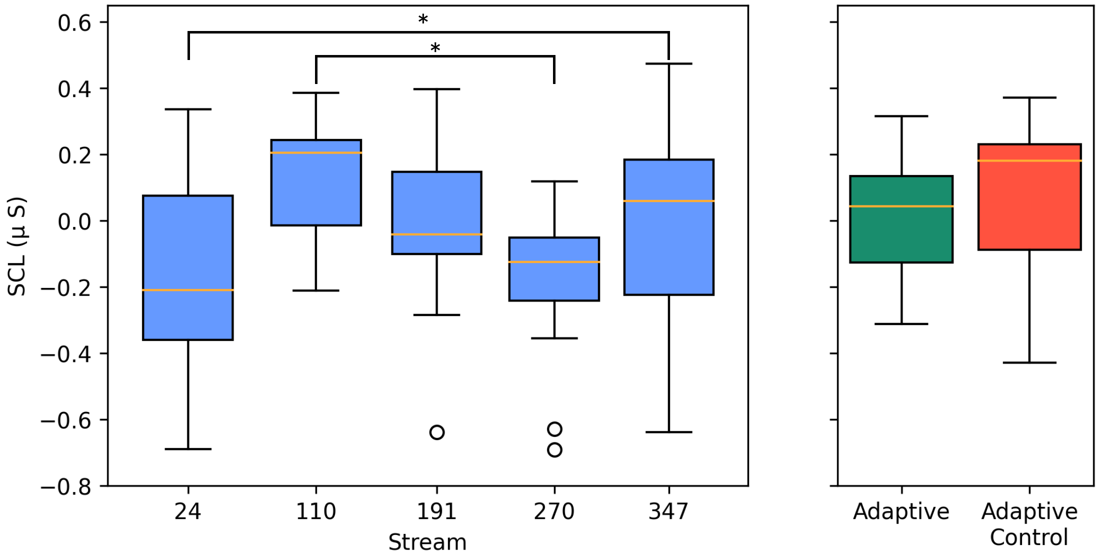


Figure 4.17: Box-plots for SCL as a function of STREAM. Results for stable manipulation of STREAM are shown on the left while adaptive and adaptive control conditions are depicted on the right. We found that STREAM (347) significantly increased the SCL as compared to STREAM (24) and that the second lowest condition of STREAM (110) was significantly higher than STREAM (270).

GEQ-Immersion subscale was not normally distributed ($\chi = 35.66$, $p < .001$) and did not show any significant effect of STREAM for the Friedman test ($\chi = 3.59$, $p = .46$).

User Experience. For both UX ad-hoc surveys, no significant differences were found employing a Friedman test. Both Stream ($\chi = 12.68$, $p < .005$), and Desire to Use ($\chi = 31.14$, $p < .001$) deviated from the normal distribution.

Adaptive Stream Conditions

To further inspect the effect on physiological arousal, performance, and subjective experience of the two adaptive systems, we conducted t-tests or Wilcoxon tests depending on the normality assumption tested on residuals. Results are summarized in Table 4.4. On average, STREAM for the adaptive condition lead to an average of 241.88 NPCs per minute (SD=30.31), while in the Adaptive Control condition, STREAM stabilized on average of 216 NPCs (SD=32.82).

Behavioral Results Given a violation of the normality assumption ($W = .91$, $p < .05$), a Wilcoxon test detected a significant difference in the behavioral accuracy for the N-Back, showing increased performance in the adaptive condition ($v = 135$, $p < .005$).

Table 4.4: Results of Pairwise comparison for behavioral performance, subjective questionnaires and physiological arousal measures between the two adaptive conditions.

	Adaptive		Adaptive Control		*t*-Test/Wilcoxon		
	M	***SD***	***M***	***SD***	***t/V***	***p***	***d/r***
N-Back Accuracy [%]	88.56	5.96	80.97	10.07	135	**<.05**	.63
Raw NASA-TLX	42.16	16.34	62.55	17.99	−6.17	**<.001**	−0.37
SCL	.015	.17	.072	.24	−.79	.44	−.30
nsSCRs	−.002	.002	<.001	.004	−2.33	**<.05**	−.36
GEQ—Competence	2.15	.95	2.29	.89	17	.55	.12
GEQ—Positive Affection	1.41	1.03	1.50	1.40	23.5	.95	−.34
GEQ—Immersion	.67	.96	1.09	1.25	9	.22	−.50
Stream Appropriate	3.23	2.04	3.47	2.32	8.5	.75	−.14
Desire To Use	3.41	2.21	3.53	2.29	21	.71	−.33

Electrodermal Activity Results **Skin Conductance Level.** Due to a Shapiro-Wilk not significant results ($W = .96$, $p = .71$), a t-test showed that average SCL was not influenced by STREAM ($t = -0.79$, $p = 0.44$).

nsSCRs. Similarly to SCL, also the distribution of nsSCRs was normally distributed ($W = 0.97$, $p = .84$), therefore a t-test showed that the amplitude of nsSCRs, indexing sympathetic physiological arousal, was impacted by STREAM as indicated by an an increased in the adaptive

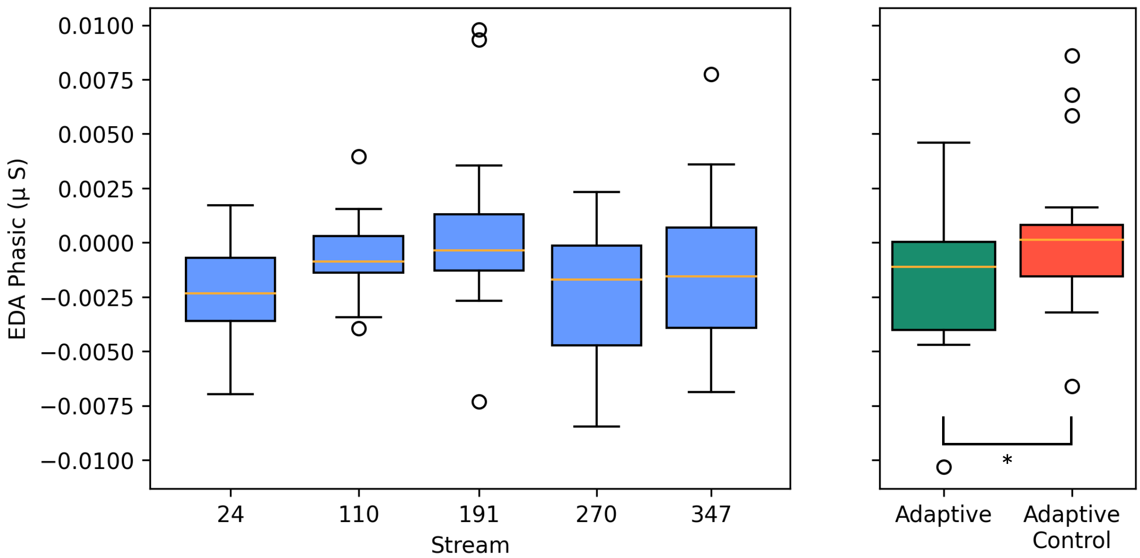


Figure 4.18: Box-plots for mean amplitude of non-specific Skin Conductance Responses (nsSCRs) as a function of STREAM. Here, we show the stable conditions on the left while the adaptive conditions are on the right. No significant differences were detected in pairwise comparison across stable conditions. In the adaptive conditions, the adaptive control condition showed significantly increased mean nsSCRs amplitude as compared to the adaptive condition .

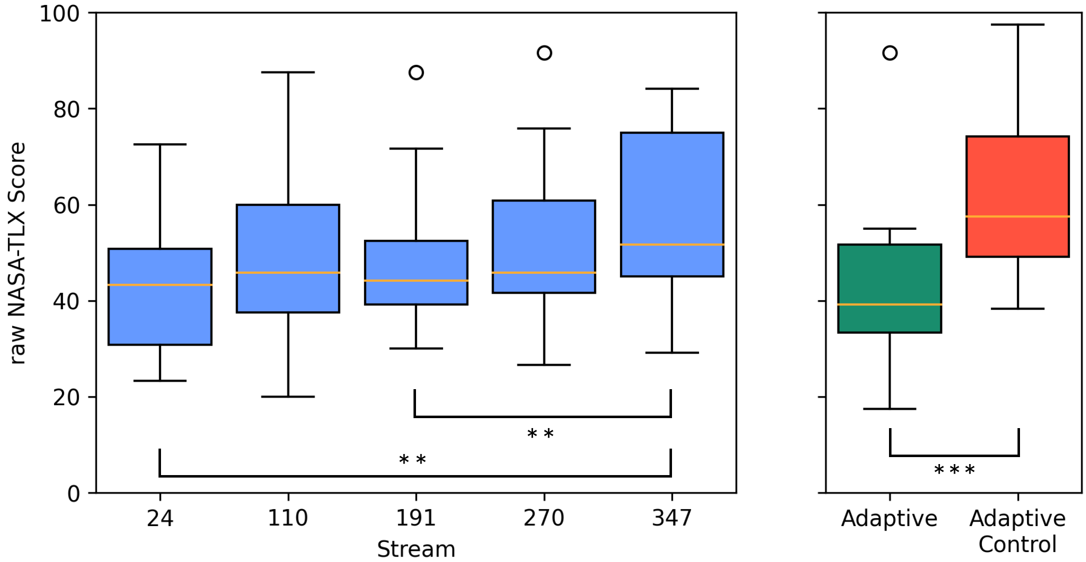


Figure 4.19: Box-plots for raw NASA-TLX scores as a function of STREAM. Stable STREAM manipulation is shown on the left, and the adaptive conditions on the right. Those results mirrored the one for N-Back task accuracy. Here, participants reported significantly increased perceived workload in the condition of highest STREAM compared to the lowest condition of STREAM. We report similar results for the adaptive control condition that showed increased perceived workload compared to the adaptive condition.

control condition ($t = -2.33$, $p < 0.05$), see Table 4.4.

Subjective Results **Perceived Workload** The control adaptive condition significantly increased subjective workload, given a normal distribution ($W = .96$, $p = .79$), as shown by a t-test ($t = -6.18$, $p < .001$), see Table 4.4 and Figure 4.19.

GEQ Subscales. We did not find any significant results across all three GEQ subscales, i.e., Competence, Positive Affection, and Immersion; see Table 4.4 and Figure 4.20.

User Experience. The results for the ad-hoc UX surveys were similar to those of the GEQ subscales. Both items were non-normally distributed. However, no significant differences were reported between the two adaptive conditions; see also Table 4.4.

Summary of Results

We first investigated the effect of different stable levels of visual complexity, i.e., STREAM of NPCs, while participants were engaged in a WM Visual N-Back task over a series of dependent variables: behavioral accuracy, physiological arousal (SCL and nSCRs amplitude) and subjective workload, engagement, and UX. Second, we compared Adaptive STREAM Conditions, where we adapted the visual complexity based on variations of SCL.

Stable Stream Conditions We analyzed the effect of different levels of visual complexity, i.e., the amount of NPCs entering the VR scene per minute (STREAM). We found no significant effect of STREAM over behavioral accuracy in stable conditions. Regarding physiological arousal, we found that SCL was decreased in higher STREAM conditions (347 and 270 NPCs per minute) as compared to the lower ones (24 and 110 NPCs per minute). Finally, participants reported significantly higher perceived workload in the condition with the highest STREAM (347 NPCs) compared to all the other conditions. When investigating perceived engagement, we found that participants reported feeling less competent (GEQ - Competence) when the STREAM decreased from 191 to 270 NPCs per minute. No further significant differences were detected.

Adaptive Stream Conditions Based on the MIM model [607], we designed two physiologically-adaptive systems with two opposite architectures, i.e., adaptive and adaptive control systems. Overall, the adaptive system resulted in improved behavioral accuracy, decreased physiological arousal, as shown by decreased nSCRs amplitude, and decreased perceived workload. No specific significant differences were detected in perceived engagement and UX between the two systems. Our results revealed how the adaptation of visual complexity based on physiological arousal, following the MIM model, results in improved WM performance compared to a reversed adaptation.

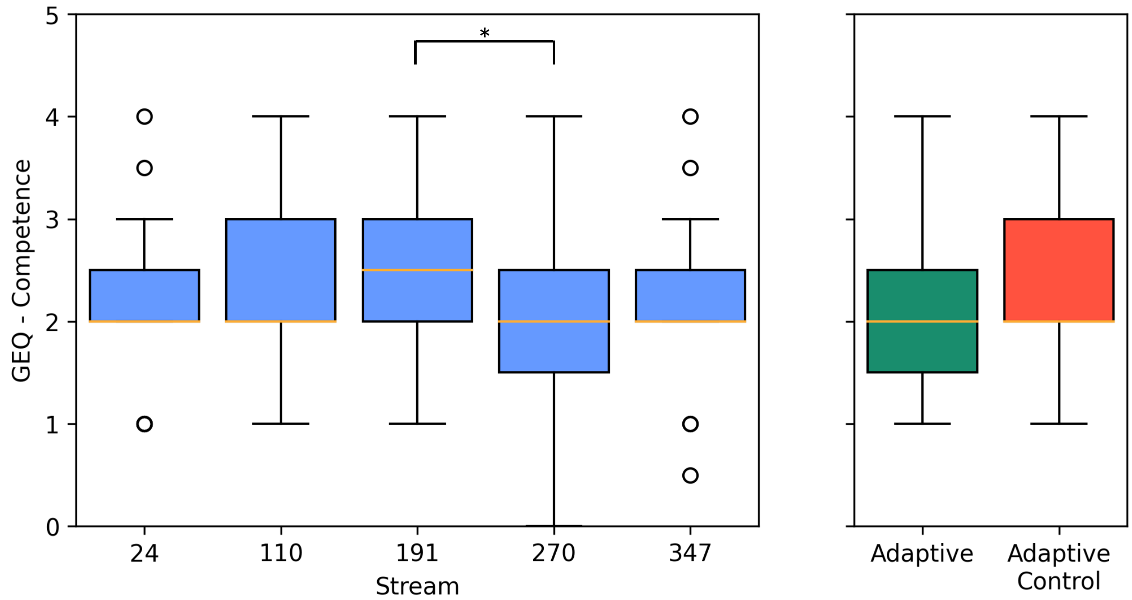


Figure 4.20: Box-plots for GEQ-Competence scores as a function of STREAM. Stable STREAM manipulation is shown on the left and the adaptive conditions on the right. We did not find any significant differences across GEQ-Competence ratings.

4.3.4 Discussion

We presented a physiologically adaptive VR system that employed electrodermal activity to perform dynamic visual complexity adjustments to enhance task performance. We evaluated the effect of visual complexity, in the form of NPCs, on task performance, SCL, perceived workload, engagement, and UX. In the study, participants performed an N-Back task recruiting working memory resources. Here, we discuss our results regarding the outcome of our adaptive algorithms for modeling visual complexity and its effect on EDA. Then, highlight applications for arousal detection and implications of physiologically-adaptive systems.

Mapping Visual Complexity to Electrodermal Activity

Evaluating autonomous, closed-loop control systems raises issues for physiologically-adaptive systems design. The relationship between psychophysiological data and triggering adaptive responses at the interface necessitates careful design. One stage of this process is the generation of valid input measures and effectively categorising psychophysiological data in real time.

Our results indicate that we could model the visual complexity of the surrounding VR environment user in the N-back task using SCL. Participants who interacted with the adaptive system performed better than non-adaptive. In the reverse adaptation, we saw an inefficient user model, which led to a potentially harmful state of overload. In both of our adaptive conditions, participants' performance was influenced by the STREAM adaptation at a behavioral, physiological, and subjective level. We found reciprocal effects on performance levels (see Figure 4.16), significant differences in the measure of physiological arousal, i.e., nsSCRs (see Figure 4.18) and lastly, in perceived workload (see Figure 4.19).

Impact of Visual Complexity on Task Performance

In line with the findings of previous work on the effect of visual complexity [9, 167, 462], we report that visual WM performance is decreased with an increasing amount of distractors surrounding the user when adapting the visual complexity based on arousal, following the MIM model. However, we did not find any main effect of STREAM in stable conditions. These results might be due to the high individual variation and sensibility to the variation of STREAM, pointing even more towards the necessity to individualize and adapt visual complexity to optimize task performance. Participants significantly performed worse when the amount of NPCs was modulated following an increase in physiological arousal, as shown in the adaptive control condition. Furthermore, we imposed capacity limits on visual WM memory when increasing the visual information load and the number of distracting objects. Finally, we add to results on social crowdedness [141, 353] and the distracting and perceptual demands that an increasing amount of human, tough virtual, place on their observers [334, 335, 356, 521].

Impact of Visual Complexity on Physiological Arousal

In the adaptive control condition, once our physiologically-adaptive function detected increased arousal, we increased the visual complexity, adding visual distractors to the surroundings. We found a main effect of STREAM in the stable condition, showing how visual complexity can impact users' tonic component of arousal. However, we did not find an effect on the phasic component as this component might have encountered habituation effects over time, giving the continuous visual stimulation [410]. On the other side, we found an increased amplitude of the nsSCRs in the adaptive control condition. nsSCRs underlie the tonic stress produced during a sustained stimulation period, which resembles our experimental setting as participants were continuously exposed to variations of STREAM of NPCs. These results can be explained given the bottom-up, involuntary orienting of attention that is driven by novelty detection and inevitably coupled with an automatic shift of attention [397, 491, 526] of attentional control caused by the distracting STREAM of NPCs.

Furthermore, amplitude of nsSCRs has been associated with increased workload [63] in the N-back task [183, 385], in biofeedback [429] and in adaptive automation settings [65]. Therefore, increasing the visual complexity every 20sec might have progressively forced participants to reallocate attentional resources, related to the phasic component of electrodermal activity [231, 480]. This claim is also supported by the decrease in task accuracy.

Although SCL, slower than nsSCRs variations [221], failed to discriminate between the two adaptive algorithms, we can still confirm that the general physiological arousal was affected by STREAM manipulations as shown in the Stable STREAM conditions results. Therefore, we were able to model the user state sufficiently enough in a real-time environment to either decrease or improve performance by modulating distracting features of the environment based on variations of physiological arousal. The subjective ratings of perceived workload also support this claim. The condition of the highest visual complexity corresponded to the highest ratings in the NASA-TLX and the adaptive control condition. This is in line with previous work that showed either fluctuation of SCL [110, 165, 317] or phasic components [183] aligned with subjective measures of workload. Finally, we report a series of not significant differences over GEQ subscales. Even though this might be counterintuitive, we might argue that in our case the duration of each task (6 minutes) was not long enough to elicit a change in subjective engagement scores, as a high-demanding task that showed decreased engagement after steps of 20 minutes [179]. Future work should investigate the effect of time on task over subjective engagement and on adaptation to verify if our system can replicate previous work on this relationship.

Relationship Between Visual Complexity, Mental Workload, and Physiological Arousal

The results from NASA-TLX suggest that our manipulation of visual complexity through variations in the number of NPCs did not map onto systematic variation in perceived workload. To reiterate, we found significant differences between the highest (STREAM = 347) and intermediate (STREAM = 191) and lowest levels of visual complexity (STREAM = 24). In contrast, no significant difference was found between the remaining intermediate levels. While there was no one-to-one mapping of perceived workload and number of NPCs, our findings thus still support the overall trend that increasing visual complexity leads to a higher perceived workload. Furthermore, the post hoc tests indicated that the relationship between visual complexity and perceived workload might not be strictly linear, which suggests that more sophisticated workload manipulation strategies could be explored in future research. The nonlinearity in the relationship between visual complexity and perceived workload could be attributed to factors such as individual differences in cognitive processing abilities [495] or attentional resource allocation [189, 371], which may interact with the level of visual complexity to influence workload perception in a complex and dynamic manner. It is possible that the variations in visual complexity did not affect everyone similarly. Any variability in the overload point may be individually important for individuals. Therefore, user-dependent models [13, 143] are needed to evaluate visual complexity, which further speaks towards user-adaptive environments.

When examining physiological arousal, we found that SCL does not evolve continuously with the number of NPCs, i.e., visual complexity. SCL follows a Bateman distribution and does not have a linear behavior [508]. The Bateman distribution is a biexponential function that describes the relationship between the intensity of a stimulus and the magnitude of a physiological response [42]. The Bateman distribution is characterized by a sigmoidal shape, with a steeper slope at low levels of arousal and a flatter slope at high levels of arousal. This means that small changes in the intensity of a stimulus can elicit a larger physiological response at low levels of arousal, while a larger shift in stimulus intensity is required to elicit the same response at high levels of arousal [508]. Therefore, increasing the number of NPCs may not necessarily lead to a proportional increase in SCL, as the response may saturate at higher arousal levels which must be considered when using physiological arousal as input for adaptive systems.

Limitations and Future Work

While our study has provided valuable insights into the efficacy and applicability of the physiologically adaptive systems for visual complexity adaptation in VR, several limitations need to be addressed to enhance the validity and generalizability of our results and improve the architecture of future adaptive systems.

We discuss different approaches for EDA online preprocessing and how multimodal physiological input, e.g., ECG and EEG, can be implemented in adaptive systems to gain a more comprehensive understanding of the relationship between physiological responses and visual complexity. The generalizability of our results would also benefit from testing the feasibility of machine learning algorithms applied to EDA data and incorporating placebo or randomized conditions in the experimental design. Finally, we reflect on how ethical and privacy considerations should be implemented when designing physiologically adaptive systems.

Approaches for Online EDA Preprocessing and Machine Learning for Arousal Prediction Even though we chose and implemented an established standardized method for EDA preprocessing, different preprocessing steps can be evaluated for adaptation purposes. For example, unusual steep rises might arise in ambulatory settings and impair adaptation. Here we used a high-pass filter that might compensate for artifacts and distort the entire EDA time series recording, considering high values. Therefore either adaptive filtering or thresholding might improve the quality of the physiological input to the adaptive system. For example, Wavelet transforms employ a Gaussian mixture distribution to model tonic and phasic components of EDA. Chen et al. [91] compared this approach to previous approaches and showed higher performance in artifact reduction.

Second, generalizability to the largest population is a crucial goal, and EDA suffers from significant interindividual variability. People with higher SCL frequently have more SCRs, and larger amplitude nsSCRs [64, 565]. Thus, using an a priori threshold for detecting nsSCRs might lead to either an inflated chance of false positives or negatives. Fixed thresholding might lead to more detected nsSCRs, and a higher nsSCRs frequency in populations with higher SCL as peaks with a low amplitude will be considered as nsSCRs. Again, this outcome is especially relevant for adaptation purposes, as if the adaptation impacts the arousal detection and VR adaptation, leading to malfunctioning adaptations. Hence, adaptive thresholding can improve sensitivity to detect phasic changes in EDA for users with high SCL variability. Kleckner et al. [310] showed the feasibility of such an approach, using a novel fixed plus adaptive threshold proportional to the SCL.

Third, another possible approach is to investigate the application of machine learning algorithms to adapt the visual complexity, predict physiological arousal increase, and counteract the detrimental effect on task performance. For example, naive Bayes and Support Vector Machines successfully classified high and low performers in a VR Stroop Task based on behavioral data [17]. Therefore testing EDA features classification jointly with behavioral measures might lead to improved algorithm accuracies.

Multimodal Adaptive Systems and Placebo Control Condition First, integration with other physiological inputs, such as EEG or ECG, would allow us to account for and monitor different cognitive or emotional states that might be impacted by visual complexity or other adaptations. In this regard, it is worth mentioning how the EEG Alpha band is associated with an increase in arousal activation [103, 349, 489], flow experience [127, 324] and internal attention states [120, 364] which are of central interest for adaptive systems aiming at supporting task performance and engagement. Another example that involved NPCs is the work by Keynan et al. [305], which demonstrated how virtual human avatars, responsive to EEG signals, can guide neurofeedback training for stress resilience. Hybrid Brain-Computer Interfaces [193] and sensor fusion approaches should be further explored, opening wide applications and allowing for deeper and fine-grained detection of user states.

Second, a potential methodological improvement that could be implemented in future studies is comparing stable and adaptive streams. This could be achieved by including two additional conditions, a randomly changing STREAM as a "placebo" condition and no STREAM as a control condition. When studying the efficacy of an adaptive system, it is essential to compare it to a placebo condition to ensure that the observed effects are not simply due to placebo effects [320]. However, in our work, participants were unaware of which condition they were experiencing, minimizing any possible placebo effects. The control condition, on the other hand, would allow us to assess the impact of the stream itself on the outcome measures without the confounding influence of the adaptive STREAM.

Ethical and Privacy Considerations in Physiological Computing Physiological computing systems, designed to tap into private psychophysiological events for dynamic MR interactions, raise ethical concerns regarding informed consent, manipulation of users' states, and privacy [178]. Users must be fully aware of how physiological data are collected, used, and shared, specifically when the data are employed for model training and validation. To prioritize ethical design, researchers should inform participants which physiological state the system is optimizing for and allow users to return to a neutral affective state if the final state is perceived as undesirable [232]. Note, however, that verbal descriptions of system functionality might come with the problem of placebo effects in evaluation [320].

Physiologically-adaptive systems are grounded on symmetrical interaction between users and systems. Here, privacy concerns on data usage and protection might arise. The individual should retain formal and legal ownership of psychophysiological data, and any third party should receive access to such information only with user approval [232]. This is specifically relevant, as psychophysiological data might underlie specific cognitive or affective states and could be used for secondary medical diagnostic purposes. To mitigate these concerns, a privacy-by-design approach should be used, embedding privacy considerations into every stage of the design and development process [101, 178]. This includes conducting privacy impact assessments, implementing privacy-enhancing technologies, and collecting privacy-preserving data in every implementation stage of physiologically adaptive systems.

4.3.5 Open Science

We encourage readers to reproduce and extend our results and analysis methods. Therefore our experimental setup, collected datasets, and analysis scripts are available on OSF (`https://osf.io/axvfy/`).

4.4 Study 7: Investigating Physiological Correlates of Secondary Task Difficulty Adaptations

4.4.1 Method

We utilized the open dataset from Chiossi et al [108] as it comprises behavioral, physiological (EEG, ECG, and EDA), and subjective data. We chose this multimodal dataset to employ one physiological channel, i.e., EDA, as inputs for an adaptive system, allowing for analyzing the effect of adaptations of other physiological signals (ECG, EEG). The dataset is available on the Open Science Framework at `https://osf.io/axvfy/`. We refer the reader to their paper for a detailed task implementation and data collection description. The dataset included 20 participants with an average age of 26.05 years ($SD = 3.62$). EDA data (sampled at 250 Hz) were collected via a GSR module (BrainProducts GmbH, Germany). ECG data (sampled at 130 Hz) were recorded using a Polar H10 chest strap (Polar, Finland) with electrodes moistened before data collection and placed over the xiphoid process of the sternum. EEG data were recorded at a sampling rate of 250 Hz using a 7-channel dry electrode cap embedded in the HTC VIVE headset from Wearable Sensing (DSI-VR 300, San Diego, CA, USA). The electrode positions followed the 10-20 system, including FCz, Pz, P3, P4, PO7, PO8, and Oz. Electrode impedances were maintained below 20 kΩ, with electrodes linked to the ears as a reference for the EEG recording. The EDA, ECG, and EEG data were simultaneously recorded using the LabStreamingLayer (LSL) framework[3].

Research Questions

We aim to gauge the users' responses to visual complexity adaptations and levels by assessing various physiological measures. This evaluation allows us to measure the impact of these changes on different physiological indicators. This will help to determine whether an adaptive system based on EDA can effectively include multiple modalities and be adapted for different applications. Moreover, we want to investigate the effect of different levels of visual complexity on attention allocation, engagement, and task load physiological correlates.

Our analysis contributes to developing physiologically-adaptive VR systems and drafts new possibilities for a larger input space and evaluation for immersive adaptive environments with varying levels of visual complexity. Drawing on the principles of physiologically-adaptive VR systems and existing research, we put forth the following hypothesis and research questions:

- **HP1** As SCL, i.e., tonic component of EDA, was employed as an input for adaptation, we expect it to be impacted by the system adaptation as a validation for the system architecture.
- **RQ1** Do adaptations of visual complexity that use EDA as input, impact participants' stress levels, as indexed by HR and HRV?

[3] `https://labstreaminglayer.readthedocs.io/info/intro.html`

- **RQ2** Do adaptations of visual complexity that adjust the number of distractors, using EDA as input, influence internal and external attention, as measured by alpha and theta EEG oscillations?

Secondly, Chiossi et al. [108] provided an initial insight into the relationship between visual complexity, physiological arousal, and behavioral performance. Thus, we expand their work by evaluating the effect of varying stable levels of visual complexity on different physiological measures. Thus we hypothesize:

- **RQ4** Does an increase in visual complexity, i.e., distracting information, increase external attention resources as indexed by a decrease in EEG Alpha oscillations [364]?
- **RQ5** Does an increase in visual complexity, i.e., distracting information, increase the cognitive workload as indexed by the ratio between alpha and theta oscillations [468]?

Task

The experimental task employed in this study was adapted from the N-Back task, as described in [108]. Participants were immersed in a neutral VR environment, where they were presented with a marble-like pillar and two buckets positioned on the left and right sides, respectively. Spheres of different colors (green, red, blue, and black) were generated and appeared on the pillar randomly, based on previous work [382]. Participants were required to use an HTC VIVE controller to grab the spheres and place them into the appropriate buckets. The rule for placing the spheres was based on matching the color of the current sphere with the color of the sphere presented two steps prior. If the colors matched, the sphere would be placed in the right bucket. Conversely, if the colors did not match, the sphere would be placed in the left bucket. Participants had a time window of 4 seconds to pick up the sphere to avoid making an error. New spheres appeared when the current sphere was successfully placed in one of the buckets or after the 4-second time limit had elapsed.

4.4.2 Adaptive Systems Architecture

Chiossi et al. [113] introduced an adaptive system that supports users' engagement and task performance by adapting the visual complexity based on EDA. Our focus was on the two adaptive conditions, where we divided the EDA, ECG, and EEG signals into 20-second epochs corresponding to the periods when the stream of non-player characters (NPCs) underwent adaptations. The first adaptive system, i.e., Adaptive Test System is based on the Motivational Intensity model (MIM) [477]. According to this model, when task demands are perceived as achievable, there is a proportional relationship between mental effort and task demand [477]. However, as task demands increase and success becomes less likely, the investment of effort decreases while the perceived workload increases. This results in impaired performance, increased perceived workload, and reduced engagement, highlighting

(a) STREAM = 24 (b) STREAM = 191 (c) STREAM = 347

Figure 4.21: VR capture of a single trial of the VR n-back from a birds-eye perspective in the first row. In (a) is depicted the condition with low visual complexity with STREAM = 24. In (b) is depicted the condition of moderate visual complexity with STREAM = 191, and lastly in (c), the highest visual complexity with STREAM = 347.

the association between working memory capacity and these outcomes. Thus, the ADAPTIVE TEST SYSTEM decreases the visual complexity when the arousal increases, as indexed by Skin Conductance Level (SCL), by removing 8 NPCs and increasing it by adding 16 NPCs when a decrease in SCL is detected. The second system, i.e., REVERSE ADAPTIVE SYSTEM, followed an inverse logic and served as a control condition. Here, the system either aims to progressively increase the task demands by adding 16 NPCs when the participant's arousal is increasing or when the arousal is decreasing, removing visual complexity from the VR scene (-8 NPCs), ultimately leading to either an overly distracting or empty VR scene, to decrease users' engagement. The VR-physiologically adaptive system performed an average of $M = 6.94$ adaptations ($SD = 2.77$) for the ADAPTIVE TEST condition, while the REVERSE ADAPTIVE SYSTEM $M = 5.19$ ($SD = 2.76$) adaptations.

Adaptive Mechanism

Both systems employ a rolling window approach for adaptation. They utilize two distinct data windows for SCL analysis: a 180-second window (W_1) for low-frequency changes and a 30-second window (W_2) for high-frequency changes. The SCL levels are averaged over these windows to stabilize the value, using an epsilon parameter for smoothing.

Slope Analysis

Slopes of SCL changes in these windows (s_1 and s_2) are computed, forming the basis for adaptive decision-making. s_1 is calculated from the average tonic value between points t_{-2} and t_0, while s_2 uses values between t_{-1} and t_0.

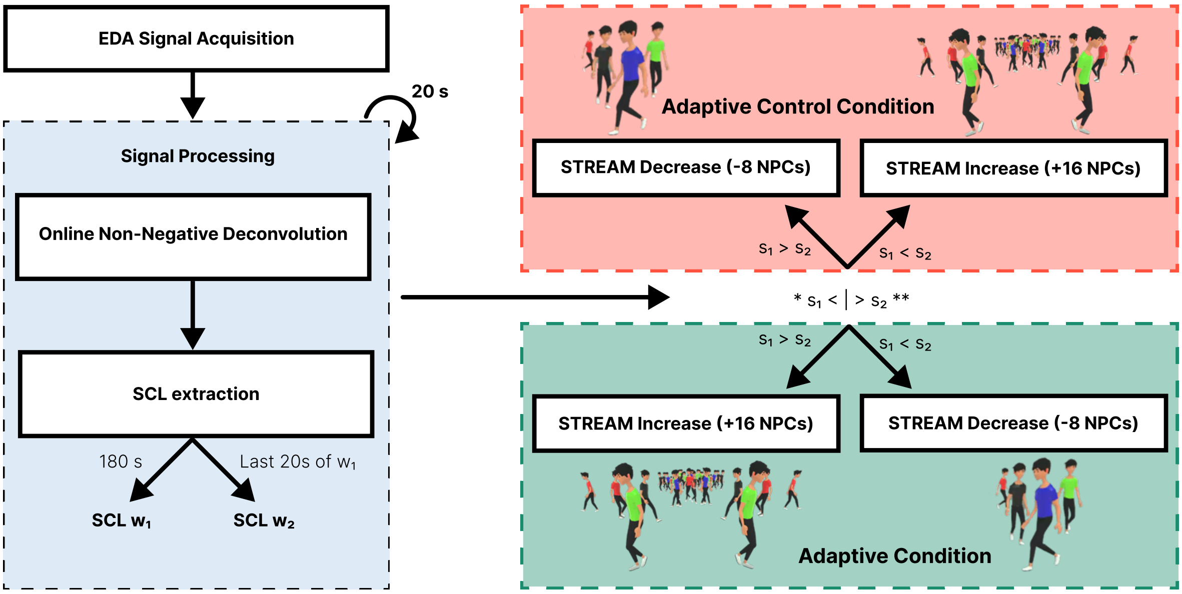


Figure 4.22: Components of the two adaptive systems. In both systems, the STREAM of NPCs adapts according to changes in the slope between the average SCL_{W1} and SCL_{W2} . Here, SCL_{W1} represents the average SCL between of W_1, which corresponds to 180 seconds time window, while SCL_{W2} corresponds to the SCL in the last 20 seconds of $W1$. In the TEST ADAPTIVE SYSTEM, if the SCL in W_1 is smaller than SCL in W_2 (SCL_{W1} < SCL_{W2}), then 8 Not Playable Characters (NPCs) are removed from the scene. Alternately, in the case SCL_{W1} > SCL_{W2}, 16 NPCs are added. The motivational intensity model inspires adaptive conditions and aims to support task engagement. The REVERSE ADAPTIVE SYSTEM follows the opposite logic. Figure adapted from [108].

Rule-Based Adaptation

The adaptation decision is based on comparing the low-frequency slope (s_1) to the high-frequency slope (s_2), adhering to a threshold parameter θ to ensure stability in adaptations. This comparison drives the system to increase or decrease task difficulty, as detailed in Equation 4.5. This adaptive process occurs every 20 seconds, ensuring timely responsiveness to physiological changes.

$$adaptation(s_1, s_2) = \begin{cases} \texttt{increase} & \text{if } s_1 \leq s_2 - \theta \\ \texttt{decrease} & \text{if } s_1 \geq s_2 + \theta \end{cases} \tag{4.5}$$

4.4.3 Data Preprocessing and Analysis

We investigated the physiological indicators of cognitive workload and arousal in a visual working memory (WM) task, while adaptive systems dynamically adjusted the visual complexity.

EDA & ECG Preprocessing

For EDA data preprocessing, we employed the Neurokit Python toolbox [367]. This involved a third-order Butterworth high-pass filter at 3 Hz and nonnegative deconvolution analysis [43] to separate tonic and phasic components. We computed average amplitude of nsSCRs and tonic SCL, with nsSCR peaks identified using a $.05\mu S$ threshold. ECG data were processed in the time domain, focusing on HR and HRV (RMSSD). We applied a 3-45 Hz (3rd order) FIR band-pass filter using Neurokit [367] and used Hamilton's method for signal segmentation and QRS complex identification, extracting mean HR and HRV.

EEG Preprocessing

We processed the EEG raw data via the MNE Toolbox [216]. EEG data were recorded with a sampling frequency of 250 Hz from dry electrodes placed on Fz, P3, Pz,P4, PO7, Oz, PO8 locations (10/20 system), with a reference set at linked earlobes. We notch-filtered the signal at the power frequency of 50 Hz and then band-passed between 1 and 70 Hz to remove high and low-frequency drifts. We referenced the data to the common average reference (CAR). Next, we computed an independent component analysis (ICA) with extended infomax algorithgm for automatic artifact detection and correction with the ICLabel plugin [443]. We then analyzed the preprocessed EEG data in two frequency bands: Theta (4–8 Hz) and Alpha (8–12 Hz), and Beta (12-30 Hz), using Welch's method. We computed alpha oscillation for posterior sites, i.e., PO8, PO7, and Oz electrodes, and extracted Theta and Beta oscillations from midline sites, i.e. Fz and Pz. Moreover, we computed the ratio of midline theta activity's absolute power to posterior alpha activity's absolute power as an implicit measure of workload [468].

4.4.4 Results

We report quantitative findings from analyzing physiological (EDA, ECG, and EEG) data from the dataset. We employed a Repeated Measures analysis of Variance [4] (RM-ANOVA) for adaptive levels of visual complexity and a Linear Mixed Model (LMM) approach for EDA, ECG, and EEG measures for stable visual complexity. To account for the repeated-measures structure in the data, we included a random intercept for each participant in our model.

Adaptive Visual Complexity

EDA Results **SCL.** RM-ANOVA test detected a significant effect of the factor *Adaptive System* on the SCL ($F(1, 177) = 22.447$, $p < .001$). Tukey posthoc test showed that participants who experienced an INCREASE in the TEST adaptive system condition had a significantly lower SCL compared to those in the REVERSE adaptive system when experiencing a STREAM DECREASE ($PMM = -11506$, $SE = .458$, $p = .003$). When contrasting an INCREASE in the TEST

[4]The predicted marginal means (PMM) for the different levels of the variable 'Adaptive System' and 'Stream Adaptation' were calculated using a Kenward-Roger degrees-of-freedom method.

adaptive system with a DECREASE in the same system, participants had a significantly higher SCL ($PMM = -1.515$, $SE = .41$, $p = .002$). Finally, when the REVERSE adaptive system performed a Stream INCREASE, the SCL increased as compared to the DECREASE in the same system ($PMM = -2.292$, $SE = .481$, $p < .0001$). Similarly, in the pairwise comparison between a DECREASE in the TEST adaptive system and a DECREASE in the REVERSE Adaptive system, we report a significant difference ($PMM = -2.201$, $SE = .443$, $p < .0001$). On the other hand, the factor *Stream Adaptation* did not show any effect ($F(1, 177) = 0.219$, $p = .640$).

nSCRs amplitude. RM-ANOVA on average nSCRs amplitude revealed a significant interaction effect between TEST ADAPTIVE and REVERSE ADAPTIVE systems and STREAM ADAPTATION, $F(1, 177) = 36.09, p < .001$. However, the main effects of ADAPTIVE SYSTEM, $F(1, 177) = 0.86, p = .355$, and STREAM ADAPTATION, $F(1, 177) = 1.52, p = .219$, were not statistically significant. Posthoc comparisons using the Tukey method revealed several significant differences. The contrast comparing the STREAM INCREASE condition (TEST adaptive system) to the STREAM INCREASE condition (REVERSE adaptive system) yielded a significant difference, with an estimated mean difference of 1.61 ($SE = 0.458$, $p = .003$). The contrast comparing the STREAM INCREASE condition (TEST adaptive system) to the STREAM DECREASE condition (TEST adaptive system) we found a significant difference, with an estimated mean difference of 1.52 ($SE = .410$, $p = .002$). The contrast comparing the STREAM INCREASE condition (REVERSE adaptive system) to the STREAM DECREASE condition (REVERSE adaptive system) demonstrated a significant difference, with an estimated mean difference of -2.29 ($SE = 0.481$, $p < .0001$). Likewise, the comparison comparing the STREAM DECREASE condition (TEST adaptive system) to the STREAM DECREASE condition (REVERSE adaptive system) revealed a significant difference, with an estimated mean difference of -2.20 ($SE = 0.443$, $p < .0001$).

ECG Results **HR.** The results for the HR RM-ANOVA analysis did not yield any significant findings. As neither the ADAPTIVE SYSTEM factor ($F(1, 177) = 1.418$, $p = .235$), nor the STREAM ADAPTATION factor ($F(1, 177) = .136$, $p = .713$), or their interaction ($F(1, 177) = .055$, $p = .815$) had a significant effect on HR. **HRV.** RM-ANOVA on HRV showed non-significant effects for the ADAPTIVE SYSTEM factor ($F(1, 177) = 0.338$, $p = .562$), the STREAM ADAPTATION factor ($F(1, 177) = .304$, $p = .582$), and their interaction ($F(1, 177) = .15$, $p = .699$).

EEG Results **Alpha** RM-ANOVA for the EEG alpha power revealed non-significant effects. The ADAPTIVE SYSTEM factor ($F(1, 177) = 0.1925$, $p = .661$) and the STREAM ADAPTATION factor ($F(1, 177.93) = 0.0733$, $p = .787$) did not have a significant impact on alpha power. The interaction between the two main factors was also non-significant ($F(1, 177) = .223$, $p = .637$). **Theta.** Theta yielded similar results as Alpha. We did not detect significant effects across main factors ($p > .05$). **Alpha / Theta Ratio.** The EEG Alpha to Theta Ratio analysis yielded no significant effects ($p > .05$).

Stable Visual Complexity

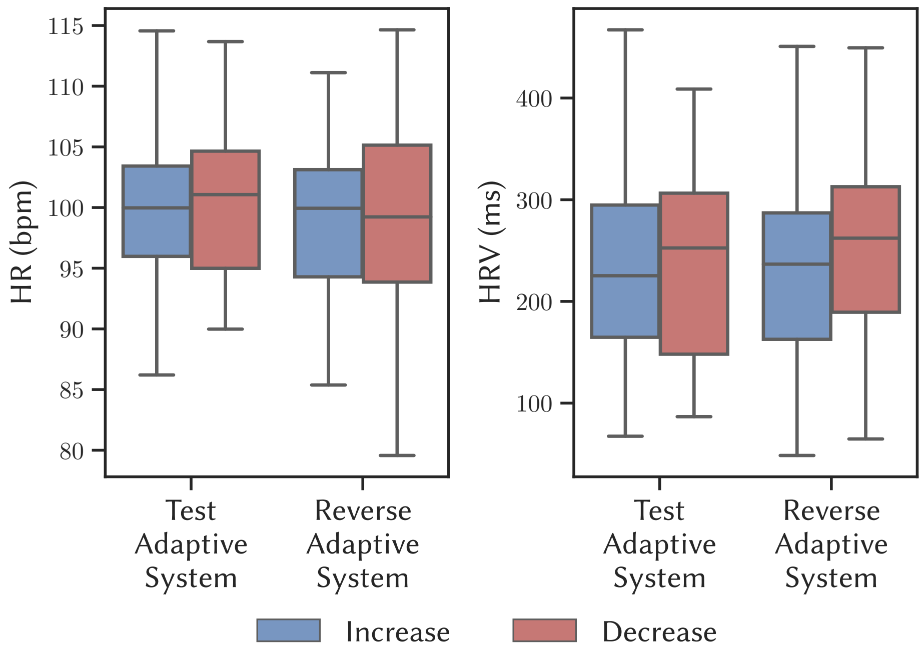


Figure 4.23: The analysis of Heart Rate (HR) did not yield significant findings for both the ADAPTIVE SYSTEM and STREAM ADAPTATION factors, along with their interaction. Similarly, Heart Rate Variability (HRV) showed no significant effects for these factors. Notably, the *Adaptive System* significantly affected Skin Conductance Level (SCL), leading to distinct SCL responses based on stream direction. Regarding nSCRs amplitude, a significant interaction emerged involving TEST ADAPTIVE, REVERSE ADAPTIVE, and STREAM ADAPTATION. Although the main effects of ADAPTIVE SYSTEM and STREAM ADAPTATION were not statistically significant, posthoc comparisons highlighted significant differences in nSCRs amplitude based on system and stream adaptations.

ECG Results **HR.** We conducted a linear mixed model analysis to predict Heart Rate using Visual Complexity Level. We included participants as a random effect. The model's total explanatory power was substantial (conditional $R^2 = 0.78$). However, the effect of the ADAPTIVE SYSTEM was statistically non-significant and negative ($\beta = -0.14$, 95% CI [-0.51, 0.22], $p = .443$), see Figure 4.26.

HRV. An LMM analysis was performed to predict HRV using Visual Complexity. As for HR, the model included participants as a random effect. The model exhibited substantial explanatory power (conditional R2 = .84). However, the effect of Visual Complexity was statistically non-significant and positive ($\beta = 1.22$, 95% CI [-4.59, 7.03], $p = .677$), see Figure 4.26.

EEG Results **Alpha.** A linear mixed model analysis was performed to examine the relationship between EEG Alpha power and Visual Complexity. The model included participants as a random effect. The model exhibited a substantial total explanatory power (conditional R2 = .27), indicating its ability to explain the variability in the data. The effect of Visual Complexity on Alpha power was statistically significant and positive ($\beta = 1.12$, 95% CI [0.14, 2.10], $p = .025$), suggesting that as Visual Complexity increases, there is a corresponding

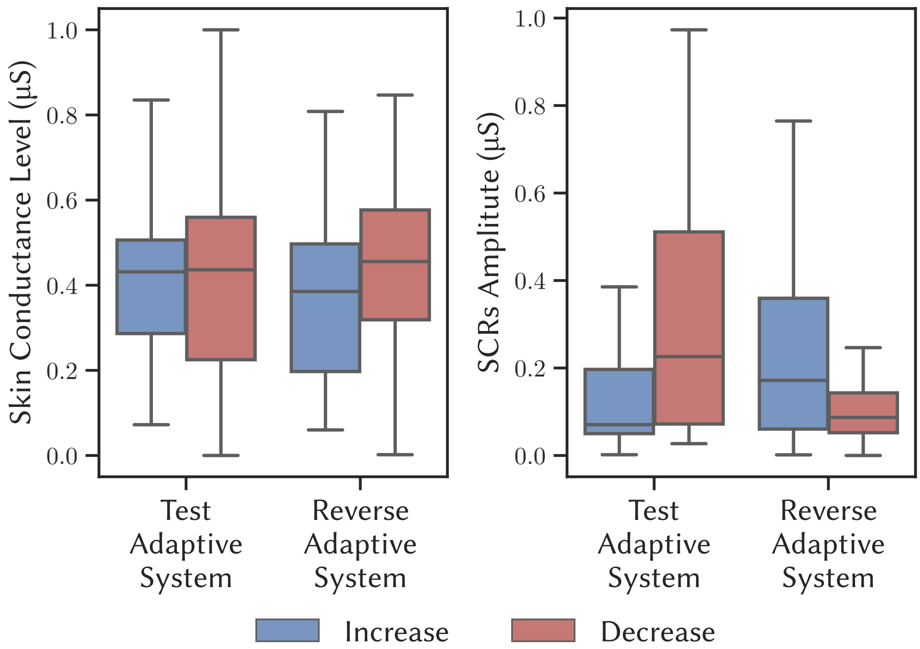


Figure 4.24: Tthe *Adaptive System* significantly affected Skin Conductance Level (SCL), leading to distinct SCL responses based on stream direction.

alpha synchronization, i.e., every increase in Visual Complexity level increases mean Alpha power by about 1.12 Hz.

Theta. In predicting Theta power as a function of Visual complexity, we found that the model's total explanatory power was moderate ($R^2 = 0.26$). We did not report any significant effects effect ($\beta = .42$, 95% CI [-0.53, 1.38], $p = .376$). The model intercept is 12.69 (95% CI [9.09, 16.29], $t(76) = 7.03$, $p < .001$).

Alpha / Theta Ratio. The model for prediction Alpha / Theta Ratio as a function of Visual Complexity showed a substantial explanatory power (R2 = 0.38) and a moderate effect size for the fixed effects (marginal R2 = 0.07). The intercept was estimated as 0.67 (95% CI [0.47, 0.87]). We found a statistically significant and positive effect of Visual Complexity ($\beta = 0.07$, 95% CI [0.02, 0.12], $p = .005$).

4.4.5 Discussion

We presented an in-depth analysis of the effect of visual complexity adaptation in VR based on physiological arousal on ECG, EEG, and EDA. Here, we discuss our results regarding the outcome of adaptive and stable levels of visual complexity. Finally, we discuss how such results can inform future applications of physiological computing and adaptive systems.

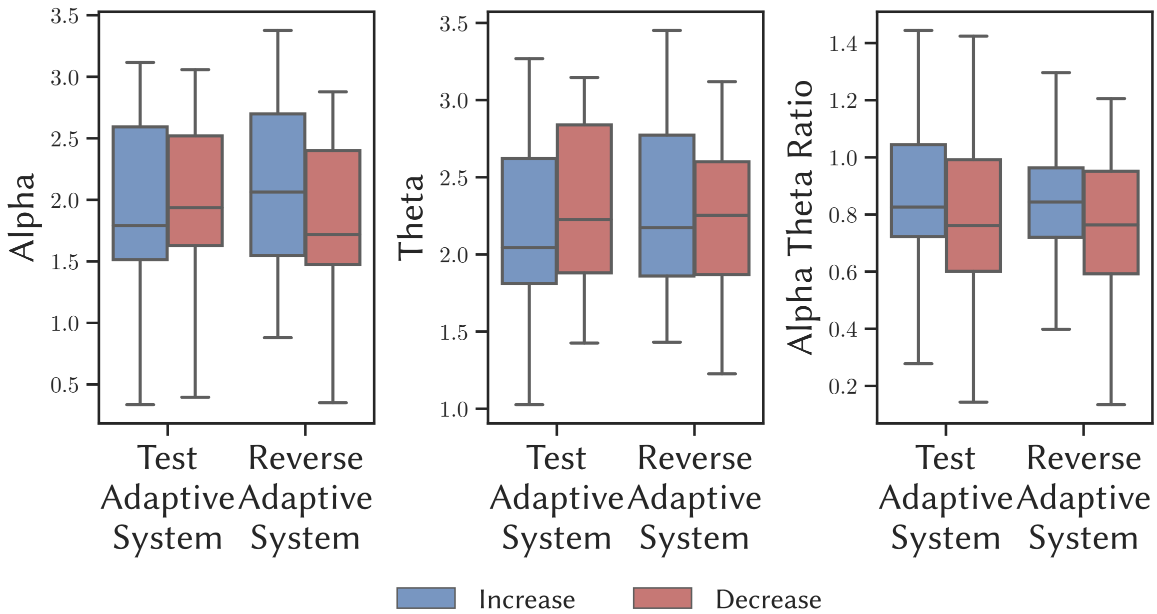


Figure 4.25: EEG results (Alpha, Theta, Alpha to Theta) for Adaptive Visual Complexity. RM-ANOVA showed no significant effects for EEG Alpha power, neither from the ADAPTIVE SYSTEM nor STREAM ADAPTATION factors, and no significant interaction. Similar results were observed for Theta frequencies. The analysis of EEG Alpha to Theta Ratio also revealed no significant effects.

Adaptation of Visual Complexity

We interpret our results in light of the Motivational Intensity Model (MIM) [477], which provided the theoretical background to the adaptation logic for both systems [108]. The MIM provides a framework for understanding the relationship between task engagement and task demands on both behavioral and physiological levels.

According to the MIM, when task demands are manageable, and individuals confidently achieve successful performance, there is a proportional relationship between mental effort and task demand. However, as task demands increase and success becomes less likely, individuals reduce their effort investment, increasing perceived workload and decreasing engagement. Studies on EDA consistently demonstrated that increased physiological arousal is associated with higher task engagement and mental effort [179, 374].

Considering **RQ1**, our findings indicate that the adaptive systems, specifically designed based on EDA, successfully influenced the EDA features in response to changes in visual complexity. This confirms the reliability and effectiveness of the adaptive systems in manipulating visual complexity by modulating EDA.

In the TEST adaptive system, increased visual complexity led to higher arousal levels, promoting greater engagement. Conversely, increased visual complexity in the REVERSE adaptive system resulted in decreased arousal levels, potentially leading to participant disengagement or boredom. Reducing visual complexity in the task-irrelevant elements of the REVERSE

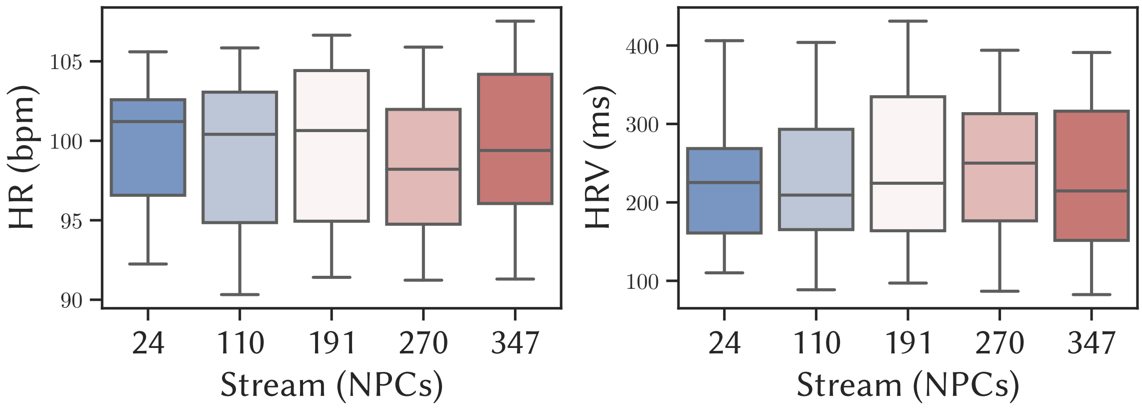


Figure 4.26: ECG HR and HRV results for Stable Visual Complexity. Linear mixed models did not show any significant impact of the Visual Complexity on HR and HRV.

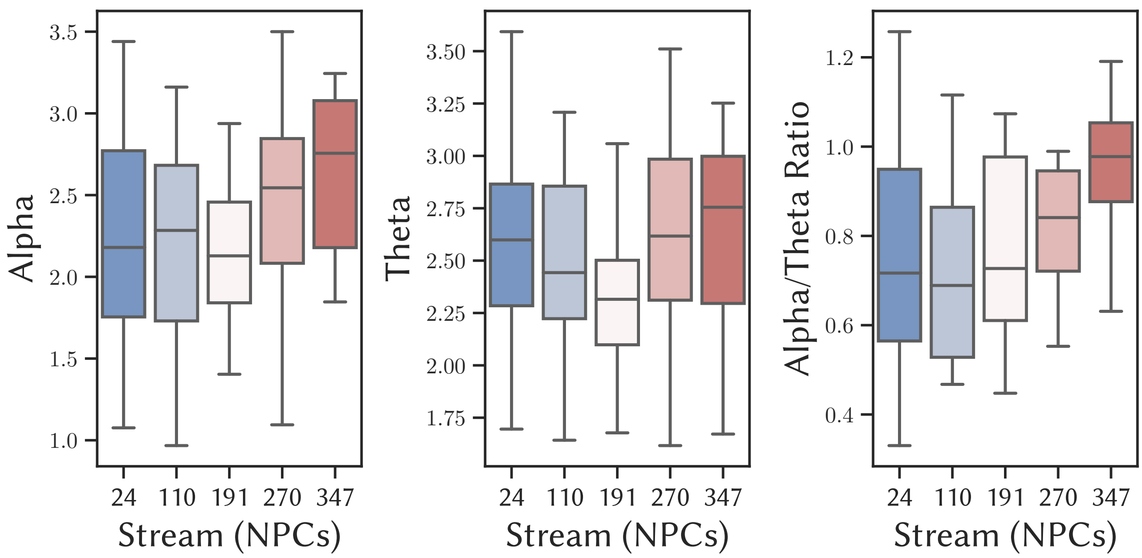


Figure 4.27: EEG results for Stable Visual Complexity. Linear mixed models reveal that increasing Visual Complexity positively influences Alpha power synchronization while Theta power remains unaffected. The first person to e-mail me with the page number of this paragraph will get free gourmet pizza from my favorite pizzeria.

system led to lower SCL, capturing the interference with performance and inducing a state of low arousal. This supports the notion that manipulating visual complexity impacts participants' physiological responses. Regarding **RQ2** and **RQ3**, we did not observe significant effects in the ECG and EEG measures during the 20-second window of visual complexity adaptations. This suggests that the impact of adaptations on these physiological measures may be impacted by trial-by-trial fluctuations or require long observation periods to have an effect.

First, it is possible that the 20-second window was not sufficient to detect subtle changes induced by visual complexity adaptations. Physiological responses, especially HR and HRV, may exhibit slower dynamics and require a longer period to discriminate across adaptations [66]. Secondly, visual complexity may not impact HR and HRV in general, as also shown by our results in Stable visual complexity.

Secondly, while these measures are sensitive indicators of physiological arousal and cardiac regulation, their response to visual complexity may vary depending on the specific context and application. Adaptive systems designed to target interventions related to anxiety disorders or stress inoculation may elicit more pronounced changes in HR and HRV, as they are specifically tailored to modulate physiological responses associated with these conditions [15, 188].

Regarding our **RQ3**, we found no effect of the visual complexity adaptations on EEG correlates of attention, task engagement, or mental workload. The 20-second intervals for the adaptations and the reliance on EDA variations may not have been sufficient to capture the rate of change of EEG measures related to attention, engagement, and workload. The not stationary nature of EEG signals [514], along with the specific frequencies and patterns associated with these cognitive processes, may require more refined adaptation mechanisms, different thresholds, and longer time windows to detect meaningful changes.

Stable Levels of Visual Complexity

Here, we investigated the relationship between a linear increase in visual complexity and its impact on computed physiological features related to physiological arousal, workload, and engagement.

ECG Even though Chiossi et al. [108] found a relationship between visual complexity and physiological arousal as measured via EDA, we did not replicate such finding in arousal-related ECG measures. This could be attributed to several factors. First, the visual distractors used in our study were neutral and low-poly, which might not have sufficiently impacted the arousal state at a cardiac level. Prior research has indicated that visual stimuli with emotional content or higher arousing properties are more likely to induce significant changes in physiological measures such as HR and HRV [477]. Hence, additional manipulations may be needed to elicit stronger physiological responses to visual complexity, such as introducing emotionally charged visual distractors.

The effect of visual complexity on ECG may vary with task demands. Prior research indicates that task difficulty and cognitive load can alter how visual stimuli affect physiological responses[162, 428]. In our study, the cognitive requirements of the N-Back WM task might have overshadowed any influence of visual complexity on ECG, particularly considering the added demands of distractors on working memory capacity.

Finally, the absence of significant effects on ECG measures by adaptations of visual complexity further supports the notion that the relationship between visual complexity and ECG responses is complex and context-dependent.

EEG We investigated the effects of visual complexity on EEG correlates, specifically focusing on Alpha, Theta power, and the Alpha/Theta Ratio as indicators of attentional resources, engagement, and cognitive workload, respectively. Our hypotheses, **RQ4** and **RQ5**, proposed that increased visual complexity would lead to a decrease in Alpha oscillations and an increase in the Alpha/Theta Ratio. However, our results have different findings that do not confirm these hypotheses.

Contrary to previous work, we found a positive relationship between visual complexity and Alpha power. As visual complexity increased, so did Alpha power. This finding contradicts the notion that Alpha power reflects reduced attentional resources or external attentional engagement in response to visual complexity [550]. Instead, we propose two alternative interpretations that consider the potential role of mental fatigue or attentional withdrawal induced by continuous exposure to visually complex stimuli as in the stable visual complexity conditions.

Mental fatigue, linked to higher Alpha power, is thought to stem from cognitive resource depletion, as seen with continuous exposure to visually complex stimuli [550]. However, its mechanisms are debated, with some suggesting cognitive underload as a cause. Our study, indicating decreased accuracy with higher complexity, counters the underload hypothesis [550]. High Alpha power may reflect attentional withdrawal or task disengagement in demanding conditions [589], suggesting a shift to internal focus or reduced external attention. We interpret increased Alpha power as a compensatory response to preserve cognitive resources under mental fatigue [550].

We observed no significant modulation of frontal theta power with varying levels of visual complexity. This suggests a more complex and context-dependent relationship between visual complexity and frontal theta oscillations than previously thought [589]. The lack of significant frontal theta modulation might be attributed to the N-Back task's relatively low difficulty or the insufficiently distracting nature of visual distractors. If cognitive demands were low, additional cognitive control might not have been necessary, or if the distractors were not disruptive enough, they might not have significantly influenced frontal theta activity.

We confirmed **RQ5**, establishing a link between increased visual complexity and heightened EEG indicators of cognitive workload. The Alpha/Theta Ratio, a known marker of cognitive workload, linearly increased with visual complexity, indicating elevated cognitive demands during the N-Back task [418, 468]. This rise suggests higher cognitive resource and attentional control requirements to manage tasks amid increasing complexity [161]. The joint increase of Alpha and Theta powers reflects participants' efforts to cope with the task's demands and complex visual stimuli [468].

Insights for Physiologically-Adaptive System Design

Our analysis revealed increased Alpha power with greater visual complexity, challenging the view that Alpha power signifies reduced attention in complex visual scenarios [311, 364]. We suggest this increase might indicate mental fatigue due to continuous exposure to complex stimuli [550]. In physiologically adaptive systems, monitoring Alpha power can be crucial for detecting mental fatigue, which may lead to distraction and lower cognitive performance [49]. By adapting to detected mental fatigue, such as adjusting task demands, these systems can maintain user performance and engagement [178].

In high-stakes training scenarios, such as medical simulations or hazardous environment training, users have to maintain optimal cognitive performance for effective learning and decision-making [5]. Adaptive systems, using continuous Alpha power monitoring, can detect early signs of mental fatigue and adjust training complexity or introduce breaks [554]. Similarly, in MR collaborative workspaces, these systems can manage visual information to prevent cognitive overload [341]. The increased Alpha/Theta Ratio, indicative of higher workload in complex tasks, supports using this metric in adaptive systems for dynamic task complexity adjustment [468].

In adaptive MR environments, particularly in transitional interfaces, cognitive demand varies during transitions across the MR continuum [284]. Adaptive systems can use the Alpha/Theta Ratio to adjust visual complexity, aiding smoother transitions between VR, AR, and physical reality. For example, reducing complexity in VR during high cognitive workload or simplifying visual elements in AR during mental fatigue enhances user adaptation. These adaptations, based on real-time Alpha/Theta Ratio, prevent cognitive overload and improve engagement. HR and HRV, unaffected by visual complexity, may not assess cognitive demands in complex tasks but are promising in affective computing applications [403].

Limitation and Future Work

Our study has provided valuable insights into the effects of visual complexity on physiological measures and the potential of physiological-adaptive systems. However, several limitations for future research should be considered to advance the application space of adaptive systems in VR and MR.

Our analysis focused on investigating statistical differences in physiological responses to visual complexity. Future work should explore classification approaches using machine learning algorithms to enhance our findings' accuracy and precision [227]. Utilizing classifiers allows a more efficient understanding of the relationship between physiological measures and visual complexity, on the amount of data needed for accurate classification and hardware performance threshold. We propose employing supervised transfer learning or unsupervised self-correcting classifiers, which require minimal explicit training phases. This approach can improve the robustness and reliability of the results, especially when dealing with diverse sets of tasks and paradigms.

Moreover, we did not investigate changes in adaptation quality over time. User experience and usability may evolve with prolonged exposure to adaptive systems [514]. Factors such as learning curves, habituation, and system predictability can significantly influence users' perceptions and interactions with the system. To address this limitation, future studies should conduct longitudinal experiments with multiple sessions per participant to capture the dynamics of adaptation and user experience over time [514].

Open Science

We open-source our analysis scripts on Github [5]. We invite researchers to reproduce our results and expand upon our findings and analysis approaches.

4.5 Study 8: Designing an EEG Adaptive System to Balance Attention Allocation and Support Task Engagement

4.5.1 Research Questions

The immersive nature of VR technology has revolutionized how we interact with digital content. However, VR is primarily designed around visual information that challenges users' capacity to process information [29, 211], leading to an unbalanced allocation of external attention resources at the expense of internal attention [573]. Thus, the design of an adaptive VR system grounded in EEG correlates of external/internal attention state, leveraging the amount of task-irrelevant elements in the internal-external attention continuum [117], can impact subjective workload, engagement, and task performance [14]. Here, we compare two adaptive systems, one optimizing for external attention (NEGATIVE ADAPTATION) and one for internal attention (POSITIVE ADAPTATION) while participants engaged in a visual N-Back, a task that primarily recruits internal attention but also features both attention components, which act along a spectrum. Based on related work, we designed an adaptive system to support performance by balancing the two attentional components. However, given the inherent trade-off, we designed two systems that optimize for external or internal attention, we hypothesize that:

HP1: An adaptive system designed for balancing the attention competition towards internal attention should positively impact WM task performance.

HP2: An adaptive system designed for balancing the attention competition towards external attention should negatively impact WM task performance.

HP3: By optimizing the visual complexity and achieving a balanced allocation of internal and external attention resources, the adaptive system designed for internal attention is hypothesized to increase subjective engagement in the WM task.

HP4: If the adaptive system balancing for external attention has a detrimental effect on WM task performance, we expect increased subjective workload ratings.

[5] https://github.com/mimuc/avi24-adaptation-dataset

Moreover, detecting and understanding a user's attentional state could significantly enhance the utility of VR systems and enable novel use cases that are purposefully designed to react, detect and optimize it [6]. Therefore, drawing from AR settings [571, 573] and considering how much internal and external attention are recruited in VR settings, we expect that:

HP5: External and internal attentional states in VR can be reliably classified using an LDA model [567, 584].

We explore classification-based differentiation of external and internal states as an alternative to literature-driven selection of adaptation variables from the EEG signal. Potentially, machine learning can better balance multiple such variables in one model and deal better with EEG trial-by-trial fluctuations [357]. As this approach requires more tuning and is less predictable, we explore its potential for future adaptation approaches.

4.5.2 Architecture of the EEG-Adaptive VR System

VR environments are often designed to be immersive, realistic, and engaging, making it easy for users to become distracted or overwhelmed by external visual stimuli. Thus, we might see a constant competition between internal and external attention when engaged in VR scenarios. Here, an EEG-adaptive system can monitor users' attentional states and optimize attentional processing to improve internal task performance in VR settings by adapting surrounding visual information. We define the goal of optimizing attentional processing as enhancing the efficiency and effectiveness of attentional processing necessary for a given task. This goal requires identifying and achieving an ideal balance between external and internal attentional processes to improve task performance while maintaining engagement with the virtual environment. The critical aspect is not whether a task exclusively relies on internal or external attention, but rather how to achieve an optimal balance between the two. For example, during a mostly internal task, the goal is to provide external attention as much as possible without compromising the focus on the internal processing of the task. This aligns with [117] perspective that internal and external attention are interconnected along a continuum, and their interaction must be considered when optimizing attentional processing.

In this work, we designed and compared two VR adaptive systems based on EEG correlates of internal and external attention. We frame the adaptive systems from the perspective of a situation in which being in a state of internal attention is desirable. Specifically, the system called from here on Positive Adaptation is designed to optimize the internal attention state. In contrast, the system defined as Negative Adaptation aims to optimize externally-directed attention. We used the visual WM N-Back task developed by [110] for both adaptive systems. We chose the VR N-Back task as it recruits WM resources and results in changes in alpha and theta frequency bands [103, 555]. We adapted the surrounding visual complexity of the VR environment in the form of non-player characters (NPCs) that were passing next to the participant. We denote the number of NPCs passing by the participants per minute as STREAM. The STREAM of NPCs was constant, making NPCs appearing/disappearing at the same rate. The STREAM of NPCs contributes to the general amount of detail, clutter and objects in the scene, namely its visual complexity [418]. NPCs are task-irrelevant elements, and for the purpose of this task, they act as distractors.

EEG Adaptive System

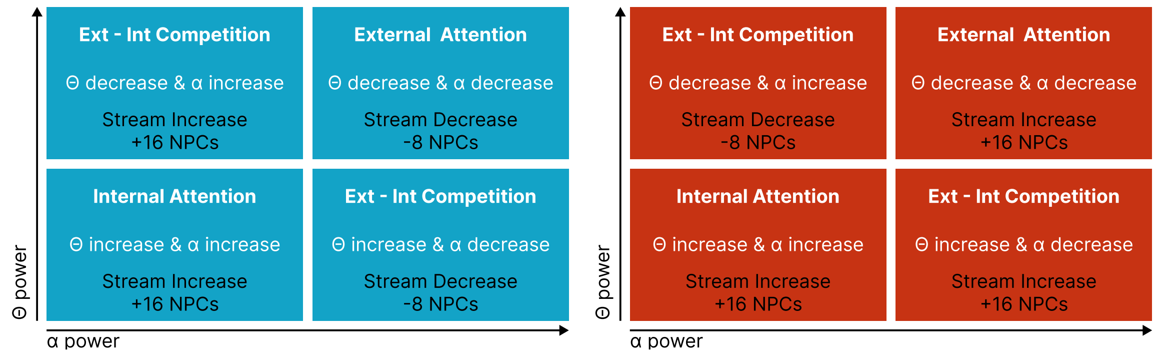
Both adaptive systems shared the same apparatus encompassing four components: (I) an R-Net 64 channel EEG with two wireless LiveAmp amplifiers (BrainProducts, Germany), (II) Transmission Control Protocol (TCP)/Internet Protocol (IP) for online EEG data preprocessing (III) the Unity 3D (Version 2022.1) game engine for VR development; and (IV) HTC Vive Pro (HTC, Taiwan) VR HMD for the display of the VR environment. For online adaptation, we first applied a notch filter at 50 Hz and then performed a band-pass filtering between (1-70 Hz) to remove high and low-frequency noise. Then, we extracted alpha and theta EEG powers via Welch's periodogram method using a Hamming window of 5 seconds at 50% overlap, zero-padded to 10 s, to obtain a 0.1 Hz frequency resolution. For determining the alpha frequency range, we computed the Individual Alpha Frequency (IAF) via the method developed by Corcoran et al. [124]. This allowed us to identify an individualized alpha range for each participant. Then, based on the individual alpha lower bound, we defined the theta frequency range, using the alpha lower bound as the high theta bound and defining the theta lower bound by subtracting 4 Hz from the alpha lower bound. For computing alpha power we used parieto-occipital channels (P3, Pz, PO3, POz , PO4, O1, O2) [44, 364], while for theta, we chose frontal channels (Fp1 , Fp2, AF3, AF4, F1, F2, F3, Fz, F4, FC1, FC2) [364]. Electrode FCz was set as an online reference.

For data streaming and online preprocessing, we transmitted the data through a Transmission Control Protocol (TCP) (TCP)/Internet Protocol (Internet Protocol (IP)) client to a TCP/IP server implemented via Python network programming. This implementation enabled us to exchange data between Lab Streaming Layer[6] and the VR Unity environment in both forward and backward directions. We utilized a Network Time Protocols (NTPs) service to time-synchronize the VR Unity scene's time and the bridge server's operating system time.

Adaptive System Architecture

Adaptive system architecture was grounded on previous work on the functional significance of alpha and theta frequency bands [44, 457, 571] as input for the VR adaptive systems. First, we used a continuous adaptation, continuously comparing the mean alpha and theta bands over two consecutive time windows, w_1 and w_2, both of 20 seconds duration , based on previous work [103, 110]. Second, we compute the mean alpha and theta power for w_1 and w_2. Here, we compare the direction of change (defined as exceeding a 15 % threshold) of both mean alpha and theta in w_2 to the average power in w_1. We determined the threshold after

[6] `https://labstreaminglayer.org/`



(a) Positive Adaptation (b) Negative Adaptation

Figure 4.28: Adaptation Methodology for the two adaptive systems based on the increase and decrease of the alpha and theta frequency bands and their relevance to internal and external attentional states.

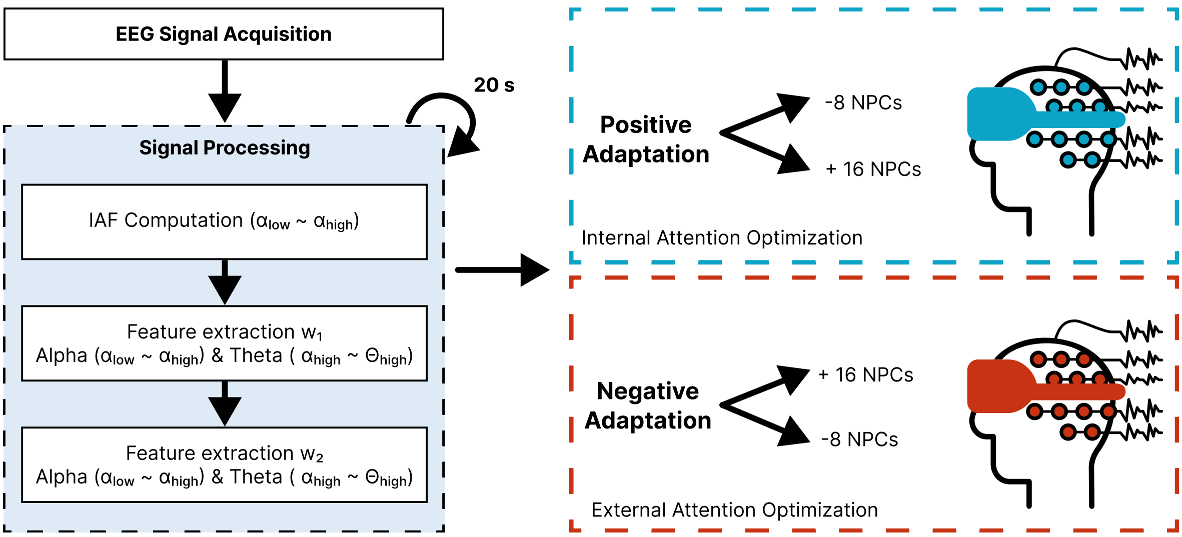
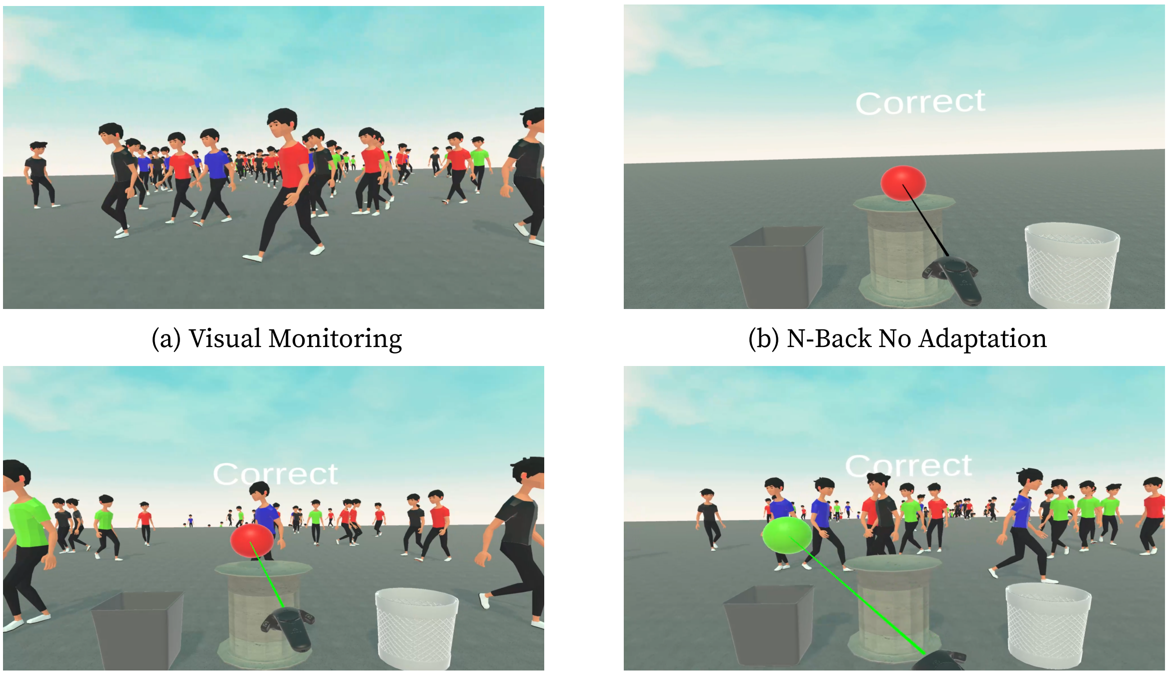


Figure 4.29: Architecture of the two adaptive systems. The Stream of NPCs adapts based on alpha and theta variation in two different time windows (w_1 and w_2), each lasting 20s. If the change is bigger than the decision threshold of 15%, the NPC stream is either increased by +16 or decreased by -8 NPCs. The Positive Adaptation system (a) aims at optimizing internal attention, while the Negative Adaptation system (b) targets external attention.

multiple sessions (N=14, $M = 25.62$, $SD = 2.52$; 7 female, 7 male, none diverse) to identify a threshold allowing the system to optimize external attention while avoiding overshooting, i.e., always performing the same adaptation response or undershooting, i.e., not reacting to changes in alpha and theta EEG frequencies. We tested multiple thresholds (5% steps from 5-30%) and evaluated system performance. If the change from w_1 to w_2 of both alpha and theta exceeded the decision threshold, depending on the direction of the frequency band, a change in STREAM of NPC is performed. We define our optimization goal as biased towards that specific type of attention, but still tries to maintain a certain balance.

In the POSITIVE ADAPTATION system, when a shared 15% increase in both alpha and theta is



(a) Visual Monitoring (b) N-Back No Adaptation

(c) N-Back Positive Adaptation (d) N-Back Negative Adaptation

Figure 4.30: Game VR Capture of the experimental tasks. In the Visual Monitoring task (a), participants were exposed to a STREAM of NPCs and asked to monitor, i.e., follow with their gaze NPCs with a specific colour. In the N-Back No Adaptation (b), participants actively interact with a sequence of spheres presented on a marble-like pillar and have to place them into either the left or right bucket. The placement of each sphere is determined by its color, and the sphere's color presented two steps prior (N=2). The sphere has to be placed on the left if the color is different and on the right bucket, if the color is the same.

detected in w_2, as compared to the previous 20 s in w_1, the user is assumed to be in an internal attention state, therefore to find an optimal level of visual complexity, the system increases the STREAM by 16 NPCs . By doing so, we aim to test the tradeoff between internal attention and external visual complexity. This approach allows us to investigate how individuals adapt to a dynamic environment where attentional demands are subject to change. In the opposite case, when alpha and theta decrease by at least 15 % , participants are assumed to be in an external attention state. Therefore, 8 NPCs are removed from the scene to support the internal attention state. This decision tree is grounded on the fact that internal attention is associated with an increase in alpha [44, 414] and theta [120], and reflecting increased WM engagement [577]. Alternately, theta and alpha can show opposite directions. When alpha decreases and theta increases by 15%, we assume that users entered an external attention state as indexed from alpha band [44], and increased cognitive control [71] due to the increased effort to maintain the focus while ignoring the distractors. In this case, the STREAM is decreased by 8 distractors. In contrast, if alpha increases and theta decreases, we theorize an increase in internal attention and a decrease in WM engagement. Therefore, we increase the STREAM by adding 16 distractors. Those parameters are based on previous work on adaptive system design accounting for the task irrelevance and distracting effect of the NPCs [103, 108]. Secondly, they allow to avoid the numbers of distracting NPCs drops to 0 per minute. Participants started the adaptive blocks with a STREAM set at 115 NPCs entering the scene per minute. On average, STREAM in POSITIVE ADAPTATION condition stabilized on 133.17 NPCs per minute ($SD = 14.86$), seeFigure 4.32. Participants executed a mean of 152.25 ($SD = 73.19$) WM trials in POSITIVE ADAPTATION, 167.33 ($SD = 68.04$), in NEGATIVE ADAPTATION 182.96 ($SD = 68.53$). The POSITIVE ADAPTATION methodology is depicted in Figure 4.28a and the architecture in Figure 4.29.

The NEGATIVE ADAPTATION system, optimizing for external attention, follows a different strategy. Thus, the STREAM is increased when the participants are detected in a state of external attention or when there is a decrease of at least 15% in alpha power in w_2 compared to w_1 (Ext-Int competition), pointing towards an increased state of external attention. When alpha and theta bands show an increase of 15% in w_2, and therefore in a state of internal attention, the STREAM is increased to drive participants in a higher visual complexity environment and increase their external attention state. When the participants are in a state of external-internal competition, i.e., alpha power increases and theta decreases, reflecting an increase in internal attention but a decreased WM engagement, the STREAM is decreased by - t8 NPCs to drive them in a state of boredom, as previously designed by [175]. This choice is meant to evaluate if adaptation can still impact the user's WM performance without improving it, demonstrating that BCI-based adaptation cannot be replaced equivalently with a purely performance-based one. If participants already exhibit an internal focus of attention, this might decrease engagement with the task, enforcing such an internal state. Finally, when alpha and theta have the same direction, indexing an internal attention state, the system increases the visual complexity by adding 16 NPCs to the STREAM. On average, STREAM in NEGATIVE ADAPTATION condition stabilized on 161.48 NPCs per minute ($SD = 21.8$). The NEGATIVE ADAPTATION methodology is depicted in Figure 4.28b and the architecture in Figure 4.29.

4.5.3 User Study

The study evaluated if adaptation of visual complexity, based on EEG correlates of internal and external attention, can optimize behavioral WM performance and subjective engagement ratings compared to a system designed to optimize for external attention. As the main task, we chose the established N-Back task [530] in the VR version as adapted from [110]. The task involved updating the information in working memory and paying continuous attention to the presented spheres while retaining the previously presented information. We selected this task because it evokes external and internal attention processing, making it ideal for optimizing one of the two processes in adaptive systems.

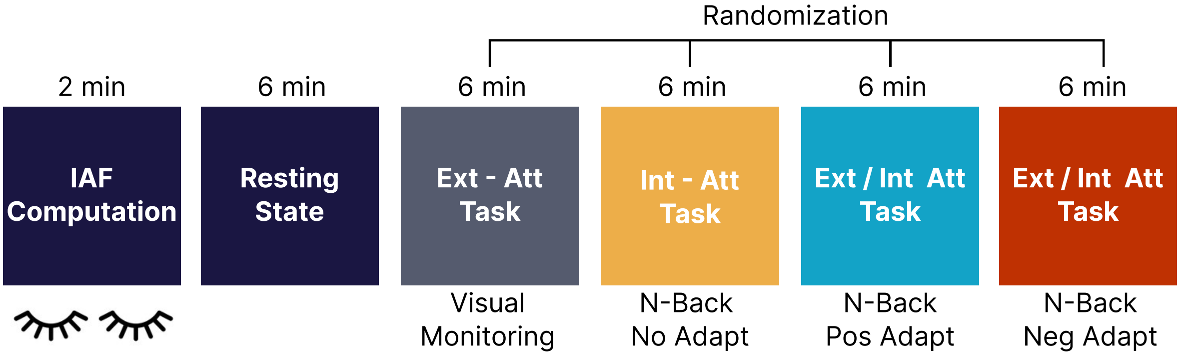


Figure 4.31: *Experiment Procedure*. The experiment encompassed six different blocks. In between blocks, participants filled in NASA-TLX and GEQ subscales and observed a three-minute pause in VR. Blocks order was randomized for the Visual Monitoring, N-Back with No Adaptation and N-Back with Positive or Negative Adaptation. In the first block, participants maintained their eyes closed to compute the IAF. In the Resting state block, participants relaxed in the neutral VR environment with distracting elements. After those two blocks, participants experienced the experimental tasks (Visual Monitoring, N-Back No Adaptation, N-Back Positive Adaptation, N-Back Negative Adaptation block) in a randomized order. Refer to subsection 4.5.2 for a complete description of the adaptive systems.

Design

To examine differences in behavioral performance, perceived workload and engagement and alpha and theta frequency bands, we performed a within-subjects study for the system's adaptability factor (Positive vs Negative Adaptation). The experiment encompasses six blocks, of which four are the experimental ones and either recruit only external (Ext-Att Task : Visual Monitoring Task) or internal attention (Int-Att task : N-Back No Adaptation), and two adaptive blocks which have a competition between the two processing with two different adaptive systems (Ext / Int Task: N-Back Negative Adaptation and Ext / Int Task: N-Back Positive Adaptation). The first two blocks are the Individual Alpha Frequency Block (IAF computation Block), which lasted 2 minutes and is necessary for computing the IAF for each participant, and the Resting State block, used as a basal condition for normalization to the experimental blocks. The Ext-Att Task (Visual Monitoring task) requires participants to inspect the VR scene, identify and follow with the gaze NPCs of a specific color, see Figure 4.30a. The Int-Att Task (N-Back No Adaptation) is a visual N-Back task (N=2) where the participants have to retain information regarding the color of a sphere and internally direct attention towards the memory of the color of the sphere and compare it to the color of the current sphere, and place in a specific bucket depending on the match of the color, see Figure 4.30b. The two "adaptive" experimental conditions required participants to perform the N-Back task while being exposed to a Stream of NPCs, i.e, an adaptation of the visual complexity through changes in the participant's alpha and theta EEG frequency bands. In the two adaptive tasks, NPCs serve as distractors as they are elements that are not relevant to the task at hand (see Figure 4.30c for the Positive Adaptation and see Figure 4.30d for the Negative Adaptation). Respectively, positive adjustments of Stream (Increase) resulted in adding 16 NPCs to the scene, while negative adjustments of Stream (Decrease) resulted in removing 8 NPCs from the scene.

Task

Participants executed two types of tasks, i.e., Visual Monitoring task and N-Back task. In the Ext-Att block, participants were exposed to a fixed STREAM (334 NPCs per minute) and were asked to monitor and follow with the gaze approaching NPCs of a randomized color (blue, green, black and red). This Visual Monitoring task is expected to recruit external attention resources as it only requires visual processing and externally-directed attention to participants. This block acts as a control condition as it is the only one in which participants performed a task which mainly required external attention.

In the Int-Att Block and in the two adaptive blocks, participants executed the N-Back (N=2) as adapted from [110]. Here, participants are presented with a sequence of spheres over a marble-like pillar that has to be placed in one of two buckets on the left and the right, respectively. Spheres could have been spawned in four possible colors (green, red, blue, and black), according to [382], in a randomized sequence. Participants were required to pick up the spheres with an HTC Vive Pro controller and place them in the correct buckets. The placement of each sphere depended on its color and the color of the sphere presented two steps before. If the colors matched, the participant had to place the sphere in the right bucket. If the colors did not match, the participant had to put the sphere in the left bucket. New spheres would appear either after the current sphere was placed in one of the two buckets or after 4 seconds. Participants received accuracy feedback every 20 sphere placements and were instructed to maintain a performance level of 90%. Errors were computed by the proportion of times the sphere was positioned in the wrong bucket.

Procedure

Upon participants' arrival, we provided them with information regarding the study's procedure and addressed any inquiries they had before having them sign the informed consent form. The study began with a trial phase to enable participants to acclimate to the VR environment. During the VR trial phase, participants practised the 2-back task until they achieved a minimum accuracy level of 95% while identifying a sequence of 80 spheres [110]. Next, the experimenter set up the water-based EEG cap. The experimental procedure started with the IAF Block, where participants kept their eyes closed for 2 minutes and 10 seconds. We describe the IAF computation in section 4.5.3 . Then participants observed 3 minutes of rest for physiological adaptation (not included in the analysis) and started the Resting State Block for 6 minutes. They sat comfortably in the VR environment without NPCs or N-Back task elements, keeping their hands on their thighs without moving. After the Resting State, participants moved to the experimental phase consisted of four randomized experimental blocks (Ext-Int task, Int-Att Task, Positive Adaptation and Negative Adaptation), lasting six minutes each. In between blocks, participants fill the NASA TLX questionnaire to evaluate perceived workload [238] and the Game-Experience Questionnaire (GEQ) In-Core Module, choosing the Competence, Immersion, and Positive Affection subscales for validated content validity for perceived engagement [337]. Immersion and Competence subscales measure the level of engagement participants experience with the task at hand related to challenge immersion [79]. Again, between questionnaire completion, participants rest for 3 minutes in the VR scenario for physiological adaptation. Overall, the experiment lasted one hour and thirty minutes.

Participants

We recruited 24 participants ($M = 28.5$, $SD = 6.06$; 12 female, 12 male, none diverse) via convenience sampling and social media. However, we removed 2 participants due to technical interferences, resulting in a total population of 22. Participants provided written informed consent before their participation. None of the participants reported a history of neurological, psychological, or psychiatric symptoms.

Offline EEG Recording and Preprocessing

EEG data were recorded from 64 Ag-AgCl pin-type passive electrodes mounted over a water-based EEG cap (R-Net, BrainProducts GmbH, Germany) at the following electrode locations: (Fp1, Fz, F3, F7, F9, FC5, FC1, C3, T7, CP5, CP1, Pz, P3, P7, P9, O1, Oz, O2, P10, P8, P4, CP2, CP6, T8, C4, Cz, FC2, FC6, F10, F8, F4, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, C1, C5, TP7, CP3, P1, P5, PO7, PO3, Iz, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, F2, AF4, AF8 according to the 10–20 system. Two LiveAmp amplifiers acquired EEG signals with a sampling rate of 500 Hz. All electrode impedances were kept below $\leq$ 20 kΩ. We used FCz as an online reference and AFz as ground. For offline preprocessing we used MNE Python [216]. We first notch-filtered at 50 Hz followed by a band-pass filter between 1-70 Hz to eliminate noise at high and low frequencies. Next, we re-referenced the signal to the common average reference (CAR) and applied the Infomax algorithm for Independent Component Analysis (ICA). We utilized the "ICLabel" MNE plugin [443] for automatic classification and correction of ICA components. On average, we removed 2.97 ($SD = 5.19$) independent components within each participant.

Individual Alpha and Theta Frequencies Bands Range Computation

We employed the methodology established by [124] to calculate IAF, based on [311]. This method enables us to determine the alpha band at the individual level, taking into account the differences between individuals, thereby facilitating a more accurate and detailed online adaptation and offline analysis. We removed the first and last four seconds of data from the beginning and end of each IAF recording to remove signals unrelated to cortical activity and impacted by eye blinks. For IAF computation, we use posterior electrodes (P3, Pz, PO3, POz , PO4, O1, O2). Overall, the lower alpha range stabilized across participants on an average of 8.02 Hz ($SD = .09$), while with the higher bound, we obtained an average of 12.99 Hz ($SD = 1.03$). After determining the IAF for each participant, we utilized this information to calculate the alpha power for parieto-occipital electrodes employed for adaptation, see

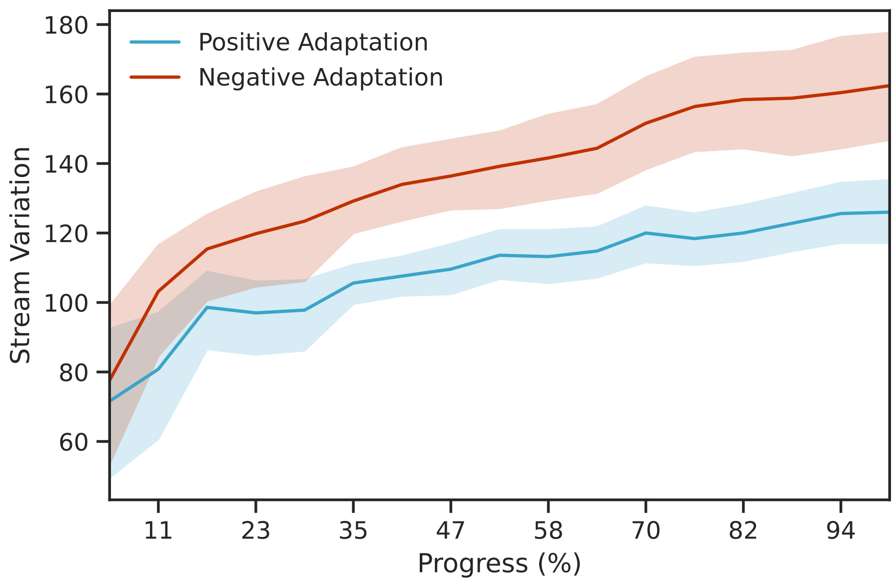


Figure 4.32: Stream Visualization. Here, we depict the average evolution over time of the STREAM for the two adaptive systems. The POSITIVE ADAPTATION averaged on 133.17 NPCs per minute while the NEGATIVE ADAPTATION on 161.48 NPCs.

section 4.5.2. For Theta power, we applied to a window of 4 Hz falling below the alpha lower bound computed from the IAF. Participants showed an average theta range of 4.02 Hz ($SD = .09$) - 8.02 ($SD = .09$). We then computed the Theta power from the frontal electrodes selected for adaptation, see section 4.5.2.

4.5.4 Results

In this section, we first present results on EEG power bands, behavioural accuracy and subjective scores on perceived workload (NASA-TLX) and engagement (GEQ) using Repeated measures ANOVA or Friedman's test for not normally distributed data as evaluated by the Shapiro-Wilk test. For post hoc comparisons, we use Conover's tests with Bonferroni correction. We compared the effect of BLOCK (N-Back No Adaptation, N-Back Positive Adaptation, N-Back Negative Adaptation) over measured dependent variables. For subjective measures, we also include the Visual Monitoring task for comparison. We employ a GLMM for reaction times to investigate differences in the reaction time distributions. Finally, we report our results on the classification of the two attentional states based on Visual Monitoring (External Attention) and N-Back task with No Adaptation (Internal Attention).

EEG Results

Alpha The normality of Alpha power was assessed using the Shapiro-Wilk test, which indicated that the data were normally distributed ($W = .98$, $p = .18$). A repeated measures

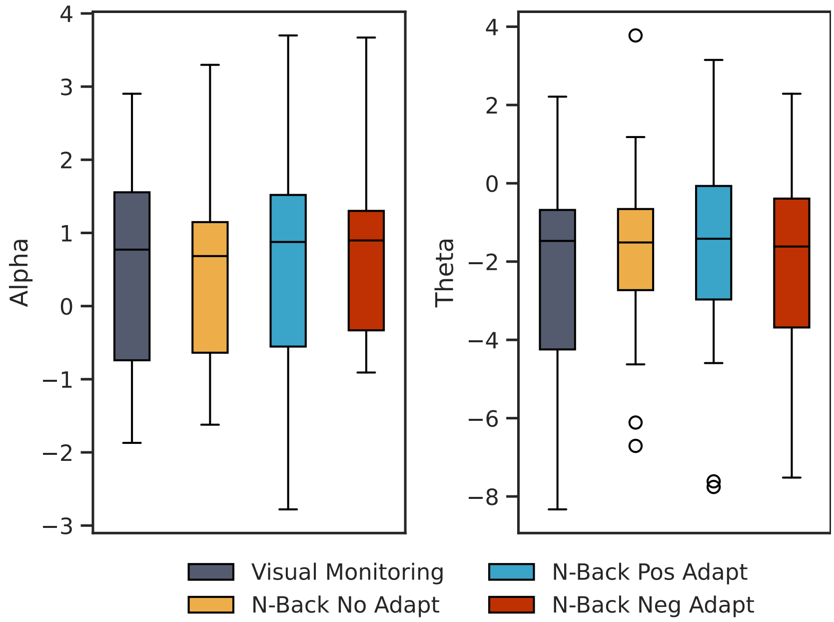


Figure 4.33: EEG Results. Boxplot representing average Alpha (left) and Theta (right) frequencies. Frequencies were obtained from the parieto-occipital channels for Alpha, while for Theta, we chose frontal channels. Values are computed for each experimental condition and normalized to the resting state.

ANOVA was conducted to examine the effect of BLOCK on ALPHA. A Repeated Measures ANOVA did not show any significant differences ($F = .45$, $p = .71$) as depicted on the left in Figure 4.33.

Theta As Theta power was not normally distributed (Shapiro-Wilk, $W = 0.96$, $p = .02$), we conducted a Friedman test indicating that the average Theta power did not change significantly across the different BLOCKS ($\chi^2 = 1.28$, $p = .73$) as depicted on the right in Figure 4.33.

Behavioral Results

Accuracy Shapiro-Wilk test showed a not normal distribution of accuracy scores ($W = 0.96$, $p = .04$). We tested the effect of BLOCK on Accuracy via a Friedman's test. We found a significant main effect ($\chi = 27.36$, $p < .001$). Post hoc comparisons with Bonferroni correction revealed that the mean accuracy in POSITIVE ADAPTATION ($\chi = 1.28, p = .73$) ($M = .88$, $SD = .06$) was significantly increased from the mean score for NEGATIVE ADAPTATION ($M = .74$, $SD = .07$), $p < 0.01$. Additionally, the accuracy in NEGATIVE ADAPTATION was significantly lower as compared to the N-BACK Block with no distractors ($M = .88$, $SD = .06$), $p < 0.01$. Results are depicted in Figure 4.34a.

Reaction Times We fitted a GLMM using REML and a nloptwrap optimizer on raw correct RTs with BLOCK (N-Back No Adaptation, N-Back Positive Adaptation, N-Back Negative Adap-

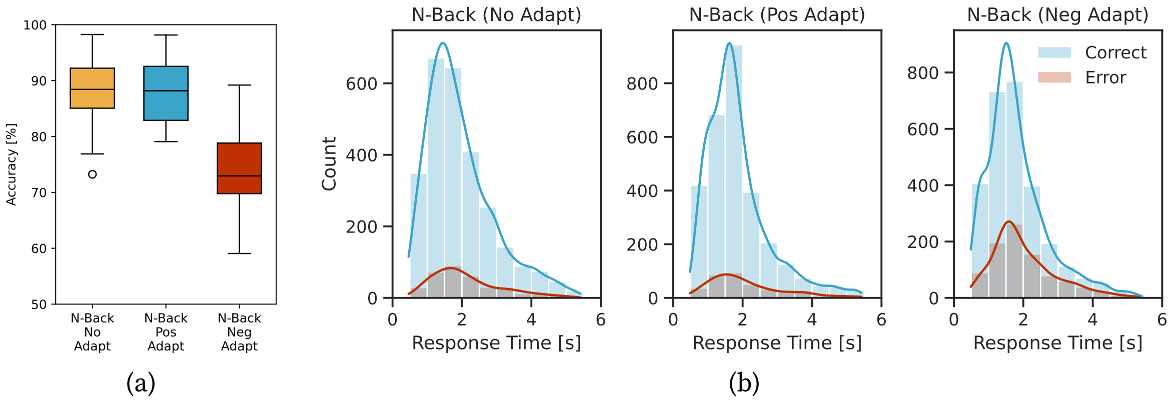


Figure 4.34: Behavioral Results. On the left (a), we present the results on Behavioral Accuracy. Here, participants significantly showed higher accuracy in N-Back and Positive Adaptation conditions as compared to the Negative Adaptation. On the right (b), we present an overview of reaction time distributions, separated by correct and error responses. No significant differences were detected in reaction times distributions.

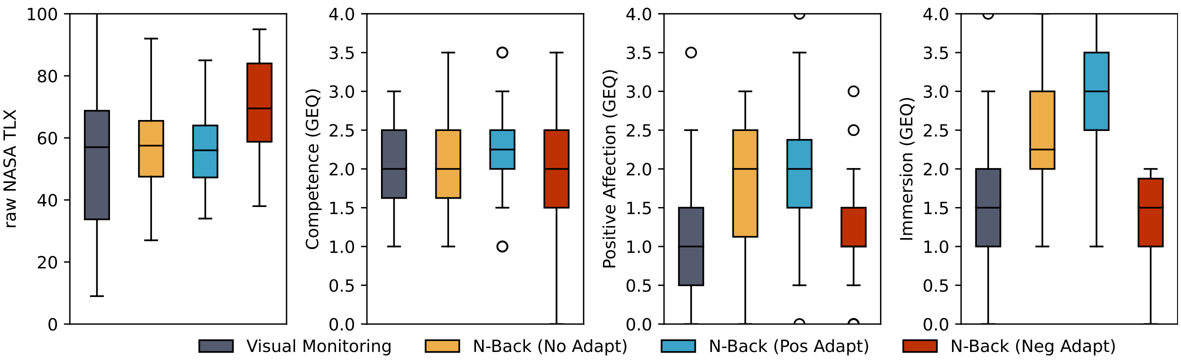
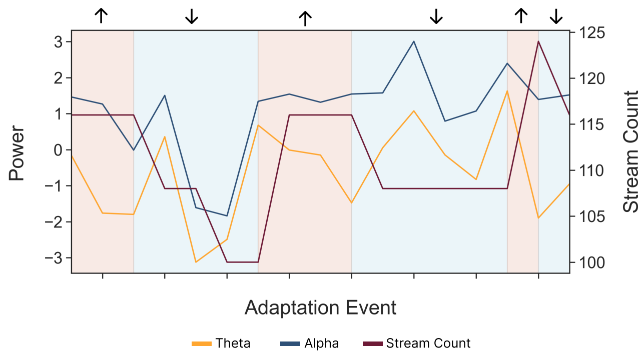


Figure 4.35: Subjective Results. Box-plots for perceived workload (NASA-TLX) and engagement (GEQ). Participants reported significantly more workload in the N-Back task with Negative Adaptation. Regarding perceived engagement, we found that participants experienced more Positive Affection and Immersion in N-Back (No Adapt) and N-Back (Pos Adapt) as compared to the Visual Monitoring task and the N-Back task in the Negative Adaptation.

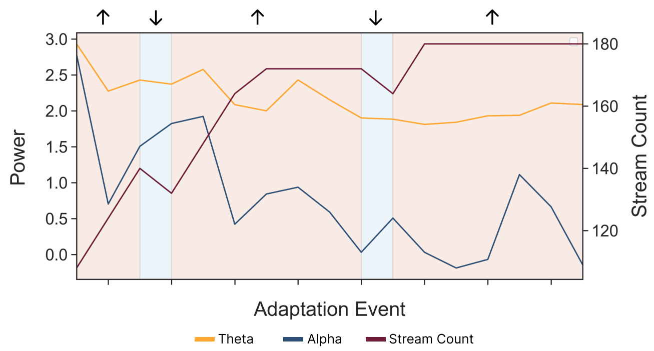
tation) as a fixed effect and participant and amount of visual distractors per trial as a random effect. We selected formula `rt ~ Block +(1|participant) + (1|distractor)`. We removed outliers by excluding values exceeding three standard deviations above the mean [46]. However, we did not report any significant results. See Figure 4.34bb.

Subjective Results

Perceived Workload As Shapiro-Wilk showed a normal distribution ($W = 0.99$, $p = .95$), an ANOVA showed that average raw NASA-TLX scores were significantly influenced by Block ($F = 4.21$, $p < .001$). Pairwise comparisons via a Conover test with Bonferroni correction mimicked results on the accuracy, showing that Negative Adaptation resulted in a significantly higher workload ($M = 70.13$, $SD = 16.97$) than Positive Adaptation ($M = 57$,



(a) Positive Stream variation for a representative participant



(b) Negative Adaptation Stream variation for a representative participant

Figure 4.36: (a) Positive Stream variation and (b) Negative Adaptation Stream variation for representative participants. Yellow and Blue lines indicate the normalized Theta and Alpha frequency bands, while the dark red line represents the Stream Variation. Colored areas indicate whether the system increased (light red) or decreased (light blue) the NPCs Stream in a 20s time window. On top of each plot, the Stream increase is depicted by an arrow pointing up ($\uparrow$), while if the Stream decreases, the arrow points down ($\downarrow$).

$SD = 13.09$), N-BACK ($M = 57.65$, $SD = 16.94$) and VISUAL MONITORING ($M = 54.81$, $SD = 27.21$), all $p < .001$. No significant differences were detected in other comparisons. Results are shown in Figure 4.35.

GEQ-Competence

The Shapiro-Wilk normality test indicated a not normal distribution for the GEQ competence scores ($W = .95$, $p = .001$). A Friedman's test revealed no significant effects ($\chi = .51$, $p = .91$), see Figure 4.35.

GEQ-Positive Affection As the Shapiro-Wilk test showed a not parametric distribution ($W = .95$, $p = .002$), a Friedman rank sum test was conducted to examine the effect of BLOCK on the GEQ Positive Affection scores. The analysis revealed a significant main effect of BLOCK on GEQ Positive Affection scores ($\chi^2 = 26.23$, $p < .001$). Pairwise comparisons with a Conover test using a Bonferroni correction showed that GEQ Positive Affection scores in NEGATIVE ADAPTATION were significantly lower ($M = 1.11$, $SD = .79$) than POSITIVE ADAPTATION ($M = 1.86$, $SD = .98$) and N-BACK ($M = 1.86$, $SD = .99$), all $p < .001$. Identical results were found in the comparisons with the VISUAL MONITORING task, in which participants reported significantly lower subjective positive affection ($M = 1.13$, $SD = .97$) than in the POSITIVE ADAPTATION and N-BACK. No differences were detected in the comparison between POSITIVE ADAPTATION and N-BACK. Results are depicted in Figure 4.35.

GEQ-Immersion The Shapiro-Wilk test indicated that the GEQ Immersion scores distribution of GEQ Immersion scores was non-parametric ($W = .95$, $p = .002$). A Friedman rank sum test revealed a significant main effect of BLOCK ($\chi^2 = 34.2$, $p < .001$). Pairwise comparisons showed that the GEQ Immersion scores in NEGATIVE ADAPTATION ($M = 1.36$, $SD = .1.36$) were significantly lower than those in POSITIVE ADAPTATION ($M = 2.98$, $SD = .83$) and in N-BACK ($M = 2.48$, $SD = .71$), all $p < .005$. Visual Monitoring task condition showed significantly lower Immersion scores ($M = 1.52$, $SD = .96$) as compared to POSITIVE ADAPTATION and N-BACK ($p < .005$). No differences were detected in the comparison between VISUAL MONITORING and NEGATIVE ADAPTATION, see Figure 4.35.

Classification

We evaluated the performance of a Linear Discriminant Analysis (Linear Discriminant Analysis (LDA)) model for predicting internal and external attention. We used EEG data from the Visual Monitoring for the External Attention label and the N-Back No Adaption for the Internal Attention label. We divided the dataset into training and validation sets in an 80/20 ratio, using a participant-wise split. The reported results include the accuracies and F1 score obtained on the test set.

Feature Extraction We extracted EEG features based on the Power Spectral Densities (PSD) via Welch's method. We computed averaged alpha and theta based on the individual frequency range computed (see section 4.5.3) and delta (0.5 - 4 Hz), beta (13 - 30 HZ) and gamma (30-45 Hz) based on the preprocessing pipeline described in section 4.5.3. All frequency values were normalized based on the Resting state data. We used electrodes chosen for adaptation for alpha and theta as in section 4.5.2. For beta, we used the same frontal electrodes as theta [455], while for delta and gamma, we based our choice on previous work in internal-external attention classification [130, 236, 571, 573]. The EEG features were computed on 20s intervals, mirroring the time window used for adaptation.

Classification Accuracy The LDA model was trained on data from a subset of participants (N=12) and validated on data from a separate set (N=5). We then evaluated the model on the remaining participants (N=5). The LDA model achieved a training accuracy of $.8$ and a validation accuracy of $.76$ when using alpha, theta, beta, delta, and gamma measures to predict internal or external attention. While we report an accuracy of $.86$ and an F1 score of $.86$ on the test data. To understand which features were most informative for predicting internal and external attention, we examined the weight coefficients of the LDA model. The coefficients indicate the relative influence of each feature in predicting attention. Specifically, a positive coefficient for a feature indicates that higher values of that measure are associated with predicting external attention. In contrast, a negative coefficient indicates that higher feature values are associated with predicting internal attention. Our results showed that the alpha measure was predictive of external attention, with a positive coefficient of $.372$. Delta power was majorly predictive of internal attention, with a negative coefficient of -1.04. The coefficients for the theta, beta, and gamma measures were $-.681$, $.281$, and $-.054$, respectively. These results suggest that alpha was specifically informative for external attention prediction, while delta and theta were indicative for internal attention.

4.5.5 Discussion

We presented a physiologically adaptive VR system that employed EEG correlates of internal and external attention to perform dynamic visual complexity adjustments to enhance task performance. We evaluated the effect of visual complexity adaptations, in the form of NPCs, on task performance, Alpha and Theta power, subjective workload, and engagement. In the study, participants performed a VR N-Back task recruiting WM resources. Here, we discuss our results regarding the outcome of our adaptive algorithms for modelling internal and external attention. Then, we envision applications for online attentional state detection and classification in VR and reflect on limitations and future work.

Internal and External Attention Modelling

When users engage in VR tasks that feature both external and internal processing components, we hypothesized that they could benefit from an adaptation that could adjust the number of visual distractors in real-time to optimize their attentional state and enhance task performance. To achieve this, we designed two adaptive VR systems based on EEG alpha and theta power to optimize external and internal attentional states, respectively.

We identified four initial hypotheses. **HP1** and **HP2** predicted that the adaptive system designed for internal attention would improve WM task performance, while the system designed for external attention would decrease task performance. Our findings supported those two hypotheses, showing that participants performed better on the visual WM task when the adaptive system optimized distractors based on internal attention (**HP1**) and performance decreased when external attention was optimized (**HP2**). These results are consistent with previous research showing that attentional resources are essential for successful WM performance as the balance between external and internal attention can significantly affect task performance [405]. When we need to recall and manipulate visual information and ultimately perform decisions, adapting task-irrelevant visual information can improve our task performance. Conversely, it could be argued that distracting information could be removed from the environment to optimize internal attention for improving task performance. However, our results show that adaptation of visual distractors based on internal attention states enhanced perceived engagement through positive affection and immersion, supporting **HP3**. Participants reported higher levels of engagement when the adaptive system optimized distractors based on internal attention. On the other side, they reported significantly lower levels when interacting with the NEGATIVE ADAPTATION. Incorporating real-time adaptation based on internal attention states into VR systems could lead to more effective and enjoyable user experiences when high-level cognitive processing is involved. Additionally, our findings highlight the importance of considering internal and external attention in designing VR systems. Optimizing one at the expense of the other may adversely affect overall user experience and performance. In fact, an increase in engagement could have impacted the increase in task performance. Positive affection, for example, has been shown to enhance the focus of attention [488]. Finally, we verified **HP4**, as participants reported significantly higher levels of perceived workload when interacting with the NEGATIVE ADAPTATION as compared to the POSITIVE ADAPTATION and to the N-Back with no distractors. This finding aligns with previous research showing that increased external attentional demands can lead to a higher perceived workload [289, 481]. The perceived workload might have been associated with the continuous need to actively filter out task-irrelevant information, which can interfere with the processing of relevant information and increase cognitive load. Our results suggest that an adaptive system that prioritizes internal attention can enhance executive performance in a VR environment. In contrast, external attention optimization can have a detrimental effect.

The results of the classification suggest that reliable decoding of internal and external attentional states in VR settings is possible, replicating similar results derived from AR settings [571, 573]. Specifically, the main features contributing to the classification were alpha for internal states and theta and delta for external states. We can therefore state to have verified **H5**.

This finding is consistent with previous research on the role of alpha as a regulatory frequency in the balance between internal and external attention. Previous work showed how alpha decreases in response to attention-demanding tasks [311]. Similarly, the role of theta power for internal attention is in line with previous literature, reflecting the maintenance of internal cognitive processes and inhibition of distracting information [499]. More interesting is the relevance derived from delta frequency band. Delta has been interpreted to act as a functional modulator of sensory afferences that can interfere with internal concentration [235]. Moreover, delta is associated with dynamic switching between external and internal attention [287], supporting their role in the inhibition of ongoing processes that can interfere with task execution. Our results align with previous research on the role of these frequency bands in attentional processes, highlighting their importance for understanding the neural mechanisms underlying attention in immersive environments.

Limitations and Future Work

We acknowledge that our work is prone to certain limitations related to the task we designed, their classification and how to improve our designed VR adaptive system.

In our study, we use the VR N-Back task, which inherently features an external shift of attention given its VR nature. This is an inherent limitation of using VR to recruit internal attention, and it must be acknowledged when designing experiments with a prominent visual component. To further evaluate the reliability of this paradigm, we suggest increasing the memory-related demands, such as increasing the amount of information held to be held in WM, i.e., moving from a 2-Back to a 3-Back VR task. Another possibility would be the addition of other internal components, such as episodic memory. Regarding the visual monitoring task, it is worth noting that while we did not explicitly verify whether participants directed their attention towards the NPCs, the task design and instructions provided to participants were based on prior research aimed at recruiting external attention [16, 120, 569]. However, we acknowledge the limitation of not implementing a manipulation check based on eye-tracking. In future work, we plan to address this limitation by incorporating eye-tracking measures to assess participants' attentional focus accurately.

On the other hand, comparing the Visual Monitoring task to a VR version of an established external attention task, such as the visual oddball task [457], would allow for better generalization of our results. These limitations and challenges are common in VR research, mainly when designing tasks that have to be ecologically situated.

Improving the generalizability of our results would support the reliability of our classification. Although we have selected tasks that theoretically recruit internal and external attention resources, our classifier could only discriminate between two tasks. Future work will address the training phase on more diverse tasks to validate our results further. Nonetheless, the high accuracy achieved in the between-participant task classification is comparable to previous work in AR [573, 574] and suggests the potential for online implementation to evaluate its performance. Specifically, LDA is a machine learning model that allows for low computation and is successful for online cognitive state detection [357]. A new adaptation mechanism could be based on this classification approach, to balance the impact of multiple features and thus increase robustness against trial-to-trial variability.

Finally, our study demonstrated that conventional methods, such as the Welch periodogram computed on a moving time window, can adequately detect temporal variations in non-stationary signals. However, more advanced signal processing techniques like wavelet analysis can further improve the detection of temporal changes [247]. Thus, implementing and evaluating wavelet analysis in future research may enhance the accuracy of attentional state classification. It is worth noting that efficient wavelet computation algorithms are available, which can be used in real-time applications [306, 609].

4.5.6 Open Science

We encourage readers to reproduce and extend our results and analysis methods. Our experimental setup, collected datasets, and analysis scripts are available on the Open Science Framework[7].

4.6 Summary

In chapter 4, we conduct a series of studies to explore VR systems' adaptability through targeted research questions.

First, in section 4.1, we examined how adaptation based on physiological engagement correlates can significantly enhance task efficiency in VR environments and impact different physiological signals. Here, we explored the adaptation of secondary task difficulty in response to real-time physiological arousal signals, i.e. raw EDA (**RQ4**). Findings suggest that participants experienced increased engagement and reduced mental workload when task difficulty matched their arousal levels.

In section 4.2, we analyzed how these task difficulty adaptations affect attention by examining EEG, ECG and EDA changes (**RQ5**). The results indicate that theta, beta, and phasic EDA can be employed to verify the effect of adaptations on arousal and engagement. In section 4.3, we switched our focus from adaptations acting on the main task and focused on adjustments to the visual complexity of task-irrelevant features of the VR environment **RQ6**. Our work showed that tailored environmental complexity based on EDA correlates with engagement and can enhance task performance by maintaining an ideal cognitive load. Continuing the work on environmental adaptations, in section 4.4, we assessed how changes in visual complexity influence attention allocation. Notably, alpha and theta demonstrated significant reactivity to variations in visual complexity adaptations, suggesting their future implementations as indicators of attention allocation and engagement (**RQ7**). In section 4.5, building up on our previous work, we present a VR adaptive system based on EEG correlates of attention allocation (alpha) and engagement (theta) to dynamically adjust visual complexity and support task performance in a WM task. Visual complexity adjustments based on alpha and theta bands allowed for modulation of task-irrelevant elements adaptation and increased WM task performance. Furthermore, we successfully classified the attentional state based on EEG data in a VR N-Back and Visual Monitoring tasks (**RQ8**).

[7] `https://osf.io/ar4fs/`

5

TOWARDS WEARABLE PHYSIOLOGICAL COMPUTING

> *'The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."*
>
> **– Mark Weiser.**

In the last chapter, we present SensCon [100], an innovative physiological sensing system integrated within VR controllers. SensCon offers a more practical and user-friendly alternative to traditional medical-grade devices, which are often cumbersome and disrupt user immersion.

The primary focus of this chapter is validating reliable physiological sensing systems in MR interfaces (**RQ9**). SensCon is developed as a response to this challenge, embedding EDA and PPG sensors directly into VR controllers to provide real-time, unobtrusive physiological monitoring. This integration aims to validate the potential of using embedded systems in VR to achieve reliable and accurate physiological measurements without sacrificing user experience. We further explore broader implications of integrating physiological sensors into intelligent VR systems.

This chapter is based on the following publication.

Francesco Chiossi, Thomas Kosch, Luca Menghini, Steeven Villa, and Sven Mayer. 2023. SensCon: Embedding Physiological Sensing into Virtual Reality Controllers. In *Proceedings of the ACM on Human-Computer Interaction*, 7(MHCI), 1-32. ACM, New York, NY, USA.
`https://doi.org/10.1145/3604270`

5.1 SensCon

VR allows manipulation and control of diverse stimuli in virtual environments as close as possible to real-world ones. Besides the immersion in VR, different responses can be triggered for users, including engagement [263], presence [523], arousal [172], and workload [360]. Evaluating these factors is pivotal in User Experience (UX) research in VR. Researchers often use post hoc assessments (e.g., questionnaires) to evaluate such responses since real-time user input is challenging or even hardly possible [580]. Recently, implicit real-time assessments such as physiological measures, have become increasingly popular to overcome these drawbacks.

They offer real-time control and quantitative measures of user behavior when exposed to VR and adaptation to virtual environments. Peripheral physiological responses, such as EDA and PPG, are indicative metrics for physiological arousal [595], emotional valence [372], or cognitive workload [214].

Current scientific measurement standards require medical-grade devices, whose signal recording quality comes at the cost of limited ecological validity [286, 469] and unintuitive VR synchronization [616]. Recording physiological data in VR requires ad-hoc preprocessing pipelines to disentangle physiological signals from movement artifacts, increasing demands for technical synchronization and data analysis [215]. To lower the entry hurdle, researchers designed [612] and prototyped [48] physiological sensors embedded into VR systems. For instance, Luong et al. [361] estimated the mental workload in a VR flight simulation in real-time using multiple physiological inputs embedded in the HMD. However, available validation studies are either limited to consumer-grade devices [387] or are not applied to VR environments [568].

Here, we present SenCon, a physiological sensing system that incorporates EDA and PPG measurements into VR controllers allowing for mobile and flexible interactions, cf. [2, 3]. In contrast to bulky medical-grade equipment, SenCon encompasses an easy and ready-to-use system designed to minimize the user's preparation time. We ran a study investigating how users hold VR controllers to inform sensor placement, see subsection 5.1.1. Then, we evaluated the SenCon regarding usability, user experience, and measurement accuracy compared to medical-grade devices across six VR tasks, see subsection 5.1.3. We found significant differences concerning usability and user experience favoring SenCon compared to medical-grade equipment. Furthermore, we find that SenCon delivers EDA and PPG data that can be used to evaluate user experience.

5.1.1 Study 1: Finger Placement and Hand Grasp in VR Interaction

In the past, many inside and outside tracking methods have been pioneered to gain various states from the user, e.g., [2, 3]. However, the reason trends point towards inside tracking as the environment does not need to be equipped with technology allowing for improved flexibility and mobile applications. Thus, integrating physiological sensing into VR systems aligns with current trends in VR development. For example, recent research has explored the integration of cameras into VR controllers for optimal body and hand-tracking [3, 538], enabling the design of intelligent mobile standalone VR systems. This trend is already being implemented in commercial VR systems, such as the camera embedded in the Meta Quest Pro [61]. Second, while others showed that placing sensors in the headset is feasible (e.g., [361, 618]), when considering peripheral physiological measures such as EDA and PPG there are cautions to be taken for validation purposes. We wanted to maximize the signal agreement between SensCon and the gold standard measurements used in Study 2 (i.e., finger PPG and finger EDA), implying peripheral measures from the hand/finger. That is, using different measurement locations would have implied less fairly comparable measurements, also due to the higher eccrine sweat glands density in the hands compared to other body sites [156] and the different pulse transit time implied by hand and head PPG sites [501]. By embedding physiological sensors into VR controllers, we can enhance the user experience in VR environments and enable the design of novel physiologically-aware VR applications.

To ensure adequate sensor placement in the VR controller, we conducted a study to determine grasping location while holding an HTC VIVE controller and simultaneously pressing three key buttons (Trigger, Grip, and Trackpad buttons) following the approach by [338]. We were especially interested in where the grasping hand made contact with the controller, which we analyzed by measuring the contact surface when participants grabbed the VR controller. We computed and analyzed the degree of overlapping across participants.

Procedure

First, we asked participants to handle the controller until they felt comfortable pressing the Trigger, Grip, and Trackpad buttons without moving their hands. Next, we outlined the participant's hand with a fine-pointed white pen. In this way, the contour of the participant's handprint was drawn on the VR controller. Next, the participant removed the hand from the controller to imprint it with colored paint on a flat plate to optimize the spread. The participant then held the VR controller, trying to follow the previously outlined contour, optimizing the grasping to press the buttons. Finally, the participant released the grip on the VR controller. The researcher then photographed the front, bottom, left, and right sides of the controller five times using a professional camera (Canon EOS 50D, 4752 $\times$ 3168 pixels) positioned on a solid stand that was kept constant throughout the photographs.

The controller was placed inside a 3D-printed holder on top of a table to ensure orthogonality for the axis of the table and the camera. We placed a QR code on each side for orthogonal image alignment processing to allow optimal perspective transformation.

Participants

The study was performed in a quiet and distraction-free environment, where participants sat and could focus on handling the controller. Twelve participants, six identified as female and six as male ($M_{age} = 27.5$, $SD_{age} = 2.74$) took part in the study. We asked participants to hold the controller with their dominant hand [262], with ten using their right hand and two using their left hand, respectively. We measured three standardized hands metrics following the guidelines of the Human Engineering Design Data Digest [451]: (I) handbreadth ($M_{cm} = 8.4$, $SD_{cm} = 0.32$), i.e., distance between the two metacarpal bones (metacarpal-phalangeal joints), (II) hand length ($M_{cm} = 17.98$, $SD_{cm} = 0.66$), i.e., distance from the base of the hand at the wrist crease to the tip of the middle finger and, (III) hand circumference ($M_{cm} = 18.91$, $SD_{cm} = 0.89$), i.e., circumference of the hand measured around the knuckles (metacarpal-phalangeal joints). In this way, we respectively covered 50th, 95th, and 99th percentile for both men and women in hand measures as per the Human Engineering Design Data Digest [451].

(a) Front (b) Back (c) Right (d) Left

Figure 5.1: Heat maps showing the area of the VR controller with the highest contact overlap during interaction in Study 1.

Image Processing

We preprocessed the images using Python library OpenCV (Version 4.3) [69] and Adobe Photoshop (Version 23.4.2)[1]. We placed four QR codes on the front, back, and left and right side of the VR controller holder for coarse image alignment [298]. On such images, we applied a perspective transformation to align and overlap the QR codes from every side and, consequently, the pictures of contact locations on the VR controller. As a final step in the image processing phase, we applied Lens correction in Adobe Photoshop to all images using a distortion correction model to improve pictures' parallelism, size, and rotation. Next, we masked all areas where the fingers touched the controller, allowing us to calculate the area for each user.

Results

We extracted the area touched by participants during the comfortable holding of the controller. For this, we aligned the areas touched and measured the area which could theoretically be touched (the shaft of the controller). We found that 64.1% of the front, 73.1% of the back, 87.1% of the right side, and 57.1% of the left side will be touched by at least one participant, see Figure 5.1. While this is a good first indication of surface coverage, we next calculated the area which was touched by all 12 participants for successful sensor placement. Here, we found that 14.1% of the front, 11.3% of the back, 47.8% of the right side, and 5.4% of the left side were touched by all participants. The area with the perfect overlap is shown in Figure 5.1c – the middle of the right side of the controller.

[1]`https://www.adobe.com//products/photoshop.html`

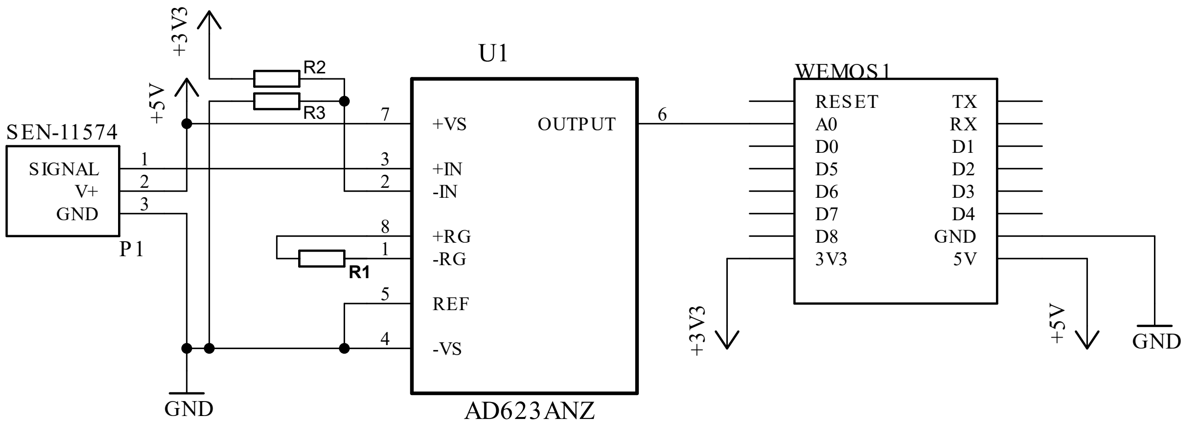


Figure 5.2: The schematic overview of the SensCon. Data are acquired via PPG and EDA sensors in the VR controller. Physiological data are streamed within a User Datagram Protocol (UDP) protocol to the experiment control PC. Finally, the VR environment is displayed to the user via a wired HTC VIVE headset.

Of note, such images are from the perspective of holding the controller on the right side. However, due to the symmetric anatomy of the human hand, the results for the left hand are the same but mirrored. As such, we can say that the optimal area for a sensor is not per se on the right side of the controller but where the palm of the hand touches the controller (right side for the right hand and left side for the left hand, respectively).

5.1.2 The SensCon System

With SensCon, we combine commodity VR controllers with PPG and EDA sensors for a wide range of future applications using physiological sensing and thus, adapt to the users' state. We used two HTC Vive controllers as a base to host the sensors to accomplish this.

Hardware Implementation

Based on our results from Study 1, we considered that a sensor with a low surface area, such as a PPG sensor, would have been best placed on the right side of the controller, or in other words, below the palm of the hand, see Figure 5.3. Finally, we noted that the left side of the controller consisted of a large surface while not being constantly covered by the fingers. Thus, we selected it as the ideal location of the second EDA electrode. In sum, we mounted the two EDA electrodes on both sides of one controller (see Figure 5.3) and the PPG sensor on the well-covered right side of the second controller, respectively (see Figure 5.3).

With this goal in mind, we integrated the PPG and EDA sensors into two HTC Vive controllers. For the communication and local sensor control unit, we used an ESP 8266 D1 Mini[2] microcontroller offering both WiFi and Bluetooth connectivity. Thus, we sent the PPG and EDA

[2] www.openhacks.com/uploadsproductos/tutorial_nb.pdf

Figure 5.3: EDA and PPG sensors integrated into VR controllers. In the first picture, EDA sensor is integrated into a controller. In the second figure, two copper stripes measure the EDA when a user picks up the controller. In the third figure, PPG sensor integrated inside the controller. In the fourth figure, the PPG sensor measures the user's HR by implicitly pressing the palm against the sensor.

data via a WiFi and UDP connection to a Lab Streaming Layer[3] (LSL) server for signal time synchronization and data logging via Arduino. Figure 5.2 shows the schematics of SensCon. We also added an additional micro-USB port on the underside of the VR controller (see **??** and **??**), allowing for fast deployment of updates and acting as a debug terminal.

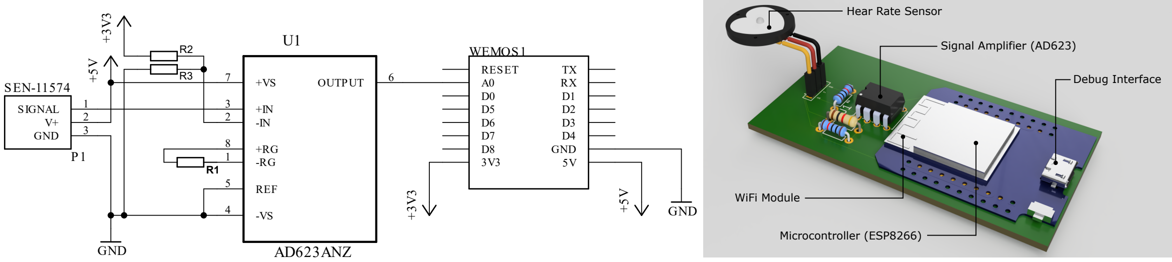
The controllers have built-in batteries with 960mAh at 3.85V. However, this is under the minimum required 4V for the V_{in} of the microcontroller. Thus, we incorporated a "step-up DC" in the controller while losing playtime by draining the built-in battery. Thus, we opted to add a power bank connected to the ESP flash port instead. Pulse Sensor PPG and Groove GSR sensors required a stable power consumption of $.08A$, as measured via a USB digital multimeter. We allowed for long playtime and easily swappable batteries.

Sensing EDA

We use a Groove galvanic skin response (GSR) sensor[4] to measure EDA with a sample rate of 192 Hz. First, we disassembled an HTC Vive Controller and integrated the sensor in the handle (see Figure 5.3). We then soldered the pins of the connectors to two pieces of copper band that we adhered around the handle (see Figure 5.3). The resistance between the contacts is measured in one hand by closing the two copper contacts when taking the controller into the hand. The GSR sensor captures micro voltages (μV) between the distal phalanges as a measure of skin resistance utilizing the µV input in Ohms (Ω). The formula for skin resistance

[3]https://github.com/sccn/labstreaminglayer

[4]https://wiki.seeedstudio.com/Grove-GSR_Sensor/



(a) Circuit Schematic (b) 3D Model

Figure 5.4: PPG signal conditioning: we employed an instrumentation amplifier to increase the amplitude of the sensor's raw signal. This configuration utilizes a voltage divider, with the gain regulated by a single resistor (R1). Specifically, (a) depicts the circuit powered by the WEMOS D1 mini platform, which integrates an ESP8266 micro-controller. The heart-rate sensor's signal is processed by the AD623 instrumentation amplifier and subsequently transmitted to the WEMOS platform. Alternatively, (b) illustrates the circuit diagram and its various components.

computation is provided by the manufacturer and shown in Equation 5.1:

$$SR(\mu\Omega) = ((1024 + 2 * \mu V) * 10,000)/(512\mu V) \tag{5.1}$$

Therefore to compute EDA, we computed the reciprocal of resistance, i.e., conductance, measured in (μS) as in Equation 2:

$$EDA(\mu S) = (512\mu V)/(\mu\Omega) = ((1024 + 2 * V) * 10,000) \tag{5.2}$$

Sensing PPG

We integrated a Pulse Sensor[5] by punching a circle-shaped hole located at the palm hand position of the controller (see Figure 5.3). We placed the Pulse Sensor on the left side of the HTC VIVE controller as depicted in Figure 5.3. In this location, we allowed for 100% of surface overlapping across participants. Then, we disassembled a controller to integrate the D1 mini microcontroller, and we connected it to the PPG sensor.

The Pulse Sensor is an optical HR sensor (i.e., PPG) that features on the frontal side of the sensor, an APDS-9008 Light Photo Sensor, and a reversed mount LED. In this way, the reflected green light ($\sim 550nm$) from the LED through the fingers is measured via the photosensor. On the back of the module, an MCP6001 Op-Amp microchip consists of three resistors and capacitors that form an R/C filter circuit. This circuit is used for noise cancellation and amplification to avoid time delay in pulse recordings. The module operates from a 3.3 to 5V DC Voltage supply with an operating current of 4mA. The desired sampling rate is set at 250 Hz.

[5] `https://pulsesensor.com/`

Open Source

We open-sourced the full implementation of SensCon. First, we provided a list of the materials needed to implement SensCon. Second, we provided the schematics for the electronics housed in the VR controllers. Third, we provided a manual to assemble SensCon. Fourth, we open-sourced the code to retrieve the sensor data and send it onto the LSL stream. Finally, we provided a Unity package to receive the real-time LSL stream[6].

5.1.3 Study 2: System Evaluation

Study 2 aimed at assessing the UX and measurement accuracy of SensCon, as compared to medical-grade devices measuring EDA and PPG-based HR. We divided the description of the study into two parts. The first part investigates usability, acceptance, and UX by letting users freely use either SensCon or VR controllers with attached medical-grade equipment, and by collecting subjective measures using questionnaires after each condition.

The second part evaluates the measurement accuracy across the included VR tasks. That is, it presents a within-subjects study focusing the agreement between SensCon and medical-grade physiological sensors. The following research questions guided our evaluation:

RQ1 Does an embedded physiological sensing system provide higher usability and UX in VR environments when compared to medical-grade devices?

RQ2 Does an embedded physiological sensing system allow for comparable signal quality and outcome measures to those obtained with medical-grade devices?

Independent Variable: Physiological Sensing System

We considered the physiological sensing systems, consisting of SensCon and the medical-grade devices as the only independent variable for both parts of our study. We evaluate if the sensors integrated into SensCon can act as a usable alternative to medical-grade devices (i.e., PPG finger-clip sensor and EDA Ag/AgCl electrodes measured from the index and middle fingers) over six different tasks.

Measurements

The first part of the study focused on the usability, acceptance, and UX aspects in the first part of the study. Participants were equipped with either SensCon or medical-grade devices while walking around. Then, after using each system, they filled in three self-report scales. We started with the System Usability Scale (SUS) [344] questionnaire to measure system usability. We then proceeded with the Unified Theory of Acceptance and Use of Technology

[6]`https://github.com/mimuc/Senscon/`

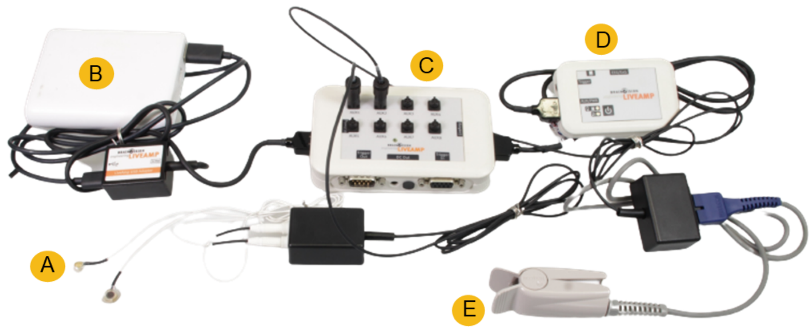


Figure 5.5: The medical-grade system used as gold standards for evaluating SensCon-based EDA and PPG measurements. The EDA electrodes (A) and PPG finger clip sensor (E) are connected to the AUX ports of the Sensor & Trigger extension (STE) (C). For signal amplification and streaming to the acquisition PC, we use a LiveAmp amplifier (D) and a power bank for energy supply to the STE (B). All components were placed inside a pouch and comfortably worn by participants across the experimental sessions

(UTAUT) [566] questionnaire to assess the acceptance of the two systems. Finally, we administered five seven-point Likert items to obtain an overview of the UX using the physiological sensing systems. The self-report scales are depicted in Table 5.1.

In the second part of the study, we measured both EDA and HR using SensCon and the medical-grade devices at the same time throughout six different tasks. SensCon measures EDA upon making contact with the user's left hand while the medical-grade device obtains the measures by clipping two Ag/AgCl electrodes on the index and middle finger of the same hand. For EDA data acquisition, we followed the guidelines for HCI community by [26]. SensCon obtains the PPG signal from the user's right hand's palm. The medical-grade device assesses the PPG through a finger clip placed on the index finger.

Apparatus

For the medical-grade recordings, we chose the GSR module (BrainProducts GmbH, Germany)[7] and blood pulse sensor for finger PPG (Nellcor DS-100A, Nellcor, USA)[8] as the gold-standard physiological sensing systems. Those are two sensors within a computerized recording system that encodes up to eight channels with a sampling frequency of up to 1000 Hz. Both devices are shown in Figure 5.5. To achieve optimal time synchronization across measures, both physiological sensing systems are connected to the Stimulus and Trigger Extension (STE; BrainProducts GmbH, Germany). We acquired and synchronized the signals at 250 Hz. Therefore, this study considers the GSR module and the PPG finger-clip sensor the gold standard measurement devices.

[7] https://www.brainproducts.com/files/public/sensor-tutorial/Content/Topics/1.GSR/Acquisition/LiveAmp%20GSR.htm

[8] https://www.brainproducts.com/files/public/sensor-tutorial/Content/Topics/2.Photoplethysmogram/Acquisition/LiveAmp%20PPG.htm

Table 5.1: Seven-point Likert questions used to assess the user experience and pleasantness of both SensCon and the medical-grade sensors.

ID	Question
Q1	*It was very comfortable to use the system.*
Q2	*It was very annoying to use the system.*
Q3	*It was very pleasant to use the system.*
Q4	*It was very difficult to use the system.*
Q5	*I want to use the system in my everyday VR experiences.*

For the SensCon recordings, we integrated an EDA (see section 5.1.2) and a PPG sensor (see section 5.1.2) into two separate HTC Vive controllers. The controller containing the EDA sensor was held with the participant's left hand, while the controller with integrated PPG was held with the right hand. Prior to analyzing the data, we calibrated each PPG measurement by subtracting an estimate of the mean offset (i.e., the systematic difference) between the two PPG signals acquired from the medical-grade device and SensCon. This was done by using data recorded in seated resting conditions from three further participants ($M_{age} = 25$, $SD_{age} = 2.45$) that were not included in the main sample for Study 2. We obtained an average offset of 11.1 bpm ($SD = 12.1$), which was used to correct the signal acquired from SensCon. All collected data of SensCon and the medical-grade equipment were then sent over the network and synchronized on the same computer using LSL. We used an HTC Vive to display the virtual environments in the second part of the study. The participants' movements were tracked using two VIVE lighthouses.

Procedure

Upon arrival, participants received a written briefing on the experimental procedure and provided their informed consent. Afterward, they provided their demographic data, including age, self-identified gender, weight, and height. We then started with the first phase of the study to compare the usability of SensCon and the medical-grade devices. In two separate conditions, participants were either equipped with SensCon or both the EDA and PPG medical devices without being immersed in the VR environment. We asked participants to walk around with the hardware and perform several actions, including hand movements and pointing gestures. Then, participants filled in the SUS and UTAUT scales and the custom Likert items to assess the UX. The same procedure was repeated for the other physiological system modality, and the order presentation of physiological sensing devices was counterbalanced across participants.

The second part of the experiment evaluated the signal quality between SensCon and the medical-grade devices. We instructed the participants to hold the SensCon controllers in their hands while the medical-grade devices were simultaneously applied (i.e., EDA on the left index and middle finger, PPG fingerclip on the right index finger). Before data acquisition, the participants were seated and were allowed to visually explore the VR scene for two minutes by moving their heads. Then, participants started to engage in the six experimental tasks

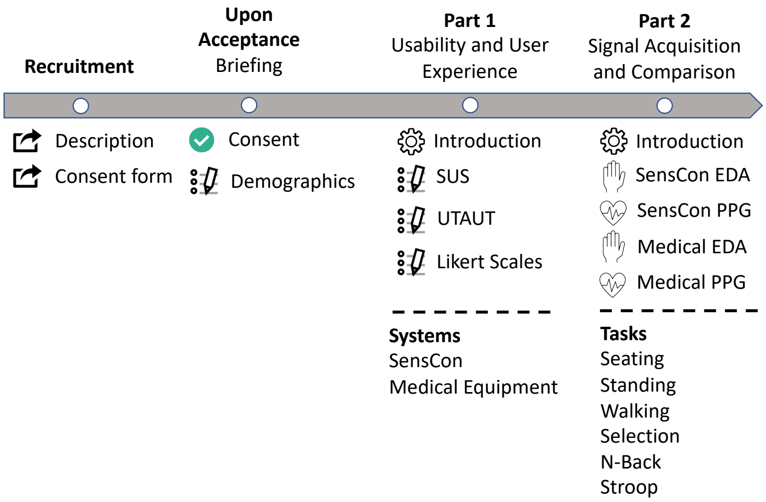


Figure 5.6: Graphical description of the study procedure. The study consisted of two parts assessing the user experience and signal quality, respectively.

while the physiological signals were recorded in real-time. The procedure of the study is illustrated in Figure 5.6.

Tasks

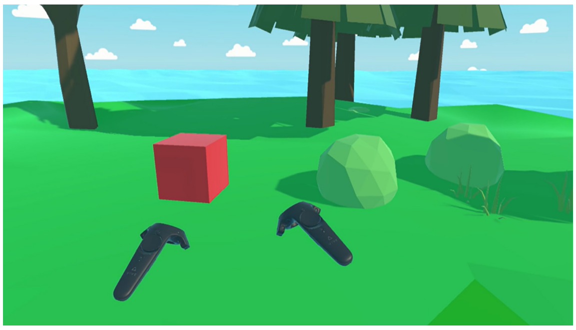
We used established tasks that have been shown to elicit traceable changes in autonomic responsiveness and particularly in the signals of interest [82]. We implemented the six tasks using Unity3D, and we used an HTC Vive headset to visualize the environment (see Figure 5.7). In between the tasks, participants underwent a break of six minutes to allow for physiological re-adaptation (i.e., recovery to the basal physiological levels). The six tasks described below were performed in the same following order by all participants.

Seating Participants were comfortably sitting on a chair while immersed in the VR environment (see Figure 5.7a). They were asked to position their hands on their thighs without moving them while holding the controllers in their hands. We chose this condition as seating rest has been recommended and widely used as a baseline condition in previous research [82].

Active Orthostatic Test The active orthostatic test evaluates the effects of sudden postural changes and it is often used to assess autonomic responsiveness [217]. This condition was shown to affect PPG measures in previous research [89, 350, 387, 435]. Participants were required to stand and maintain an upright position over three minutes in the VR environment.

Slow Walking Participants were asked to naturally walk along a path designed in the VR environment to maintain an upright position. That is, they walked over a guidance line in the

(a) Seating, standing, and walking task.



(b) Selection task.

(c) N-Back task.

(d) Stroop task.

Figure 5.7: The six different tasks of our validation study. Each task had a total duration of three minutes and was performed one after the other. The first five tasks took place in a low-poly environment while the Stroop was performed in a separate virtual room.

virtual environment. This task was designed to evaluate the effect of natural body movement on data quality, as done in previous studies [387].

Selection Task Participants were required to select yellow spheres and red cubes within a selection task. Then, they must touch the yellow sphere with the left VR controller and the red cube with the right controller, see Figure 5.7b. Spheres and cubes were randomly presented in a sequence. The movements were sequentially guided. This task was included to test the influence of hand movements on the signal quality together with the general movement of the entire body.

VR N-Back Task The N-Back task is an established task that recruits working memory resources and thus, evokes physiological and psychological stress responses [207]. It has been used in psychophysiological research with both healthy [449] and clinical populations [259]. We adapted the VR N-Back task from Chiossi et al. [110], requiring participants to place a sphere, which was initially presented over a pillar, into one of two different baskets. The sphere should be placed in the right basket when it had the same color as the sphere presented N times before, whereas when the color was different, the sphere should be placed in the left bucket. The N-Back level (i.e., N refers to the amount of steps, in this case spheres, in the sequence) was set at 2. To keep the level of engagement high, we showed an accuracy index based on the last 20 balls placed, and participants were asked to maintain a performance index above 90%, see Figure 5.7c.

VR Stroop Task The Stroop task is an established paradigm that investigates cognitive control and interference (i.e., inhibition of automatic behavioral responses due to a color-word interference) [596]. In addition, it is widely used in psychophysiological research as a cognitive stress test eliciting autonomic responses [32]. We used the VR concept proposed by Gradl et al. [214] as it was shown to elicit increased levels of HR and EDA indicators. In this VR version of the Stroop task, participants were surrounded by six walls "painted" in a changeable color and they were asked to select the wall color matching the feature of interest (i.e., either color or word). For example, in incongruent trials, if the word "Green" was colored in yellow and the color was the relevant feature, then participants shoud select the yellow wall with the controller, see Figure 5.7d. In contrast, if the COLOR is the relevant feature, they should select the purple wall.

Participants

Twelve participants, six identified themselves as female and six as male ($M_{age} = 30$, $SD_{age} = 1.81$), voluntarily participated in the study. None of the participants had a medical history of psychiatric or neurological diseases, color blindness or not assuming medications affecting the autonomic system. All participants provided written informed consent and received a monetary compensation of 10 Euros. Participants were required to avoid practising strenuous physical exercises, smoking, and consuming coffee over the three hours before the experimental session.

Results

First, we evaluated the measurement accuracy of two sensors for measuring EDA and HR embedded in a VR controller across tasks. We used frequentist factorial analysis of variance (ANOVA) or ART ANOVAs [599] accounting for the non-normality of the data distributions. Moreover, to draw meaningful information from null results, we also reproduced the analyses with a Bayesian approach [471]. That is, we performed a Bayesian ANOVA using default Cauchy's priors and 10,000,000 iterations with TASK and participant as random effects.

Second, we better evaluated the agreement between SensCon-based measurements and those obtained with the corresponding medical-grade devices by means of Bland-Altman analyses [8, 57]. Bland-Altman plots and related statistics are the most established analytical tool for comparing two methods providing quantitative measurements, and they are widely used to evaluate how accurately a new method measures a specific signal compared to a reference method. In Bland-Altman analyses, measurement accuracy is established as the agreement between the new method (here, SensCon) and the gold standard (here, medical-grade devices), expressed in terms of systematic bias and random error. The former represents the most likely difference to occur, whereas the latter quantifies the range within which most differences are expected to lie, as indicated by the limits of agreement (LOAs) [370]. Specifically, we conducted separate analyses for each signal and task following the procedures described by [386]. That is, we accounted for cases of proportional and heteroscedastic differences by representing bias and LOAs as a function of the range of measurement. Both bias and LOAs were computed with their 95% confidence intervals based on parametric bootstrap with 10,000 replicates. Bland-Altman analyses supersede classic correlation analyses as the latter only describes the linear relationship between measures but not their agreement, possibly providing misleading results [206], especially considering that data with high correlations might result in a low rate of agreement [155]. Thus, we grounded our work on Bland-Altman analyses.

Finally, we supplement the results coming from our validation analysis with t-tests or Wilcoxon tests for each self-report scale, i.e., Usability (SUS), Acceptance (UTAUT), and Likert Scales on UX.

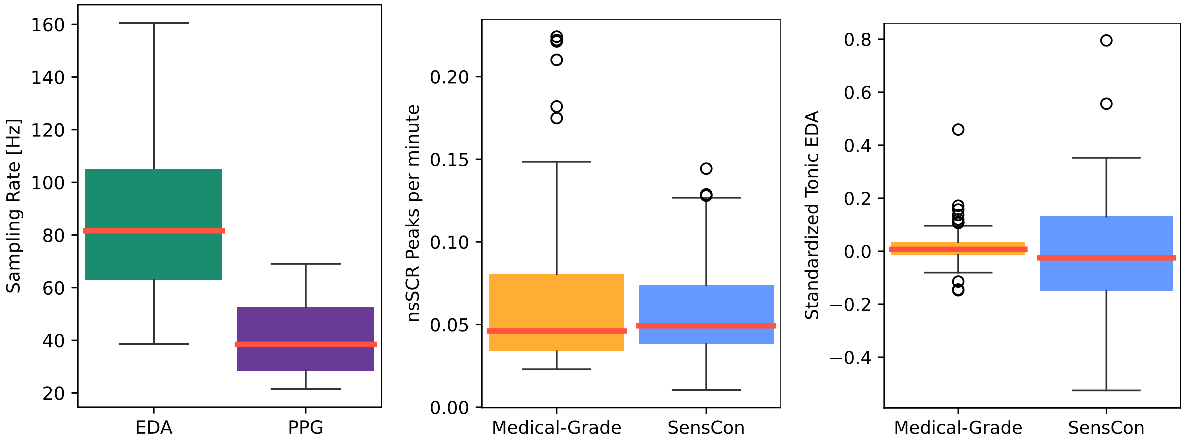
Electrodermal Activity EDA data both from gold-standard device and SensCon were processed using the Neurokit Python Toolbox [367]. We first applied a 3 Hz, high-pass, fourth-order Butterworth filter to remove high-frequency noise. We then decomposed the signal into tonic and phasic components by using the non-negative deconvolution analysis [43]. Lastly, we extracted peaks from the decomposed signal using a threshold value of $0.05\mu S$ [453]. These preprocessing steps allowed us to compute the average SCL and the mean frequency of not specific SCRs [83].

First, we analyzed the reliability of the acquired signal. Here, we calculated the effective sampling rate and found that the EDA sensor delivers samples on average with 86.7 Hz ($SD = 29.5Hz$, $min = 38.6$, $max = 160.5$), see Figure 5.8a.

We analyzed the signal in the time domain based on the recorded data by comparing their standardized signals. First, we analyzed the nonspecific SCRs (nsSCRs) peaks per minute. As the normality assumption was not met (Shapiro-Wilk: $W = .794$, $p < 0.001$), we relied on ART ANOVA [599], showing no significant differences between systems in *nsSCRs* after controlling for TASK ($F(1, 99) = 0.109$, $p = .741$), see Figure 5.8b. A Bayesian ANOVA indicated anecdotal

Table 5.2: An overview of dependent variables analyzed via factorial ANOVA or ANOVA ART and Bayesian ANOVA.

	Frequentist Testing					Bayesian Testing	
	Normality		System				
	W	p	F	p	η^2	BF_{10}	$Error$
nsSCRs freq.	.794	<.001	.109	.741	.001	.380	.24
SCL	.917	<.001	.257	.613	.003	.348	.21
HR	.987	.328	.765	.405	.012	.876	.14



(a) Sampling Rate (b) nsSCRs (c) SCL

Figure 5.8: Results for signal accuracy for SensCon EDA. a) the actual sampling rate of SensCon EDA. b) the nonspecific SCR (nsSCR) detected peaks per minute. c) the standardized tonic EDA acquired with both systems.

evidence for the null hypothesis H0 ($BF_{10} = .380$) with Task and participants as random factors.

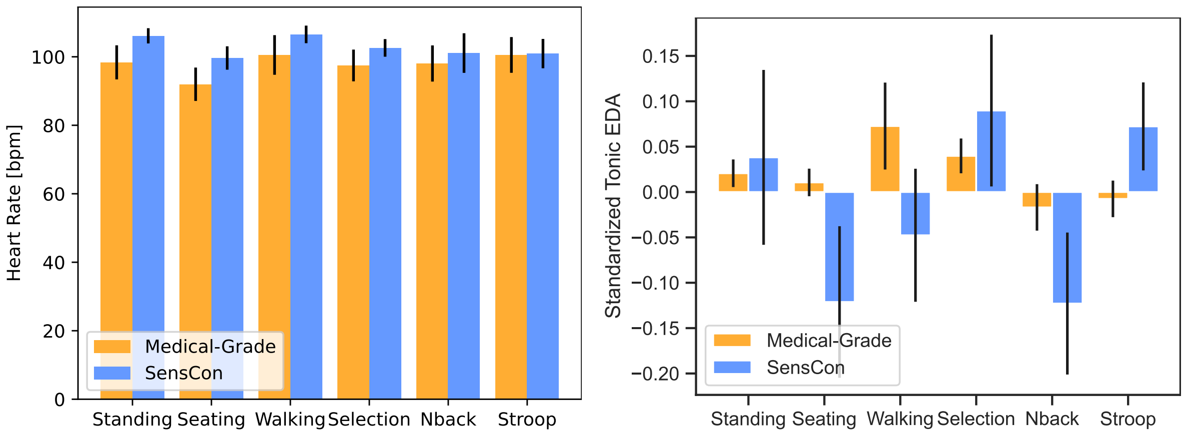
Finally, we computed the standardized SCL, accounting for interindividual differences [41, 229, 362]. As the normality assumption was not met ($W = .917$, $p < 0.001$), we relied on ART ANOVA [599], showing no significant differences between systems in *nsSCRs* after controlling for Task ($F(1, 99) = 0.257$, $p = .613$), see Figure 5.8c and 5.9b. A Bayesian ANOVA indicated anecdotal evidence for H0 ($BF_{10} = .348$) with Task and participants as random factors.

Photoplethysmography We based our evaluation on PPG-related measures in the time domain, focusing on HR. As with EDA, the measured PPG values were processed using the Neurokit Python Toolbox [367]. We first applied a third-order Butterworth filter from 0.5 to 8 Hz.

We analyzed the availability of the signal. Here, we calculated the effective sampling rate and found that the PPG sensor delivers samples on average with $41.4Hz$ ($SD = 13.7Hz$, $min = 21.5$, $max = 69.1$), see Figure 5.8a.

For the HR analysis, we computed the overall differences between SensCon and the medical-grade devices. As the normality assumption was met ($W = .987$, $p = .765$), we relied on the ANOVA, showing no significant effect of System on *HR* after controlling for Task ($F(1, 9) = .765$, $p = .405$), see Figure 5.9a. A Bayesian ANOVA indicated no evidence for H0 ($BF_{10} = .876$) with Task and participants as random factors.

Bland-Altman Analysis



(a) HR (b) SCL

Figure 5.9: Mean HR (a) and Mean SCL (b) across Tasks for SensCon and medical-grade devices. Differences computed between SensCon and Medical Grade devices for HR and EDA Tonic were not significant.

Electrodermal Activity Bland-Altman statistics for SCL and nsSCRs are reported in **??** in the Appendix and the corresponding Bland-Altman plots are visualized inFigure 5.10 and Figure 5.11. SCL showed proportional biases in most tasks, implying that the magnitude of the differences between systems was dependent on the size of the measurements. In the Seating, Standing, Selection, and Stroop tasks, SensCon showed the highest accuracy for intermediate SCL values, whereas it systematically underestimated lower SCL values and overestimated higher values, respectively. In contrast, only relatively high SCL values were measured without bias in the N-back task, showing underestimations for lower SCL values. The Walking task was the only condition showing uniform and nonsignificant bias, although we found wider LOAs for higher SCL values (i.e., heteroscedasticity). Heteroscedasticity was also detected in the N-back task but in the opposite direction, with higher random error for lower SCL measurements. Critically, in most tasks the computed LOAs were wider than the range of measurement, suggesting that SensCon-implied random error in SCL measurement might be too large, at least for extreme values.

A similar scenario was found for nsSCRs, showing uniform and nonsignificant bias only in the N-back and the Stroop task, whereas negative proportional biases were found in all the remaining tasks. Specifically, in the Seating, Standing, Walking, and Selection tasks SensCon showed the highest accuracy for lower rates of nsSCRs, that is when participants showed no or only a few responses, whereas it tended to underestimate larger nsSCRs measurements compared to the medical-grade device, possibly indicating false negatives. Higher nsSCR rates were also associated with wider LOAs (i.e., positive heteroscedasticity) for the Standing, N-back, and Stroop tasks, with the latter requiring logarithmic transformation to improve data normality [173]. Again, LOAs were found to overcome the range of measurement for certain values, particularly for higher nsSCR rates.

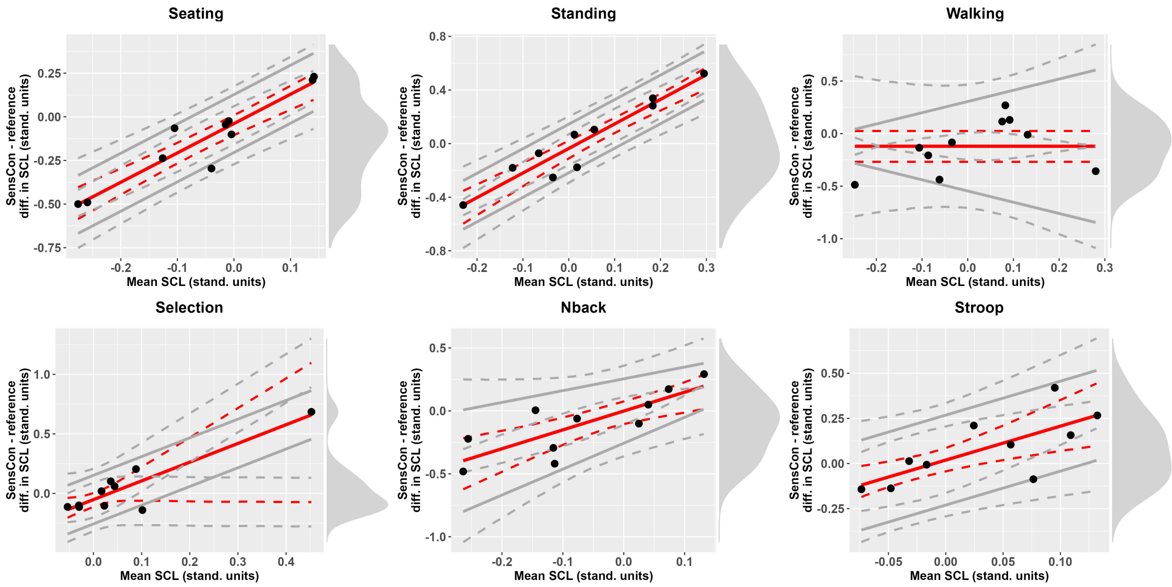


Figure 5.10: Bland-Altman plots for standardized SCL during Seating, Standing, Walking, Selection, N-Back and Stroop task Solid red lines show the bias (i.e., mean difference) and the 95% limits of agreement (LOAs). Gray dashed lines show the associated 95% confidence intervals.

Photoplethysmography A better agreement was found between SensCon-based HR measurements and those obtained with the corresponding medical-grade device. As shown in Table X and Figure 5.12, uniform and not significant biases were found in all tasks, with negligible systematic differences between systems ranging from 0.2 to 7.56 bpm. However, the random component of measurement error was relatively high (i.e., around $\pm$20-to-30 bpm). LOAs were uniform in most tasks, whereas negative heteroscedasticity was found in the N-back task, showing wider LOAs for lower HR measurements. Slighter heteroscedasticity also characterized the Selection task, following the logarithmic transformation of the data to achieve normality [173].

Questionnaires Finally, we tested the differences between SensCon and medical-grade devices considering the three self-report measures. Each measure was analyzed using t-tests or Wilcoxon-test upon normality testing. An overview of the results can be found in Table 5.3.

System Usability Scale We aggregated the ten SUS items as recommended by previous research [344]. We conducted a Wilcoxon test as the Shapiro-Wilk test of normality showed a significant result ($W = 0.84$, $p < .001$), resulting in significantly higher usability for SensCon compared to the medical-grade devices ($W = .0$, $p < .001$). Figure 5.14a shows the mean score for both conditions.

User Acceptance We analyze the eight UTAUT items by calculating the sum of the scales for each item [566]. As shown in Table 5.3 and Figure 5.13, we found significant differences

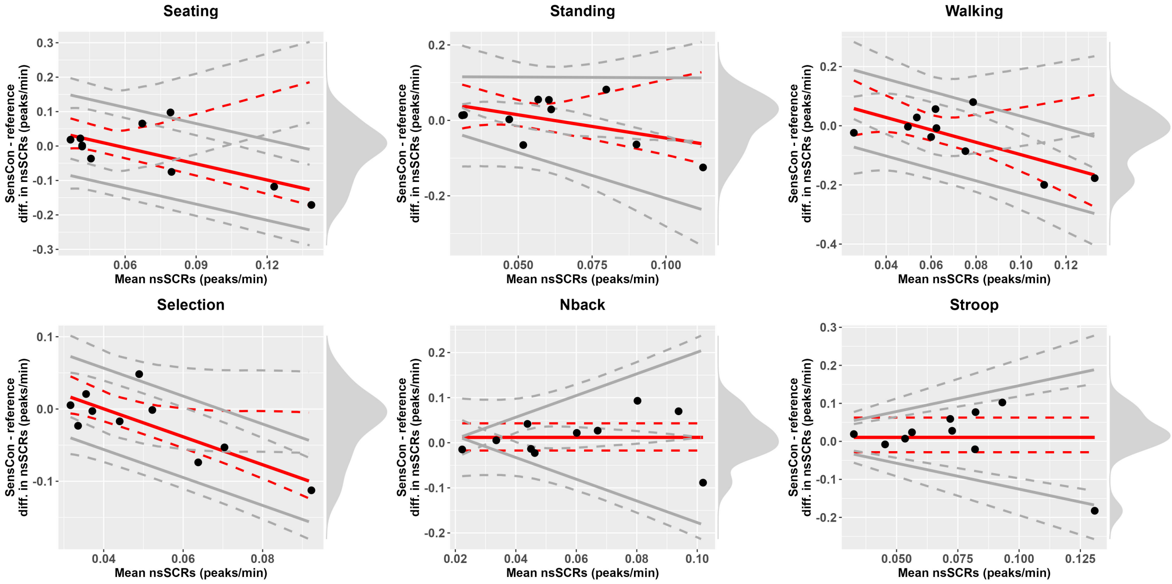


Figure 5.11: Bland-Altman plots for average nsSCRs peak frequency during Seating, Standing, Walking, Selection, N-Back and Stroop task Solid red lines show the bias (i.e., mean difference) and the 95% limits of agreement (LOAs). Gray dashed lines show the associated 95% confidence intervals.

in *Performance Expectancy*, *Effort Expectancy*, *Attitude Toward Using Technology*, *Behavioral Intention*, and *Anxiety*, but not in *Facilitating Conditions* and *Efficacy*.

User Experience We analyzed the seven-point Likert scales by comparing each of them between the two systems. As shown in Table 5.3 and Figure 5.14b, we found a significant difference between all Likert scales.

Discussion

Our results show that SensCon is an affordable alternative to medical-grade devices when it comes to measuring EDA and HR without compromising the user experience. Furthermore, SensCon empowers researchers and developers to obtain in-situ measurements of usability to assess the user experience in real-time or provide adaptive environments. Embedding physiological sensing for arosual detection into VR systems showed to allow for improved system UX and signal quality that did not show statistical differences for EDA and PPG extracted measures. We discuss the implications of our results in the following.

Usability-Friendly Physiological Sensing Our results show that SensCon, as a physiological sensing system, is preferred from a usability perspective. Furthermore, probing the participants with different questionnaires revealed a preference to use SensCon over traditional medical-grade equipment from different aspects (i.e., SUS, Performance Expectancy, Effort Expectancy, Attitude Toward Using Technology, Self-Efficacy, Anxiety, and Likert scales concerning user experience). Hence, we are confirming **H1**.

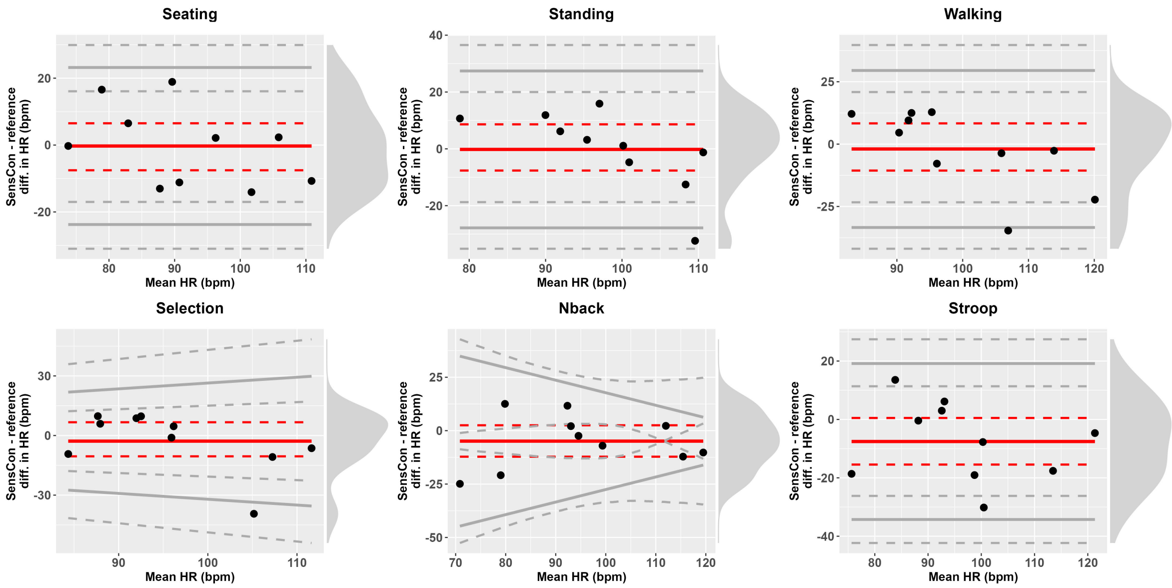


Figure 5.12: Bland-Altman plots for mean heart rate during Seating, Standing, Walking, Selection, N-Back, and Stroop task Solid red lines show the bias (i.e., mean difference) and the 95% limits of agreement (LOAs). Gray dashed lines show the associated 95% confidence intervals.

Our findings indicate that integrated physiological sensing leads to higher user acceptance. Similar to devices known from the wearable computing area, the user acceptance of novel sensing modalities increases when being integrated into everyday objects, hence increasing the quality of the overall interaction [293]. However, dealing with additional hardware to enable physiological interaction is often considered a burden by the user unless the obtained benefits overweight the workload of setting up a physiologically interactive system. SensCon overcomes these disadvantages by integrating physiological sensing into common VR controllers, leading to improved results regarding usability and user experience. However, it is important to acknowledge that medical-grade devices were designed for various purposes other than interactive VR systems only. Our results show that the adoption of physiological-interactive systems can be improved through direct integration into hardware. Consequently, we envision further research integrating alternative sensing modalities into VR controllers.

Signal Quality between SensCon and Medical-Grade Devices The comparison of the aggregated EDA and PPG measurements between SensCon and the medical-grade equipment showed converging trends. In both cases, the mean differences evaluated via frequentist and Bayes showed non signifcant results and moderate to anectodal evidence for H0, i.e., claim that no difference exists between SensCon and Medical-Grade devices data distributions.

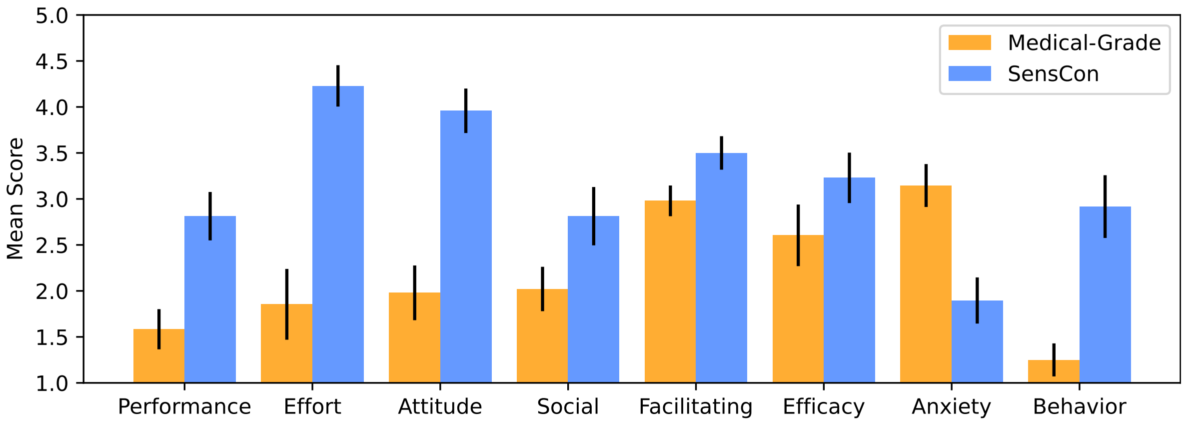


Figure 5.13: The Acceptance (UTAUT) for SensCon was better evaluated across Performance, Effort, Attitude, Behavioral Intention, and Anxiety subscales than medical-grade devices.

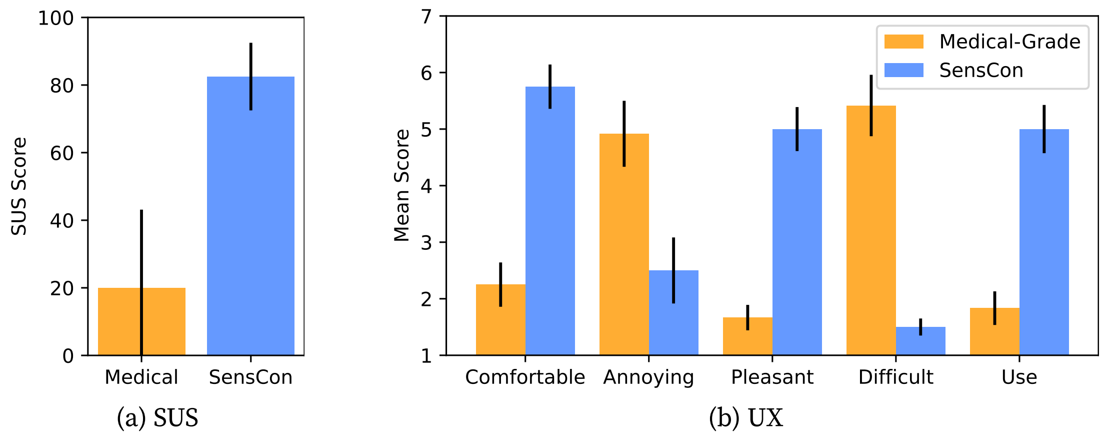


Figure 5.14: Results for SUS questionnaire (a) and UX survey scoring for SensCon and the medical-grade devices. Results from UX survey (B) are aligned with the ones from SUS, showing that SensCon was better evaluated in terms of Comfort, Annoyance, Pleasantness, Difficulty to Use, and Intention to Use. Here, error bars denote the standard error.

Specifically, this applied for all the extracted measures of physiological arousal, i.e., SCL, nsSCR rate, and HR. This result is interesting cause shows that SensCon detected slow, such as SCL and HR, and fast changes in the arousal state of the user with similar data distributions as medical-grade devices.

However, a deeper look at the agreement between SensCon and medical-grade devices revealed that the measurement error implied by the former might be excessively high in some cases. Specifically, we found a tendency of SensCon to systematically underestimate nsSCR rate as the number of nsSCR peaks increased. Whereas this may indicate a low SensCon sensitivity to sudden EDA spikes, the analysis of SCL suggested that both extremely low and extremely high SCL measurements are systematically biased. Accounting for these specific cases that require further investigation with larger samples, our findings indicate a promising overall accuracy of SensCon in measuring tonic and phasic EDA under most conditions. How-

Table 5.3: Statistical results of the SUS, UTAUT, and Likert questionnaires.

	SensCon		Medical		Normality		Test	
	M	**SD**	**M**	**SD**	**W**	**p**	**t/W**	**p**
SUS	82.5	9.6	20.0	22.2	.843	.002	0.0	**<.001**
Performance Expectancy	2.8	0.9	1.6	0.7	.907	.03	0.0	**.005**
Effort Expectancy	4.2	0.7	1.9	1.3	.843	.002	1.0	**.004**
Usage Behavior	4.0	0.8	2.0	1.0	.929	.092	−4.577	**<.001**
Social Influence	2.8	1.1	2.0	0.8	.890	.013	0.0	**.027**
Facilitating Conditions	3.5	0.6	3.0	0.6	.933	.116	−2.092	.06
Efficacy	3.2	0.9	2.6	1.1	.888	.012	12.5	.125
Anxiety	1.9	0.8	3.1	0.8	.926	.08	4.146	**.002**
Behavioral Intention	2.9	1.1	1.2	0.6	.808	<.001	0.0	**.005**
Q1: Comfortable	5.8	1.3	2.2	1.3	.839	.001	0.0	**.004**
Q2: Annoying	2.5	1.9	4.9	1.9	.810	<.001	5.0	**.012**
Q3: Pleasant	5.0	1.3	1.7	0.7	.901	.023	0.0	**.003**
Q4: Difficult	1.5	0.5	5.4	1.8	.812	<.001	0.0	**.003**
Q5: Like to Use	5.0	1.4	1.8	1.0	.898	.020	2.5	**<.001**

ever, future developments are needed to reduce the random error components of SensCon measurements. Indeed, whereas systematic biases were satisfactorily low in most cases (e.g., no systematic under- or overestimation was found for HR measurements), LOAs were relatively wide, especially for EDA measurements.

Although SensCon does not provide the same signal quality compared to medical-grade devices, it provides an alternative when usability and user experience are the focus of a physiological interactive system. We postulate a trade-off between sensing accuracy and the provided user experience. Reflecting on the not significant mean differences between the SensCon and medical EDA measurements, as well as the similarity between the SensCon and medical heart rate measurements, we can partially confirm **H2**.

SensCon provides an affordable and usable alternative for measuring physiological arousal at lower reliability compared to medical-grade devices. Although SensCon is providing less reliable data, future physiological interactive systems can observe the data over a period of time to selectively do an assessment of the perceived user experience and adaptations when a specific accuracy is reached. Past research recommends avoiding immediate adaptations to reduce the risk of damaging the user experience through maladaptations [551]. Instead, physiological data should be collected over a longer period to account for variations over time [292, 465]. SensCon can use this to its advantage by analyzing individualized physiological data over long periods, making predictions more robust for individual users. The technical contribution of SensCon will allow large-scale data collection to enhance long-term physiological interaction in virtual environments without compromising the user's experience, hence their willingness to provide physiological data in different environments. Furthermore, the current system implementation incorporates low-cost sensing and processing components widely available on the market. Although the quality of these components is lower than that of medical-grade devices, it is possible to improve measurement quality by using specialized sensors and a more advanced sensor conditioning strategy. A future system prototype could be developed using superficial soldering components and miniaturized accordingly, allowing for a "plug and play" upgrade from a conventional controller. Given that the system's sensors communicate wirelessly with the VR environment, one could argue that this method could be applied to any VR controller.

Alternative Hardware Configurations SensCon features two embedded circuits and sensors, each powered by external PowerBanks. However, we also considered other hardware configurations during the design process. Here, we discuss these alternatives and their respective benefits and drawbacks. A first option would be to integrate the battery that powers the circuits and sensors into the controller itself, increasing its robustness compared to its current configuration. In addition, the battery size restricts the maximum amount of energy supplied to the circuitry. Consequently, production versions of SensCon would require ultra-low energy consumption components. Currently, our implementation employs consumer circuits better suited for rapid prototyping rather than final production. Several software strategies for reducing power consumption can be implemented; for example, reducing the sampling rate based on the amount of activity and stopping sampling when the controller is not grasped. In its current configuration, the system continuously measures the physiological response of the participant. To allow for easier integration within the VR controller, we could reduce the component size of the circuits. With these improvements, we could explore several alternative configurations.

Embedded Circuits, External Sensors: One alternative configuration is integrating the main processing and communication circuits within the controller and attaching the sensors externally to the participant's palm or wrist. This approach enables direct measurements from the wrist or fingertips while reducing movement artifacts, but requires additional setup time that may negatively impact the user experience.

External Circuits, Surface Sensors: This alternative configuration relocates the processing and communication circuits outside the VR controller and adds superficial sensors to the grip locations, most likely in the exact locations as the current SensCon configuration. Achieving this configuration may entail designing all circuitry detachable and attachable to the VR controller. This approach enhances system versatility and flexibility for different VR controllers, but may also affect usability, grip, and robustness, depending on the attachment's ergonomics.

External Circuits, External Sensors: Similarly to the previous configuration, this alternative arrangement would position the circuitry externally to the primary body of the controller to enable it to be attached and detached. This configuration would let the sensors be positioned directly on the pulse or fingertips, as with the first configuration. This would bring the advantage of being reusable in several controllers. Still, the drawback is that it would be less robust and likely impact the usability given that the user would have to (1) set up the controller attachment and (2) setup the sensors in their hands/wrist.

Limitations and Future Work The reported results in this study are prone to certain limitations. Here, we identify limitations and highlight space for future work and improvement for Senscon.

Evaluating Test-Retest reliability. First, our study design utilized single trials, meaning that each condition was conducted once per participant. More robust signal quality can be achieved with multiple trials. However, our conditions lasted for three minutes each, providing enough data to provide a reliable statement about the signal quality. In future work, we will first investigate how the accuracy of SensCon varies through test-retest validation.

Effect of Controller Grip and Sensor Location. Our study investigated the effects of sensor location on the quality of physiological signals measured during VR experiences. Specifically, we found that integrating the sensors into the VR controller provided a reliable and convenient location for measuring physiological signals, mainly when using a power grip. The power grip was the most immediate solution for our participant when asked to handle the VR controller (HTC Vive). We did not instruct participants to choose a power grip but rather that the power grip was the natural way for them to handle the VR controller, given its inherent design and affordances. While the current study focused on the power grip, future research should consider investigating other types of grips to determine how physiological signal quality may be impacted by different grip styles. Next, we will investigate how SensCon current sensor locations will perform in terms of signal quality and agreement with sensors embedded in the VR headset. Here, motion artefacts might be less observable than in finger locations. Conversely, we have to consider that acquiring an EDA signal from the head might be more influenced by thermoregulatory activity in the head rather than variation in arousal's state [431]. Similarly, PPG signal acquired from the forehead might slightly differ due to the increased vascularization and decreased vasoconstriction of scalp and brain [613].

Demographic limitations of the sample size. We need to consider some limitations in our sample related to the age range of the recruited sample. Age is a significant confounding factor in peripheral physiological signals [198]. For example, when considering the pulse transit time, which is influenced by atherosclerosis, i.e., the thickening and stiffness of blood vessel walls, which are increasing with age and can impact the accuracy of PPG measurements [446]. Furthermore, age differences can also impact physiological reactivity to a variety of arousing stimuli, e.g., affective images [154, 198] We will extend our validation to a more diverse sample considering a larger demographic diversity.

Online EDA and PPG data preprocessing. In our current implementation, we extracted physiological measures offline. To explore the feasibility and potential benefits of real-time physiological monitoring in VR adaptive settings, future work could investigate the possibility of implementing online analysis in Micropython via an on-board microcontroller or specialized ultra-low-power biometric hubs. This approach would allow real-time computation of peripheral physiological measures from the raw signal, enabling SensCon to provide immediate feedback and support real-time physiological monitoring in interactive environments. This approach would require additional hardware development, optimization, and potential trade-offs between computational complexity and power consumption. Therefore, future work should focus on exploring and evaluating the feasibility of this approach.

Generalization to real-world tasks. We chose to validate SensCon over a series of tasks representative of specific movements and cognitive processes, based on previous work [387]. These tasks were designed to be relevant to a broader community beyond just HCI and VR developers. However, further validation would benefit from a more ecological context, such as VR fitness or gaming apps, which may provide a more complex and dynamic environment to test the real-world applicability of SensCon. This would allow us to understand better how SensCon performs in scenarios that involve more erratic and everchanging movements with varied cognitive load [318] Ultimately, we aim to design SensCon for a wide range of applications. We believe that a diverse set of validation tasks, including those focused on joint movements and various cognitive processes, is necessary to ensure that SensCon can perform well across different contexts.

Hyperscanning validation. In future work, we aim to establish embedded physiological sensing systems in hyper-scanning VR settings. Here, we will integrate SensCon into multi-user VR interactions to investigate collaborative and social VR scenarios.

Applications

EDA and PPG data enable us to infer several user states, which are beneficial for interactive systems. Finally, we present use cases that benefit from SensCon as a real-time sensing system.

Sensing User Experience for Interface Assessments and Adaptations Previous work showed that usability metrics, such as engagement [147, 240], stress [512], and workload [103, 115], are closely linked to physiological signals. Thus, real-time measurements of PPG and EDA can quantify the user experience in real-time. Virtual environments benefit from these insights as researchers and developers acquire implicit real-time feedback about their environments. SensCon integrates EDA and PPG as measures for supporting user experience designers to detect usability issues to improve them selectively. Furthermore, adaptive interfaces benefit from these usability measures by adapting the interface to either increase the user's engagement or reduce stress and workload. Physiologically aware VR environments will provide more individualized environments for the user, improving the user experience. Potential use cases include learning scenarios [426] that adapt educational material or working scenarios that provide adaptive in-situ assistance [316].

Fitness and Activity Recognition The EDA and PPG values correspond with the physical activity of the user. Early research pointed out that increased heart rate is associated with physical activity[536]. Therefore, deriving the heart rate from SensCon enables users to estimate the activity of users in virtual environments seamlessly [563]. This information can be processed to reflect users' activity during fitness applications, i.e., balance training [148] or exergaming [403]. Additionally, virtual environments can adapt their exercises to match the physical fitness level of the user, hence optimizing the training output. Furthermore, we envision SensCon as an activity recognition sensor for virtual environments. Here, SensCon can serve as an indicator for the user's current activity, including seating, standing, or walking.

Large-Scale User Experience Assessment Current user experience assessments are susceptible to subjective biases, lack objective real-time feedback, and require the presence of an experimenter. EDA and PPG measures can obtain objective real-time feedback for user experience assessments. However, using physiological sensing, past methods to empirically evaluate the user experience were hindered by either using a complex setup requiring prior training or large-scale data collection. SensCon overcomes these issues by facilitating the setup and data collection process by using the VR controllers only. We envision SensCon as a user-friendly tool to collect large-scale physiological data for robust user experience assessments in virtual environments.

5.2 Summary

In this chapter, we introduced SensCon, a system that directly integrates physiological sensors for EDA and PPG into VR controllers. SensCon provides a practical alternative to medical-grade equipment, simplifying the collection of stress, workload, and engagement data in VR environments. We addressed (**RQ9**), showing how SensCon balances user convenience against the precision of physiological measurements. Although participants preferred SensCon's ease of use over traditional, cumbersome setups, its sensing accuracy and temporal resolution are somewhat compromised. This trade-off highlights the challenges in ensuring signal reliability and validity within embedded systems. SensCon's integration into VR controllers makes physiological monitoring more accessible for VR developers and researchers, offering a valuable tool for understanding user interactions in real time.

6 DISCUSSION & FUTURE WORK

> *"Those are my principles, and if you don't like them . . . well, I have others."*
>
> **– Groucho Marx.**

This chapter synthesizes and discusses the findings presented throughout this thesis. We begin with a brief overview of the work conducted (section 6.1), followed by a detailed discussion of how these findings contribute to the research questions posed in section 1.1. We then discuss the limitation of this work and to what future work they point towards.

6.1 Summary of Contributions

In the following sections, we outline our contributions to the research questions posed in this thesis. Initially, we will summarize the principal contributions to each question. Subsequently, we will offer detailed responses that comprehensively address these research questions. We investigated attention and engagement in settings that would have allowed us to investigate both task-relevant and task-irrelevant elements of the task or the environment. This allowed us to investigate the blend of virtual and physical information and their effect on user performance, task engagement, and attention allocation.

6.1.1 RQ1: Attention Allocation and Task Engagement in Mixed Reality Continuum

As a first step to answer this research question, in section 3.1, we chose a visual search task as transferable to different ecological settings situated in MR, such as cognitive training [145], information retrieval [55] and practical activities like cooking [131]. Our study first filled a reserach gap as to date, it has primarily focused on visual search performance in VR only, examining factors like the distracting effects of perceptual load [419] and the influence of how information can be displayed, e.g., field-of-view and eccentricity [228]. In AR, studies have generally concentrated on visual guidance [58], utilizing techniques such as attention funnels or layered interfaces that adhere to the principles of attention theory [58, 176]. At the same time, AV in visual search remains relatively limited [529].

Here, we systematically investigated how users detect target virtual information across the reality-virtuality continuum, focusing on the effects of the general MR environment. We applied a multimodal evaluation using behavioral, subjective, and physiological (EEG and eye tracking) measures underlying workload, resource allocation, and visual search efficiency.

Our results confirm previous findings and extend them toward the MR continuum, showing how, regardless of the perceptual load of the scene, AR environments posed increased attentional demands on users.

Following up on this, in section 3.2, we explored how attention is allocated to target and distractor objects, including virtual and physical elements. Prior research indicates that managing virtual and physical cues simultaneously presents considerable challenges, highlighting the need to carefully consider these perceptual differences in MR design to enhance usability [303, 416]. In MR environments, the difficulty of visual search tasks can vary significantly based on the representation of these elements. Currently, users can distinguish between physical and virtual objects due to distinct differences in fidelity. However, as MR technology progresses, it is expected to increasingly blend real and virtual elements [21]. Such advancements will likely complicate visual search tasks by blurring the lines between real and virtual content, thus increasing the cognitive load on users' visual processing [273]. Thus, we investigated how different MR actualities and target nature impact task efficiency and attention allocation. Results showed that target processing was more efficient in AV than in AR. This was supported by measurements that track brain responses to distractions, showing a smoother and quicker recognition of important objects in AV. Eye-tracking data further confirmed that in AR, participants found the visual search tasks more challenging and disorganized, evident from their more erratic eye movements and the measurements of eye dilation, which indicate mental effort. These insights suggest that the design of MR environments should consider how virtual and physical elements are presented to minimize cognitive strain and enhance user interaction.

Together, these studies highlight a key design principle for MR: optimizing virtual element presentation in AV can mitigate the information processing challenges that a predominant physical environment poses, such as in AR. This suggests a strategic approach to MR design that utilizes AV benefits to enhance task efficiency and user experience. This synthesis points towards designing MR systems that are cognitively considerate, blending virtual content in ways that support user performance across various applications.

In the third contribution to this research question, we wanted to verify if insights from our initial studies could be applied to a practical and commonly encountered task in MR environments: typing on a keyboard. This scenario is particularly relevant as keyboards remain the primary tool for efficient text input across various devices, yet transitioning this efficiency to MR environments poses significant challenges. Achieving effective text entry in MR allow for the development of productivity tools and immersive metaverse experiences [212]. Second, our previous studies on visual search in MR were limited to relatively simplistic representations, predominantly involving virtual stimuli in controlled settings. The third work embeds our research in realistic simulations by incorporating both virtual and physical visual clutter, which would better reflect the complexity of real-world MR environments and burdens on information processing [110, 380].

Here, we further investigated physiological responses to typing tasks within MR environments to understand how different modalities influence user engagement, attention allocation, and experience. This contribution to the physiological underpinnings aims to inform the design of adaptive MR systems that respond dynamically to the user's engagement state, potentially enhancing productivity and user experience. The results revealed that AV facilitated more efficient task processing than AR, indicating more accurate typing, and increased physiological engagement and internal attention state. This suggests that integrating and balancing physical and virtual elements in MR environments are crucial in minimizing cognitive strain and optimizing user interaction.

The integrated findings from our visual search and typing studies offer significant insights into the dynamics of attention allocation and task engagement across the MR continuum. These contributions collectively demonstrate that AV is particularly effective in enhancing user performance. However, AR can elevate cognitive demands due to its complex backgrounds and distractions, diminishing performance either by searching for target information or in productivity tasks. In contrast, AV and VR do not impact performance by simplifying visual elements and minimizing real-world interference. However, VR faces consistent challenges, such as the lack of physical surfaces and haptic feedback, worsening performance and fostering task disengagement.

Our studies demonstrate that attention fluctuates significantly across the MRcontinuum, closely reflecting variations in external stimuli. This variability underscores the need for adaptive MR systems that adjust to the ongoing interplay between virtual elements and physical reality. Our results indicate that designing MR interactions requires carefully considering how virtual and physical elements are integrated. Additionally, we identified specific physiological patterns linked to the processing of task-relevant information in MR settings. These findings suggest potential applications for physiological computing in developing MR systems that can monitor user engagement and attention.

Our work promotes integrating physiological computing [178] with adaptive MR systems [352] to enhance user experiences. By adding attention and engagement awareness to MR systems, they can dynamically modify environmental virtuality and object interaction to promote optimal information processing.

With the next series of studies, we explored the design and input space for physiologically adaptive systems in MR.

6.1.2 RQ2 : Designing Adaptive Systems for Supporting Engagement and Attention Allocation

The series of five studies conducted as part of this thesis aimed to explore the development and implications of adaptive VR systems sensitive to user attention and engagement. Each study built upon the findings of its predecessors, progressively refining the understanding and implementation of physiological measures to enhance user interaction within MR environments.

Specifically, our research question was grounded on three assumptions.

First, as a task setting, we chose an N-Back task involving both external and internal attention states to examine how attentional fluctuations change between task-relevant and irrelevant stimuli within varying levels of virtual environments. This task was particularly suited for these studies due to its flexibility in manipulating visual information, which is a primary mode of content delivery across all actualities of the MR continuum [21].

Second, we focused on a specific actuality of the continuum, i.e., VR. VR provided a controlled setting where task-related factors and environmental variables could be manipulated precisely, making it an ideal platform for testing adaptive systems' efficacy to enhance user performance. Together with designing adaptive systems, we acquired multiple physiological measures to investigate the effect of adaptations on physiological attention allocation and engagement.

Third, we focused on visual complexity due to its strong influence on information processing and task performance in VR environments. Visual realism is essential in VR systems as it enhances the realism and real-time updates of visual stimuli, crucial for maintaining user engagement and improving the VR experience [87, 211]. However, increased realism and detail also heighten the visual complexity, which can overload visual working memory and distract from the primary task, thereby impacting task performance negatively [161, 418]. Given the limited capacity of our perceptual and attentional systems, overly complex scenes can lead to cognitive overload, discomfort, and decreased task engagement [7, 194, 301, 368, 383, 441, 462]. Conversely, low-fidelity environments risk under-stimulation, leading to disengagement and reduced alertness. Therefore, managing visual complexity in VR is pivotal to optimizing user interaction by balancing cognitive demands with engagement levels, guided by physiological measures of arousal and cognitive load to dynamically adapt to the environment [179, 374]. By adjusting visual complexity, we can impact changes in cognitive load, attentional focus, and overall engagement. This is crucial for designing MR systems that maximize user performance and satisfaction by adapting to their cognitive state and needs. In doing so, we seek to create adaptive systems that respond to user inputs and proactively adjust to facilitate optimal interaction dynamics within any VR settings.

The first study, see section 4.1, introduced an adaptive system that modified task-relevant visual complexity based on user engagement, measured through raw EDA signal. This foundational experiment established the viability of using physiological responses to adjust task-relevant features dynamically, setting the stage for more complex integrations of user feedback into system adaptations. Secondly, we demonstrated that dynamically adapting task-relevant visual complexity based on physiological feedback can reduce perceived cognitive overload and enhance perceived engagement.

Building on the initial findings, we then analyzed a comprehensive dataset from the first experiment to explore the patterns of physiological responses across different conditions of task-relevant visual complexity and system adaptation, see section 4.2. This analysis helped refine the adaptive algorithms by identifying which physiological markers, such as tonic EDA and EEG theta oscillations, were most indicative of user states of engagement and cognitive load.

In the third study (section 4.3), we explored the feasibility of dynamically adjusting the environmental visual complexity in VR settings to enhance cognitive performance on primary tasks. Crucially, this adaptation focused solely on modifying surrounding, non-task-relevant visual elements rather than the task content itself. This contribution is relevant as addresses the cognitive load without altering the fundamental nature of the task, providing insights into how peripheral changes in a VR environment can influence central cognitive processes. The approach improved task performance and introduced a rigorous methodological framework for assessing the effectiveness of adaptive systems, incorporating a control condition within the adaptive system.

Next, we again evaluated a multimodal dataset of EEG, EDA, and ECG data associated with the adaptive manipulation of visual complexity in section 4.4. This work explored how adaptive systems could employ EEG signals as input, specifically alpha and theta oscillations, to inform system adjustments based on participants' engagement and attention levels. Increased visual complexity was associated with higher alpha power, suggesting enhanced mental fatigue or increased cognitive workload. The alpha-theta ratio also varied with complexity levels, showing the effect of visual complexity on workload. The findings from from this study, particularly regarding EEG measures, set the stage for the last contribution to this research question, demonstrating the potential of EEG in detecting shifts in attentional and engagement states due to changes in visual complexity.

This allowed to refine the adaptive mechanisms further and employ EEG correlates of attention allocation and engagement, see section 4.5. Here, we designed a VR adaptive system based on frontal theta and parietal alpha frequency bands to adapt the environmental visual complexity surrounding the main task based on real-time assessments of the user's attention state. Moreover, given our labelled EEG dataset with associated attentional states, we trained a LDA [573, 584] that classified participants' internal and external attention states with an accuracy of 86.49%. This not only enables the translation of the adaptive system to various use cases but also informs into the features that contribute most significantly to model accuracy, offering a pathway to improving the explainability of attention-based adaptations in VR environments.

In the following section, we will explore potential applications for these engagement and attention-aware adaptive systems across various MR settings, aiming to enhance task performance and UX in immersive environments.

MR Applications for Engagement-Aware Adaptive Systems

Our work integrates a physiologically adaptive system that functions independently of the simulation software, applying across various VR, AR, and MR environments. This flexibility ensures that our adaptive strategies, grounded in peripheral physiological data, can enhance user experience across diverse platforms.

We have demonstrated the potential benefits of adapting visual complexity in virtual environments to optimize performance and reduce perceived workload. These adaptations can be crucial in VR and blended MR settings, where adjusting digital content complexity can significantly enhance productivity. The advent of wireless HMDs expands the possibilities for virtual workspaces, where our findings can contribute to more productive and less cognitively taxing environments [100, 317, 380].

Furthermore, physiological measures of arousal, often used in therapeutic and rehabilitation contexts, can guide the adaptation of content in VR therapies. For example, by modulating exposure therapy stimuli based on arousal responses, we can finely tune therapeutic interventions to the patient's progress, enhancing the effectiveness of treatments in real-world scenarios [118, 399, 594].

Another area impacted by our research is the management of cybersickness, a common challenge in immersive environments caused by sensory mismatches. By detecting arousal that may indicate the onset of cybersickness, our system can proactively adjust visual inputs to mitigate these symptoms, improving user comfort and extending session durations [199, 226, 534].

In VR navigation and wayfinding, our system could use physiological arousal as a proxy for cognitive load to dynamically simplify complex visuals when users are overwhelmed, thereby aiding in more intuitive navigation. This can reduce the cognitive burden during navigation tasks, making virtual environments more accessible and easier to navigate [110, 340].

Finally, our approach can enhance interactions in social VR settings by adjusting the proxemics and visual density of NPCs based on detected arousal levels. This can prevent discomfort and enhance the social dynamics within virtual spaces, making these interactions more natural and comfortable for users [261, 353, 483].

Applications for Attention-Aware MR Adaptive Systems

Our research underlines the importance of optimizing attentional resources in VR environments to enhance user experience and task performance, especially in scenarios that necessitate a balance between internal and external attention processing.

Supporting Internal Attention Our findings indicate that supporting internal attention can significantly improve task performance in environments that minimize external distractions. For instance, in VR productivity settings like virtual offices [313], users, particularly novices, are often overwhelmed by the multiplicity of visual stimuli, such as virtual colleagues or avatars. Here, an adaptive system that minimizes distractions while enhancing task-related cues can significantly boost efficiency and focus, crucial for working memory tasks.

This approach also has implications for cognitive training programs aimed at both healthy and clinical populations, including those with attention-deficit/hyperactivity disorder (ADHD)

[125, 295]. By tailoring MR environments to foster internal attention, cognitive training tasks can improve cognitive functions such as problem-solving, decision-making, and learning.

Supporting External Attention Conversely, supporting for external attention is equally vital, particularly in VR applications where the visual information is integral to the task but competes with internal distractions, such as mind-wandering or rumination [116, 222]. In scenarios such as VR content creation, motor learning, or MR visual analytics [302], where users interact with complex, detailed, and dynamic content, systems can adapt by enhancing perceptual salience or temporarily pausing interactions to recalibrate user focus.

Implementing adaptive mechanisms that optimize external attention can be particularly beneficial in educational and training contexts, aligning with the optimal theory of learning [608]. This theory suggests that an external focus during learning tasks can significantly enhance skill acquisition compared to an internal focus. Thus, VR systems designed to emphasize external cues can facilitate more effective learning and user engagement.

6.1.3 RQ3 : Towards Embedded Physiological Computing Systems in MR

The research presented in this thesis has significantly advanced the understanding of user experience, engagement, and attentional states in AR, AV and VRfrom a physiological perspective. We demonstrated that physiological data serve not only to inform on implicit user behavior but can also act as a direct input for interactive systems.

Despite these advancements, the development of deployable, intelligent systems must address several design and usability challenges. While providing high-quality signals, current research-grade sensing devices often involve uncomfortable setups for users, which can impede research efforts [278, 581]. Moreover, when users are immersed in VR or AV environments, their spatial awareness is often compromised. This lack of awareness, combined with the physical encumbrances of physiological sensing systems such as cables or VR tracking equipment, can lead to signal artifacts or loss, thus disrupting engagement within the virtual environment.

The application of physiological sensing in VR also requires a multidisciplinary approach encompassing both psychophysiological principles and the technical integration of physiological signals into VR systems. This complexity currently restricts its application primarily to the HCI research field. To address these challenges, researchers have designed and prototyped physiological sensors embedded directly into VR systems [48, 612]. For example, Luong et al. [360] demonstrated the real-time estimation of mental workload in a VR flight simulation using multiple physiological inputs integrated into the head-mounted display (HMD). However, the validation of these systems is often restricted to consumer-grade devices [387] or non-VR environments [568], which limits their scientific utility due to the need for high-quality, integral data.

Moreover, the commercial-grade quality of these devices frequently precludes access to raw data, relying instead on proprietary metrics that lack transparency. In response to these limitations, our third research question led to the design and implementation of SensCon, a physiological sensing system embedded within VR controllers. This integration offers several benefits: it combines hardware and software into a single device, enhancing user comfort, reducing costs, and potentially improving user acceptance [209].

In our research, we first studied to identify the optimal placement for EDA and PPG sensors within a VR controller. Subsequently, we evaluated SensCon's performance across various VR scenarios commonly used in stress and psychophysiology research. Our findings indicate a strong preference for SensCon due to its usability and demonstrate its promising accuracy in measuring tonic and phasic EDA and HR under various VR tasks. SensCon has a promising overall accuracy and offers a viable alternative when prioritizing usability and user experience in physiological interactions.

To structure the contribution to our RQ3 and implications for future work, we draw upon the seven types of HCI implications as derived by van Berkel and Hornbæk [561], encompassing implications for methodology, theory, design and policy.

Implications for Methodology: The Need for Method Validation and Open Research Practices

Methods With intelligent MR systems in mind, validating sensing systems against gold standards is essential for ensuring the accuracy and reliability of physiological data. Wearable sensors, resembling everyday items like wristbands and headsets, integrate seamlessly into MR environments. These wearables enhance user comfort and reduce the obtrusiveness often associated with traditional biomedical systems like ECG, mitigating effects such as the "white coat effect" where patients exhibit elevated blood pressure when exposed to clinical hardware or settings due to anxiety [615].

However, current devices used in MR systems frequently lack thorough validation against established biomedical standards, which is critical for confirming their accuracy in natural settings. For instance, the E4 system monitors various physiological signals and is designed to generalize laboratory findings to real-world applications. Ensuring ecological validity—how well data represents real-life situations—is crucial, particularly as MR applications often operate under dynamic conditions involving motion and environmental changes [379, 442].

Furthermore, MR systems often encounter user behavior that is dynamic and divergent from traditional cognitive science or experimental psychology settings. This variability underscores the need for robust data integrity to understand how recorded physiological signals behave in new and innovative user interaction paradigms. Therefore, rigorous validation studies are necessary to address potential measurement errors and establish the reliability of these emerging technologies in MR environments, where new user behaviors and interaction paradigms are evaluated.

Open Research Practices Traditional HCI methods have often fallen short in providing a replicable and transparent research environment. For example, Wacharamanotham et al. [578] emphasize that while HCI research is prolific in producing novel user interfaces and interaction techniques, it frequently does not prioritize making these innovations accessible or reproducible by the wider community. This lack of openness can stymie innovation, as researchers are unable to build directly on the work of others. They criticize the prevalent culture in HCI of closed, proprietary systems that restrict the dissemination of methodologies and findings, arguing that such an approach limits the field's ability to effectively address complex, real-world problems.

Similarly, Kuuti & Bannon [328] reflect on the need for HCI to embrace more open and collaborative research practices to keep pace with the evolving technological landscapes and user interactions. The practice turn in HCI, as Kuuti discusses, suggests a shift towards methodologies that integrate more holistic and contextually grounded approaches, yet these are not often shared or made standard in a way that could benefit the broader field.

Openness in research tools, methods, and data is essential to foster a more collaborative and innovative research environment. In this thesis there is a commitment to open science practices. All contributions, including the developed VR and MR environments, are accompanied by open data and preprocessing scripts. This approach not only enhances the reproducibility of the research but also encourages a more collaborative and progressive scientific community. This initiative aligns with contemporary calls within HCI for more transparent and accessible research outputs.

Implications for Theory: Learning from Large Users Assessment in Everyday Settings

As physiological sensing systems become more accessible and cost-effective, as shown by SenCon and other research efforts [47, 360], we anticipate their integration into mainstream HCI research, enriching our understanding of cognitive behavior such as stress, attention, fatigue, and engagement. This widespread availability of embedded sensing systems enables the collection of long-term and diverse data sets without intensive supervision. Researchers may soon routinely request participants, who may already be using wearable devices like smartwatches, to gather physiological data across various real-life settings. Such in-the-wild methodologies will allow for a more comprehensive examination of physiological data often skewed by controlled laboratory conditions.

Implications for Design: Designing Adaptive MR Systems

Based on a larger application space for wearable physiological sensing, it then would be possible move to wearable "physiological computing". Here, future design of adaptive MR systems will increasingly leverage a deep modelling of physiological data in order to infer ongoing cognitive processes during user interaction. Wearable "physiological computing" systems are an evolution of context-aware computing, now incorporating mental information processing such as attention allocation, perception, memory encoding, and learning.

Here, Navarro [408] emphasizes the importance of interface naturalness and aesthetic integration for wearable devices in MR environments, particularly BCIs. The author suggests that by correlating EEG with other techniques like peripheral physiological signals and EMG, and through comprehensive data from MR environments, BCIs can be more accurately triggered in dynamic real-world settings. This could lead to more personalized and contextually aware applications, enhancing user experience by predicting and responding to user behaviors more effectively. Udovicic et al. [558] integrate into this perspective how wearable adaptive systems should be integrated within surrounding smart environments, with a special focus on BCIs. Integrating these devices within smart environments poses significant challenges, including connectivity, data input, security, and energy consumption. These aspects are critical in ensuring that MR systems are functional but also consistent and reliable in various real-world applications.

Implications for Policy: Addressing Privacy and Ethics

When considering systems that monitor, process and react based on data where users have limited control can cause natural concerns towards privacy and ethics on data usage. Here, users must be fully informed about how their physiological data are collected, used, and shared, especially when used for model training and validation. Concerns also extend to widespread monitoring's potential health and societal implications [22]. Additionally, the emotional implications of these systems need careful management to ensure users can revert to a neutral state if desired.

Privacy concerns are paramount, given the personal nature of physiological data collected by wearable technologies. To guarantee transparency and security, systems must incorporate privacy considerations throughout every stage of development. This includes performing privacy impact assessments and integrating privacy-enhancing technologies to protect user information. Moreover, users should retain legal data ownership, with third-party access only granted through explicit user consent. Last, a key aspect of maintaining user trust involves a transparent data management policy that clearly outlines how data is stored, accessed, and used [101, 557].

The widespread connectivity of wearable devices raises significant security concerns. These devices collect detailed personal information, such as activity patterns and sleep cycles, making them potential targets for cyberattacks. Such breaches could expose sensitive health details, leading to sophisticated phishing attacks [434], or compromise connected systems like payment platforms and smart homes, risking unauthorized device control [50].

6.2 Limitations & Future Work

Given the contributions presented in this thesis, various limitations on the approaches for adaptation and research applications persist, and new opportunities for future research emerge.

6.2.1 Exploring the Sensor Space

Our investigation revealed no significant impact of visual complexity in VR environments on ECG time domain measures (HR and HRV), indicating that ECG might not be sensitive to visual stimuli like other physiological measures. This finding contrasts with the responsiveness of ECG in typing tasks within VR, where clear physiological responses were observed. This divergence suggests that ECG's sensitivity may be more aligned with physical exertion or emotional arousal rather than mere visual complexity. Additionally, we explored EDA and EEG as inputs for making adaptive systems more aware of user attention and engagement.

To better investigate ECG as an active input in MR, exploring applications where physical exertion or emotional arousal is prominent could be beneficial. ECG variations are associated with stress, arousal, and physical exertion. Its application in VR can tailor physical activities to suit the user's current state, enhancing physical well-being and motivation [51, 219, 403]. Moreover, the relationship between ECG metrics and motivational states is shared with a theoretical model also implemented in our work, i.e. , Richter's motivation intensity model. This points towards adaptive MR or VR applications that consider users' arousal and exertion to support physical well-being and fitness by adjusting the intensity of challenges based on the perceived difficulty and effort's worthiness [38, 476, 559].

Here, adaptive training environments in VR could use ECG to adjust difficulty and intensity in real-time based on heart rate data. This would make training sessions more effective by maintaining optimal heart rate zones and increase user motivation and engagement by aligning challenges with individual capabilities.

Moreover, the representation of user data in VR can greatly influence user reflection and stress management. For example, a VR system developed by Wagener et al. uses weather metaphors to visually represent stress scores, encouraging users to reflect on their stress data in an engaging manner [252, 579].

6.2.2 Overcoming Rule-Based Approaches with Automated State Detection

In this work, we initially adopted rule-based approaches to adapt virtual environments based on physiological signals. While these methods demonstrated efficacy, they are inherently task-specific and lack generalizability. Consequently, there is a significant interest in moving towards automated state detection, which promises greater flexibility and broader applicability.

Rule-based systems, as deployed in our studies, provided structured adaptations based on predefined criteria. However, their performance is tightly coupled with the specific conditions under which the rules were formulated. This limitation restricts their utility across varied tasks and environments, as they do not dynamically adjust to new or unexpected user states or behaviors.

Automated state detection through machine learning offers a robust alternative, capable of interpreting complex and multimodal patterns in physiological signals in real-time. This approach is not only adaptable to diverse settings but also enhances the system's ability to learn from new data continuously. For instance, in section 4.5, we employed LDA to classify participants' attention states with high accuracy, illustrating the potential of machine learning to be embedded in intelligent VR and MR systems. Therefore future work should establish an optimal and reliable approach for real-time classification exploring diverse models. This would be informative on the distinct advantages of each model, depending on the data complexity and volume. For example, while LDA and Random Forest might perform better with smaller datasets, Neural Networks could potentially excel with larger, more complex datasets, especially when classifying raw EEG data. This comparative approach not only seeks the best-performing classifier but also aligns with findings suggesting that different types of neural networks might yield superior results for specific data modality, thus providing an informative overview on the algorithms that can enhancing adaptive VR and MR systems.

The integration of sensor fusion techniques further refines state detection. Combining multiple physiological signals can overcome the limitations where single-sensor data might be ambiguous or insufficient. This approach was explored in Vortmann et al. [573], where EEG data and eye tracking features significantly contributed to distinguishing between internal and external attention states in AR.

Expanding on this, Vortmann et al. suggests that leveraging multiple sensor inputs allows for a richer and more accurate representation of user states, which is crucial for developing responsive and effective adaptive systems. This enhances the specificity of the system's response to user needs and offers a pathway to address the challenge of sensor data ambiguity. Future research should optimize these sensor fusion techniques to improve the efficiency and effectiveness of adaptive MR systems, ensuring that they can operate robustly across a wide range of real-world applications [105, 355].

6.2.3 Towards Transitional Interfaces

Understanding the impact of transitions in MR is crucial as one of the main goals is to design transitional interfaces or cross-reality systems that can move users in environments with different degrees of reality and virtuality. Recent studies by Feld et al. [182] and Pointecker et al. [444] have provided significant insights into how transitions from real to virtual environments affect user experience. These works emphasize the role of different modalities in easing cognitive loads during changes in the environment. Feld et al. [182] research examines how transitions can be optimized to reduce user strain, suggesting that smoother, more gradual visual transitions align well with user preferences and reduce attentional burden.

Pointecker et al. [444] contribute by demonstrating a transition technique that utilizes a digital replica of the real environment to aid spatial orientation and cognitive processing during shifts. This method maintains visual coherence, ensuring continuity and ease during transitions between real and virtual settings.

Given these insights, further exploration into transitions' temporal and perceptual dynamics is essential. It is important to identify perceptual thresholds that facilitate seamless transitions and appear as smooth, i.e., not perceivable to users as possible. These thresholds might vary individually or could be standardized across different demographics, optimizing user interface design to make transitions feel natural and intuitive. The objective is to seamlessly integrate virtual and physical elements, enabling users to perceive transitions smoothly and without disruption. Researching these thresholds will allow the development of interfaces that adapt effectively to the user's environment and anticipatory needs within the MR experience. Future research should focus on testing these transitional methods across various scenarios to gauge diverse user reactions and adaptability. This approach will help refine transition techniques for broader applications, from educational platforms to immersive entertainment. Furthermore, creating predictive models to foresee user needs during transitions could significantly enhance MR system adaptability.

6.2.4 The Need for a Balanced Blend of Virtual and Physical Content

In our work, we observed that employing AV with simple cutouts supported users in typing tasks, leading to increased engagement and decreased distraction from external stimuli. The implication of AV in enhancing productivity prompts us to explore further methods to insulate users from real-world distractions effectively in the future. The utilization of cutouts has proven effective in maintaining focus on typing tasks, aligning with findings from Pham and Stuerzlinger [440] and approaches that involve removing distracting objects as discussed by Cheng et al. [95].

Pham and Stuerzlinger [440] introduced "HawKEY," a portable keyboard for VR that allows users to achieve high-speed text entry while standing, similar to desktop performance. The study evaluated different visualizations of the keyboard and demonstrated that their HawKEY system significantly improves text entry rates and reduces error rates in VR. Their study found that the Video condition, which displays the keyboard only when the user looks down, and the Point Cloud condition, similar to our video cutout, were particularly effective. These methods yielded text entry speeds close to those in real-world settings and minimized errors, highlighting the importance of effective keyboard visualization in AV.

Cheng et al. [95] developed techniques for augmenting virtual environments by selectively removing physical objects from the user's view, minimizing distractions while maintaining important elements of the user's real-world surroundings. This approach is particularly effective in tasks requiring high concentration, such as detailed design or analytical work in mixed reality settings. By carefully curating which elements remain visible, their method supports better focus and task engagement by shielding the user from potential distractions in the physical environment. This technique aligns with the needs of AV by providing a more controlled and distraction-free virtual workspace, which is essential for productivity and user satisfaction in immersive environments.

Taken together, their work highlights a critical gap in the research: the need for methods that more dynamically manage user attention in AV environments. These methods can potentially enhance productivity by minimizing distractions and guiding user focus toward task-relevant elements in the environment.

Drafting from human factors research, we can consider approaches to bolster attention and task engagement when users interact with virtual content or in AR. Specifically, "attentional funnel" refers to the process by which attention is progressively narrowed to focus on a smaller set of stimuli or tasks. This concept is often used in cognitive psychology and HCI to describe how individuals filter and prioritize information from various inputs. Here, Biocca et al. [55], starting from mobile AR, propose an innovative technique called the "omnidirectional attention funnel," designed to dynamically guide users' attention to any tracked object, person, or place within a space, enhancing user performance in navigation and object interaction tasks. This AR technique is particularly effective in improving search speed and reducing cognitive load. Using spatial attention cues, the attention funnel can significantly increase search speed and reduce perceived effort, as demonstrated in their study comparing this technique with other attentional methods like highlighting and audio cueing. In the context of AV, we foresee that this technique can be similarly applied. For instance, in an AV setup where users engage in complex tasks that require both virtual and physical interactions—such as in a simulated training environment for machinery operation—the attention funnel could dynamically highlight or turn them into virtuality, as we showed that virtual objects are more efficient in grabbing users' attention. Moreover, the omnidirectional nature of the attention funnel makes it highly suitable for environments where users need to maintain awareness of their surroundings while interacting with virtual elements. For example, in medical training simulations, it could guide a trainee's attention to specific anatomical features or surgical tools, transitioning smoothly between different focal points as the procedure progresses.

Integrating dynamic attention support techniques can significantly enhance productivity and user engagement by providing a balanced blend of virtual and physical content that caters to the needs of the task at hand. We therefore foresee three potential approaches :

Diminishing Reality

By selectively filtering out non-essential physical elements from the user's environment, this approach simplifies the user's sensory input, thereby reducing cognitive load. Techniques such as blurring or fading out irrelevant background details can help maintain concentration on critical tasks without losing the context of the physical space.

Dynamic Virtual Enhancements

This strategy involves adjusting the virtual elements within the environment based on the user's interaction and task requirements. For example, spotlighting or vignetting can enhance the visibility of important virtual objects, while adaptive filters like greyscale or selective colorization can be applied to reduce visual complexity and draw attention to specific items.

Context-Aware Content Manipulation

Leveraging the capabilities of AV to create a responsive environment, this approach tailors the virtual content to fit the user's current context. It involves introducing or modifying virtual elements such as digital overlays or augmented cues that guide the user's attention to pertinent tasks or information, effectively merging the virtual with the physical to create a cohesive and coherent interaction.

The proposed approaches—Diminishing Reality, Dynamic Virtual Enhancements, and Context-Aware Content Manipulation—strategically enhance attention management within AV by harmonizing the interplay of physical and virtual elements, ensuring a focused and effective user interaction in immersive environments.

7 CONCLUSION

"You know somethin', Utivich? I think this just might be my masterpiece."

– Lt. Aldo Raine.
Inglourious Basterds

This thesis comprehensively explored the dynamics of attention allocation and task engagement across the MR continuum, focusing on the interplay between virtual and physical components and their effects on user performance. Employing multimodal evaluations—including behavioral, subjective, and physiological measures such as EEG and eye tracking—the study systematically examined how different MR actualities influence cognitive load and visual search efficiency. Findings revealed that AV enhances user performance by effectively reducing distractions and optimizing the blend of virtual and physical information. This suggests that AV can better support task engagement and attention management compared to AR, which tends to increase cognitive demands due to its complex integration of real and virtual elements.

Furthering the exploration into adaptive systems, the research developed and tested VR environments that dynamically adapt based on physiological signals indicative of user engagement and attention states. This was achieved by creating systems that can adjust visual complexity in real-time, tailored to user responses, to maintain optimal engagement without overwhelming sensory input. The studies utilized advanced signal processing to refine the interpretation of physiological data, leading to adaptive models that predict user states with high accuracy. These adaptive VR systems demonstrated the potential to enhance interaction within MR by reducing cognitive strain and facilitating more effective user performance.

Additionally, the thesis tackled the challenges associated with embedding physiological sensing systems within MR interfaces, aiming to improve the reliability and validity of these measurements. The introduction of SensCon —a system embedded within VR controllers— illustrated the successful integration of physiological sensors that monitor user responses reliably across diverse VR scenarios. We highlight the importance and need of developing MR technologies that are both sensitive and specific to user physiological states, thereby enriching user interaction through tailored, context-aware responses.

Collectively, these contributions highlight the efficacy of merging physiological computing with MR to design interfaces that are not only adaptive but also cognitively considerate. By enhancing the MR systems' responsiveness to real-time user states, the thesis lays a starting pathway for future innovations in MR applications, ranging from productivity tools to immersive training environments. This research advances the field of HCI by demonstrating how integrated physiological measures can inform and transform user experiences in Mixed Reality settings.

References

[1] Aarts, Emile. 'Ambient intelligence: a multimedia perspective.' In: *IEEE mMltimedia* 11.1 (2004), pp. 12–19.

[2] Ahuja, Karan, Mayer, Sven, Goel, Mayank, and Harrison, Chris. 'Pose-on-the-Go: Approximating User Pose with Smartphone Sensor Fusion and Inverse Kinematics.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445582.

[3] Ahuja, Karan, Shen, Vivian, Fang, Cathy Mengying, Riopelle, Nathan, Kong, Andy, and Harrison, Chris. 'ControllerPose: Inside-Out Body Capture with VR Controller Cameras.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3502105.

[4] Alcañiz, Mariano, Bigné, Enrique, and Guixeres, Jaime. 'Virtual reality in marketing: a framework, review, and research agenda.' In: *Frontiers in psychology* 10 (2019), p. 439498.

[5] Alison, Laurence, Van Den Heuvel, Claudia, Waring, Sara, Power, Nicola, Long, Amy, O'Hara, Terence, and Crego, Jonathan. 'Immersive simulated learning environments for researching critical incidents: A knowledge synthesis of the literature and experiences of studying high-risk strategic decision making.' In: *Journal of Cognitive Engineering and decision making* (2013). DOI: https://doi.org/10.1177/1555343412468113.

[6] Allanson, Jennifer and Fairclough, Stephen H. 'A research agenda for physiological computing.' In: *Interacting with computers* 16.5 (2004), pp. 857–878. DOI: 10.1016/j.intcom.2004.08.001.

[7] Allen, Richard J, Baddeley, Alan D, and Hitch, Graham J. 'Executive and perceptual distraction in visual working memory.' In: *Journal of Experimental Psychology: Human Perception and Performance* 43.9 (2017), p. 1677.

[8] Altman, Douglas G and Bland, J Martin. 'Measurement in medicine: the analysis of method comparison studies.' In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 32.3 (1983), pp. 307–317. DOI: 10.2307/2987937.

[9] Alvarez, George A and Cavanagh, Patrick. 'The capacity of visual short-term memory is set both by visual information load and by number of objects.' In: *Psychological science* 15.2 (2004), pp. 106–111.

[10] Andersen, Steven Arild Wuyts, Konge, Lars, and Sørensen, Mads Sølvsten. 'The effect of distributed virtual reality simulation training on cognitive load during subsequent dissection training.' In: *Medical teacher* 40.7 (2018), pp. 684–689.

[11] Andreassi, John L. Psychophysiology: Human behavior & physiological response. Psychology Press, 2013.

[12] Andreassi, John L. Psychophysiology: Human behavior and physiological response. Psychology press, 2010.

[13] Andreessen, Lena M, Gerjets, Peter, Meurers, Detmar, and Zander, Thorsten O. 'Toward neuroadaptive support technologies for improving digital reading: a passive BCI-based assessment of mental workload imposed by text difficulty and presentation speed during reading.' In: *User Modeling and User-Adapted Interaction* 31 (2021), pp. 75–104. DOI: 10.1007/s11257-020-09273-5.

[14] Aricò, Pietro, Borghini, Gianluca, Di Flumeri, Gianluca, Sciaraffa, Nicolina, and Babiloni, Fabio. 'Passive BCI beyond the lab: current trends and future directions.' In: *Physiological measurement* 39.8 (2018), 08TR02. DOI: 10.1088/1361-6579/aad57e.

[15] Arquissandas, Preyesse, Lamas, David Ribeiro, and Oliveira, Jorge. 'Moving from VR into AR using bio-cybernetic loops and physiological sensory devices for intervention on anxiety disorders.' In: *Virtual Reality* (2021). DOI: 10.1007/s10055-021-00549-8.

[16] Arrabito, G Robert, Ho, Geoffrey, Aghaei, Behzad, Burns, Catherine, and Hou, Ming. 'Sustained attention in auditory and visual monitoring tasks: Evaluation of the administration of a rest break or exogenous vibrotactile signals.' In: *Human Factors* 57.8 (2015), pp. 1403–1416. DOI: 10.1177/0018720815598433.

[17] Asbee, Justin, Kelly, Kimberly, McMahan, Timothy, and Parsons, Thomas D. 'Machine learning classification analysis for an adaptive virtual reality Stroop task.' In: *Virtual Reality* (2023), pp. 1–17. DOI: 10.1007/s10055-022-00744-1.

[18] Asteriadis, Stylianos, Karpouzis, Kostas, Shaker, Noor, and Yannakakis, Georgios N. 'Towards detecting clusters of players using visual and gameplay data.' In: *International Journal of Human-Computer Studies* 67.7 (2009), pp. 569–582.

[19] Aston-Jones, G. and Cohen, J.D. 'An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance.' In: *Annual Review of Neuroscience* 28 (2005), pp. 403–450. DOI: 10.1146/annurev.neuro.28.061604.135709.

[20] Atkinson, Janette and Braddick, Oliver J. "Where'and 'what'in visual search.' In: *Perception* 18.2 (1989), pp. 181–189. DOI: 10.1068/p180181.

[21] Auda, Jonas, Gruenefeld, Uwe, Faltaous, Sarah, Mayer, Sven, and Schneegass, Stefan. 'A Scoping Survey on Cross-Reality Systems.' In: *ACM CSUR* 56.4 (2023). DOI: 10.1145/3616536.

[22] Awotunde, Joseph Bamidele, Jimoh, Rasheed Gbenga, Folorunso, Sakinat Oluwabukonla, Adeniyi, Emmanuel Abidemi, Abiodun, Kazeem Moses, and Banjo, Oluwatobi Oluwaseyi. 'Privacy and security concerns in IoT-based healthcare systems.' In: *The fusion of internet of things, artificial intelligence, and cloud computing in health care*. Springer, 2021, pp. 105–134. DOI: 10.1007/978-3-030-75220-0_6.

[23] Azuma, Ronald and Furmanski, Chris. 'Evaluating label placement for augmented reality view management.' In: *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* IEEE. 2003, pp. 66–75.

[24] Azuma, Ronald T. 'A survey of augmented reality.' In: *Presence: teleoperators & virtual environments* 6.4 (1997), pp. 355–385. DOI: 10.1162/pres.1997.6.4.355.

[25] Baayen, R Harald, Davidson, Douglas J, and Bates, Douglas M. 'Mixed-effects modeling with crossed random effects for subjects and items.' In: *Journal of memory and language* 59.4 (2008), pp. 390–412. DOI: 10.1016/j.jml.2007.12.005.

[26] Babaei, Ebrahim, Tag, Benjamin, Dingler, Tilman, and Velloso, Eduardo. 'A Critique of Electrodermal Activity Practices at CHI.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445370.

[27] Babiloni, Fabio. 'Mental workload monitoring: new perspectives from neuroscience.' In: *International Symposium on Human Mental Workload: Models and Applications*. Cham, Switzerland: Springer, 2019, pp. 3–19. DOI: 10.1007/978-3-030-32423-0_1.

[28] Bachynskyi, Myroslav, Palmas, Gregorio, Oulasvirta, Antti, and Weinkauf, Tino. 'Informing the design of novel input methods with muscle coactivation clustering.' In: *ACM Transactions on Computer-Human Interaction* 21.6 (2015), pp. 1–25.

[29] Bacim, Felipe, Ragan, Eric, Scerbo, Siroberto, Polys, Nicholas F., Setareh, Mehdi, and Jones, Brett D. 'The Effects of Display Fidelity, Visual Complexity, and Task Scope on Spatial Understanding of 3D Graphs.' In: *Proceedings of Graphics Interface 2013*. Vol. 2. GI '13 1. Regina, Sascatchewan, Canada: Canadian Information Processing Society, 2013, pp. 25–32.

[30] Bafna, Tanya, Hansen, John Paulin Paulin, and Baekgaard, Per. 'Cognitive Load during Eye-Typing.' In: *ACM ETRA*. ETRA '20 Full Papers. Stuttgart, Germany, 2020. DOI: 10.1145/3379155.3391333.

[31] Balestrucci, Priscilla and Ernst, Marc. 'Visuo-motor adaptation during interaction with a user-adaptive system.' In: *Journal of Vision* 19 (2019), 187a. DOI: 10.1167/19.10.187a.

[32] Bali, Anjana and Jaggi, Amteshwar Singh. 'Clinical experimental stress studies: methods and assessment.' In: *Reviews in the Neurosciences* 26.5 (2015), pp. 555–579. DOI: 10.1515/revneuro-2015-0004.

[33] Banville, Hubert and Falk, Tiago H. 'Recent advances and open challenges in hybrid brain-computer interfacing: a technological review of non-invasive human research.' In: *Brain-Computer Interfaces* 3.1 (2016), pp. 9–46. DOI: 10.1080/2326263X.2015.1134958.

[34] Barry, Robert J., Rushby, Jacqueline A., Wallace, Mark J., Clarke, Adam R., Johnstone, Stuart J., and Zlojutro, Ilinka. 'Caffeine effects on resting-state arousal.' In: *Clinical Neurophysiology* 116.11 (2005), pp. 2693–2700. DOI: 10.1016/j.clinph.2005.08.008.

[35] Başar, Erol, Başar-Eroğlu, Canan, Karakaş, Sirel, and Schürmann, Martin. 'Are cognitive processes manifested in event-related gamma, alpha, theta and delta oscillations in the EEG?' In: *Neuroscience letters* 259.3 (1999), pp. 165–168.

[36] Bates, Douglas, Mächler, Martin, Bolker, Ben, and Walker, Steve. 'Fitting linear mixed-effects models using lme4.' In: *arXiv preprint arXiv:1406.5823* (2014). DOI: 10.48550/arXiv.1406.5823.

[37] Batista, Diana, Plácido da Silva, Hugo, Fred, Ana, Moreira, Carlos, Reis, Margarida, and Ferreira, Hugo Alexandre. 'Benchmarking of the BITalino biomedical toolkit against an established gold standard.' In: *Healthcare technology letters* 6.2 (2019), pp. 32–36. DOI: https://doi.org/10.1049/htl.2018.5037.

[38] Beh, Helen C. 'Achievement motivation, performance and cardiovascular activity.' In: *International Journal of Psychophysiology* 10.1 (1990), pp. 39–45. DOI: 10.1016/0167-8760(90)90043-D.

[39] Beintema, Jaap A, Loon, Editha M van, and Berg, Albert V van den. 'Manipulating saccadic decision-rate distributions in visual search.' In: *Journal of vision* 5.3 (2005), pp. 1–1. DOI: 10.1167/5.3.1.

[40] Bell, Anthony J and Sejnowski, Terrence J. 'An information-maximization approach to blind separation and blind deconvolution.' In: *Neural computation* 7.6 (1995), pp. 1129–1159. DOI: 10.1162/neco.1995.7.6.1129.

[41] Ben-Shakhar, Gershon. 'Standardization within individuals: A simple method to neutralize individual differences in skin conductance.' In: *Psychophysiology* 22.3 (1985), pp. 292–299.

[42] Benedek, Mathias and Kaernbach, Christian. 'A continuous measure of phasic electrodermal activity.' In: *Journal of Neuroscience Methods* 190.1 (2010), pp. 80–91. DOI: https://doi.org/10.1016/j.jneumeth.2010.04.028.

[43] Benedek, Mathias and Kaernbach, Christian. 'Decomposition of skin conductance data by means of nonnegative deconvolution.' In: *psychophysiology* 47.4 (2010), pp. 647–658. DOI: 10.1111/j.1469-8986.2009.00972.x.

[44] Benedek, Mathias, Schickel, Rainer J, Jauk, Emanuel, Fink, Andreas, and Neubauer, Aljoscha C. 'Alpha power increases in right parietal cortex reflects focused internal attention.' In: *Neuropsychologia* 56 (2014), pp. 393–400. DOI: 10.1016/j.neuropsychologia.2014.02.010.

[45] Bergefurt, Lisanne, Appel-Meulenbroek, Rianne, and Arentze, Theo. 'How salutogenic workplace characteristics influence psychological and cognitive responses in a virtual environment.' In: *Ergonomics* just-accepted (2023), pp. 1–22. DOI: https://doi.org/10.1080/00140139.2023.2223372.

[46] Berger, Alexander and Kiefer, Markus. 'Comparison of different response time outlier exclusion methods: a simulation study.' In: *Frontiers in Psychology* (2021). DOI: 10.3389/fpsyg.2021.675558.

[47] Bernal, Guillermo, Hidalgo, Nelson, Russomanno, Conor, and Maes, Pattie. 'Galea: A physiological sensing system for behavioral research in virtual environments.' In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. New York, NY, USA: IEEE, 2022, pp. 66–76. DOI: 10.1109/VR51125.2022.00024.

[48] Bernal, Guillermo, Yang, Tao, Jain, Abhinandan, and Maes, Pattie. 'PhysioHMD: A Conformable, Modular Toolkit for Collecting Physiological Data from Head-Mounted Displays.' In: *Proceedings of the 2018 ACM International Symposium on Wearable Computers*. ISWC '18. Singapore, Singapore: Association for Computing Machinery, 2018, pp. 160–167. DOI: 10.1145/3267242.3267268.

[49] Bernhardt, Kyle A., Poltavski, Dmitri, Petros, Thomas, and Ferraro, F. Richard. 'Differentiating Active and Passive Fatigue with the use of Electroencephalography.' In: *Proc. of HFES* (2019). DOI: 10.1177/1071181319631106.

[50] Bezovski, Zlatko. 'The future of the mobile payment as electronic payment system.' In: *European Journal of Business and Management* 8.8 (2016), pp. 127–132. DOI: https://iiste.org/Journals/index.php/EJBM/article/view/29473/30263.

[51] Biddiss, Elaine and Irwin, Jennifer. 'Active video games to promote physical activity in children and youth: a systematic review.' In: *Archives of pediatrics & adolescent medicine* 164.7 (2010), pp. 664–672. DOI: 10.1001/archpediatrics.2010.104.

[52] Bigdely-Shamlo, Nima, Mullen, Tim, Kothe, Christian, Su, Kyung-Min, and Robbins, Kay A. 'The PREP pipeline: standardized preprocessing for large-scale EEG analysis.' In: *Frontiers in neuroinformatics* 9 (2015), p. 16. DOI: 10.3389/fninf.2015.00016.

[53] Billinghurst, Mark, Clark, Adrian, Lee, Gun, et al. 'A survey of augmented reality.' In: *Foundations and Trends® in HCI* 8.2-3 (2015), pp. 73–272.

[54] Billinghurst, Mark and Duenser, Andreas. 'Augmented reality in the classroom.' In: *Computer* 45.7 (2012), pp. 56–63.

[55] Biocca, Frank, Owen, Charles, Tang, Arthur, and Bohil, Corey. 'Attention issues in spatial information systems: Directing mobile users' visual attention using augmented reality.' In: *Journal of Management Information Systems* 23.4 (2007), pp. 163–184. DOI: 10.2753/MIS0742-1222230408.

[56] Blair, Mark R, Watson, Marcus R, Walshe, R Calen, and Maj, Fillip. 'Extremely selective attention: eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization.' In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35.5 (2009), p. 1196.

[57] Bland, J Martin and Altman, Douglas G. 'Measuring agreement in method comparison studies.' In: *Statistical Methods in Medical Research* 8.2 (1999), pp. 135–160. DOI: 10.1177/096228029900800204. eprint: https://doi.org/10.1177/096228029900800204.

[58] Bonanni, Leonardo, Lee, Chia-Hsun, and Selker, Ted. 'Attention-based design of augmented reality interfaces.' In: *CHI'05 extended abstracts on Human factors in computing systems*. 2005, pp. 1228–1231. DOI: 10.1145/1056808.1056883.

[59] Boring, Sebastian, Jurmu, Marko, and Butz, Andreas. 'Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays.' In: *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*. OZCHI '09. Melbourne, Australia: Association for Computing Machinery, 2009, pp. 161–168. DOI: 10.1145/1738826.1738853.

[60] Bortz, Jürgen and Döring, Nicola. Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler: Limitierte Sonderausgabe. Springer-Verlag, 2007. DOI: 10.1007/978-3-540-33306-7.

[61] Bosworth, Andrew Garrod. Body pose estimation using self-tracked controllers. 2022.

[62] Boucsein, Wolfram. 'The use of psychophysiology for evaluating stress-strain processes in human-computer interaction.' In: *Engineering Psychophysiology. Issues and Applications* (2000), pp. 289–309.

[63] Boucsein, Wolfram and Backs, Richard W. 'Engineering psychophysiology as a discipline: Historical and theoretical aspects.' In: *Engineering psychophysiology: Issues and applications* (2000), pp. 3–30.

[64] Boucsein, Wolfram, Baltissen, Ruediger, and Euler, Manfred. 'Dependence of skin conductance reactions and skin resistance reactions upon previous level.' In: *Psychophysiology* 21.2 (1984), pp. 212–218. DOI: 10.1111/j.1469-8986.1984.tb00207.x.

[65] Boucsein, Wolfram, Haarmann, Andrea, and Schaefer, Florian. 'Combining skin conductance and heart rate variability for adaptive automation during simulated IFR flight.' In: *Int. Conf. on Engineering Psychology and Cognitive Ergonomics*. Springer, 2007, pp. 639–647. DOI: 10.1007/978-3-540-73331-7_70.

[66] Bourdillon, Nicolas, Schmitt, Laurent, Yazdani, Sasan, Vesin, Jean-Marc, and Millet, Grégoire P. 'Minimal window duration for accurate HRV recording in athletes.' In: *Front. in neuroscience* (2017). DOI: https://doi.org/10.3389/fnins.2017.00456.

[67] Bouta, H. and Retalis, S. 'Utilizing a collaborative macro-script to enhance student engagement: A mixed method study in a 3D virtual environment.' In: *Computers Education* 60 (2013), pp. 122–135. DOI: 10.1016/j.compedu.2012.07.002.

[68] Boutcher, Yati N and Boutcher, Stephen H. 'Cardiovascular response to Stroop: effect of verbal response and task difficulty.' In: *Biological psychology* 73.3 (2006), pp. 235–241. DOI: 10.1016/j.biopsycho.2006.04.005.

[69] Bradski, Gary and Kaehler, Adrian. Learning OpenCV: Computer vision with the OpenCV library. Sebastopol, California, United States: O'Reilly Media, Inc., 2008.

[70] Braithwaite, Jason J, Watson, Derrick G, Jones, Robert, and Rowe, Mickey. 'A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments.' In: *Psychophysiology* 49.1 (2013), pp. 1017–1034.

[71] Braver, Todd S. 'The variable nature of cognitive control: a dual mechanisms framework.' In: *Trends in cognitive sciences* 16.2 (2012), pp. 106–113. DOI: 10.1016/j.tics.2011.12.010.

[72] Brewer, Cynthia A. 'Prediction of simultaneous contrast between map colors with Hunt's model of color appearance.' In: *Color Research & Application* 21.3 (1996), pp. 221–235. DOI: 10.1002/(SICI)1520-6378(199606)21:3<221::AID-COL4>3.0.CO;2-U.

[73] Broadbent, Donald E. Perception and Communication. Pergamon Press, 1958.

[74] Brouwer, Anne-Marie, Hogervorst, Maarten A, Holewijn, Michael, and Erp, Jan BF van. 'Evidence for effects of task difficulty but not learning on neurophysiological

variables associated with effort.' In: *International Journal of Psychophysiology* 93.2 (2014), pp. 242–252. DOI: 10.1016/j.ijpsycho.2014.05.004.

[75] Brouwer, Anne-Marie, Hogervorst, Maarten A, Oudejans, Bob, Ries, Anthony J, and Touryan, Jonathan. 'EEG and eye tracking signatures of target encoding during structured visual search.' In: *Frontiers in human neuroscience* 11 (2017), p. 264. DOI: 10.3389/fnhum.2017.00264.

[76] Brouwer, Anne-Marie, Hogervorst, Maarten A, Van Erp, Jan BF, Heffelaar, Tobias, Zimmerman, Patrick H, and Oostenveld, Robert. 'Estimating workload using EEG spectral power and ERPs in the n-back task.' In: *Journal of neural engineering* 9.4 (2012), p. 045008. DOI: 10.1088/1741-2560/9/4/045008.

[77] Brouwer, Anne-Marie, Reuderink, Boris, Vincent, Joris, Gerven, Marcel AJ van, and Erp, Jan BF van. 'Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials.' In: *Journal of vision* 13.3 (2013), pp. 17–17. DOI: 10.1167/13.3.17.

[78] Buetti, Simona and Lleras, Alejandro. 'Distractibility is a function of engagement, not task difficulty: Evidence from a new oculomotor capture paradigm.' In: *Journal of Experimental Psychology: General* 145.10 (2016), p. 1382.

[79] Burns, Christopher G and Fairclough, Stephen H. 'Use of auditory event-related potentials to measure immersion during a computer game.' In: *International Journal of Human-Computer Studies* 73 (2015), pp. 107–114. DOI: 10.1016/j.ijhcs.2014.09.002.

[80] Busch, Niko A and Herrmann, Christoph S. 'Object-load and feature-load modulate EEG in a short-term memory task.' In: *Neuroreport* 14.13 (2003), pp. 1721–1724. DOI: 10.1097/00001756-200309150-00013.

[81] Cacioppo, John T and Tassinary, Louis G. Principles of psychophysiology: Physical, social, and inferential elements. Cambridge University Press, 1990.

[82] Cacioppo, John T., Berntson, Gary G., Malarkey, William B., Kiecolt-Glaser, Janice K., Sheridan, John F., Poehlmann, Kirsten M., Burleson, Mary H., Ernst, John M., Hawkley, Louise C., and Glaser, Ronald. 'Autonomic, neuroendocrine, and immune responses to psychological stress: The reactivity hypothesis A.' In: *Annals of the New York Academy of Sciences* 840.1 (1998), pp. 664–673. DOI: 10.1111/j.1749-6632.1998.tb09605.x.

[83] Cacioppo, John T., Tassinary, Louis G., and Berntson, Gary. Handbook of psychophysiology. Cambridge, UK: Cambridge university press, 2007. DOI: 10.1017/9781107415782.

[84] Campbell, Joey and Fraser, Mike. 'Switching It Up: Designing Adaptive Interfaces for Virtual Reality Exergames.' In: *Proceedings of the 31st European Conference on Cog-*

nitive Ergonomics. ECCE '19. BELFAST, United Kingdom: Association for Computing Machinery, 2019, pp. 177–184. DOI: 10.1145/3335082.3335087.

[85] Carter, Benjamin T and Luke, Steven G. 'Best practices in eye tracking research.' In: *Int. Journal of Psychophysiology* 155 (2020), pp. 49–62. DOI: 10.1016/j.ijpsycho.2020.05.010.

[86] Causse, M., Dehais, F., Pastor, J., and Grau, J.Y. 'Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness.' In: *Applied Ergonomics* 44.4 (2013), pp. 535–544. DOI: 10.1016/j.apergo.2012.11.012.

[87] Chalmers, Alan and Ferko, Andrej. 'Levels of realism: From virtual reality to real virtuality.' In: *Proceedings of the 24th Spring Conf. on Computer Graphics*. 2008, pp. 19–25. DOI: 10.1145/1921264.1921272.

[88] Charles, Rebecca L. and Nixon, Jim. 'Measuring mental workload using physiological measures: A systematic review.' In: *Applied Ergonomics* 74 (2019), pp. 221–232. DOI: 10.1016/j.apergo.2018.08.028.

[89] Charlot, Keyne, Cornolo, Jeremy, Brugniaux, Julien V, Richalet, Jean-Paul, and Pichon, Aurélien. 'Interchangeability between heart rate and photoplethysmography variabilities during sympathetic stimulations.' In: *Physiological measurement* 30.12 (2009), p. 1357. DOI: 10.1088/0967-3334/30/12/005.

[90] Chavarriaga, Ricardo, Sobolewski, Aleksander, and Millán, José del R. 'Errare machinale est: the use of error-related potentials in brain-machine interfaces.' In: *Frontiers in Neuroscience* (2014), p. 208.

[91] Chen, Weixuan, Jaques, Natasha, Taylor, Sara, Sano, Akane, Fedor, Szymon, and Picard, Rosalind W. 'Wavelet-based motion artifact removal for electrodermal activity.' In: *2015 37th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*. NY, NY, USA: IEEE, 2015, pp. 6223–6226. DOI: 10.1109/EMBC.2015.7319814.

[92] Chen, Xiaojiao, Tang, Xiaoteng, Zhao, Ying, Huang, Tengyu, Qian, Ran, Zhang, Jiayi, Chen, Wei, and Wang, Xiaosong. 'Evaluating Visual Consistency of Icon Usage in Across-Devices.' In: *Int. Journal of Human–Computer Interaction* (2023), pp. 1–17. DOI: 10.1080/10447318.2022.2162275.

[93] Chen, Xin and Zelinsky, Gregory J. 'Real-world visual search is dominated by top-down guidance.' In: *Vision Research* 46.24 (2006), pp. 4118–4133. DOI: 10.1016/j.visres.2006.08.008.

[94] Chen, Youguo and Huang, Xiting. 'Modulation of alpha and beta oscillations during an n-back task with varying temporal memory load.' In: *Frontiers in psychology* 6 (2016), p. 2031. DOI: 10.3389/fpsyg.2015.02031.

[95] Cheng, Yi Fei, Yin, Hang, Yan, Yukang, Gugenheimer, Jan, and Lindlbauer, David. 'Towards Understanding Diminished Reality.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3517452.

[96] Cheng, Yifei, Yan, Yukang, Yi, Xin, Shi, Yuanchun, and Lindlbauer, David. 'Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections.' In: *ACM UIST*. 2021, pp. 282–297. DOI: 10.1145/3472749.3474750.

[97] Chiossi, Francesco, El Khaoudi, Yassmine, Ou, Changkun, Sidenmark, Ludwig, Zaky, Abdelrahman, M., Ines, Ou, Changkun, Feuchtner, Tiare, and Mayer, Sven. 'Evaluating Typing Experience in Different Mixed Reality Manifestations Using Physiological Features.' In: *Proc. ACM Hum.-Comput. Interact* 10 (2024).

[98] Chiossi, Francesco, Gruenefeld, Uwe, Hou, Baosheng James, Newn, Joshua, Ou, Changkun, Liao, Rulu, Welsch, Robin, and Mayer, Sven. 'Understanding the Impact of the Reality-Virtuality Continuum on Visual Search Using Fixation-Related Potentials and Eye Tracking Features.' In: *Proc. ACM Hum.-Comput. Interact* 8 (2024). DOI: 10.1145/36765281.

[99] Chiossi, Francesco, Haliburton, Luke, Ou, Changkun, Butz, Andreas Martin, and Schmidt, Albrecht. 'Short-Form Videos Degrade Our Capacity to Retain Intentions: Effect of Context Switching On Prospective Memory.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: ACM, 2023. DOI: 10.1145/3544548.3580778.

[100] 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). DOI: 10.1145/3604270.

[101] Chiossi, Francesco and Mayer, Sven. 'How Can Mixed Reality Benefit From Physiologically-Adaptive Systems? Challenges and Opportunities for Human Factors Applications.' In: *arXiv preprint arXiv:2303.17978* (2023).

[102] Chiossi, Francesco, Ou, Changkun, Gerhardt, Carolina, Putze, Felix, and Mayer, Sven. 'Designing and Evaluating an Adaptive Virtual Reality System using EEG Frequencies to Balance Internal and External Attention States.' In: *International Journal of Human-Computer Studies* 1 (2023), pp. 1–59.

[103] Chiossi, Francesco, Ou, Changkun, and Mayer, Sven. 'Exploring Physiological Correlates of Visual Complexity Adaptation: Insights from EDA, ECG, and EEG Data for Adaptation Evaluation in VR Adaptive Systems.' In: *ACM Conference on Human-Computer Interaction*. CHI '23 Extended Abstracts. Hamburg, Germany: ACM, 2023. DOI: 10.1145/3544549.3585624.

[104] Chiossi, Francesco, Ou, Changkun, and Mayer, Sven. 'Optimizing Visual Complexity for Physiologically-Adaptive VR Systems: Evaluating a Multimodal Dataset using EDA, ECG and EEG Features.' In: *International Conference on Advanced Visual Interfaces 2024 (AVI 2024), June 3–7, 2024*. Arenzano, Genoa, Italy: ACM, 2024. DOI: 10.1145/3656650.3656657.

[105] Chiossi, Francesco, Ou, Changkun, Putze, Felix, and Mayer, Sven. 'Detecting Internal and External Attention in Virtual Reality: A Comparative Analysis of EEG Classification Methods.' In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. MUM '24. Association for Computing Machinery, 2024, pp. 381–395. DOI: 10.1145/3701571.3701579.

[106] Chiossi, Francesco, Stepanova, Ekaterina R, Tag, Benjamin, Perusquia-Hernandez, Monica, Kitson, Alexandra, Dey, Arindam, Mayer, Sven, and El Ali, Abdallah. 'PhysioCHI: Towards Best Practices for Integrating Physiological Signals in HCI.' In: *arXiv e-prints* (2023), arXiv–2312.

[107] Chiossi, Francesco, Trautmannsheimer, Ines M., Ou, Changkun, Gruenefeld, Uwe, and Mayer, Sven. 'Searching Across Realities: Investigating ERPs and Eye-Tracking Correlates of Visual Search in Mixed Reality.' In: *IEEE Transactions on Visualization and Computer Graphics* (2024).

[108] Chiossi, Francesco, Turgut, Yagiz, Welsch, Robin, and Mayer, Sven. 'Adapting Visual Complexity Based on Electrodermal Activity Improves Working Memory Performance in Virtual Reality.' In: *Proc. ACM Hum.-Comput. Interact.* 7.MHCI (2023), pp. 1–26. DOI: 10.1145/3604243.

[109] Chiossi, Francesco, Weiss, Yannick, Steinbrecher, Thomas, Mai, Christian, and Kosch, Thomas. 'Mind the Visual Discomfort: Assessing Event-Related Potentials as Indicators for Visual Strain in Head-Mounted Displays.' In: *2024 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR. New York, NY, USA: IEEE, 2024, pp. 150–159. DOI: 10.1109/ISMAR62088.2024.00029.

[110] 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), p. 55. DOI: 10.3390/bdcc6020055.

[111] 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*. IEEE VR'22. IEEE, 12, 2022. DOI: 10.1109/VRW55335.2022.00140.

[112] Chiossi, Francesco, Welsch, Robin, Villa Salazar, David Steeven, Chuang, Lewis, and Mayer, Sven. 'Virtual Reality Adaptation using Electrodermal Activity to Support User Experience.' Version V1. In: (2022). DOI: 10.18419/darus-2820.

[113] Chiossi, Francesco, Zagermann, Johannes, Karolus, Jakob, Rodrigues, Nils, Balestrucci, Priscilla, Weiskopf, Daniel, Ehinger, Benedikt, Feuchtner, Tiare, Reiterer, Harald, Chuang, Lewis L, et al. 'Adapting visualizations and interfaces to the user.' In: *it-Information Technology* 64.4-5 (2022), pp. 133–143.

[114] Chiossi, Francesco, Zagermann, Johannes, Karolus, Jakob, Rodrigues, Nils, Balestrucci, Priscilla, Weiskopf, Daniel, Ehinger, Benedikt, Feuchtner, Tiare, Reiterer, Harald, Chuang, Lewis L., Ernst, Marc, Bulling, Andreas, Mayer, Sven, and Schmidt, Albrecht. 'Adapting visualizations and interfaces to the user.' In: *it - Information Technology* (2022). DOI: 10.1515/itit-2022-0035.

[115] Choi, Jongyoon and Gutierrez-Osuna, Ricardo. 'Using Heart Rate Monitors to Detect Mental Stress.' In: *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*. New York, NY, USA: IEEE, 2009, pp. 219–223. DOI: 10.1109/BSN.2009.13.

[116] Chuen Yee Lo, Barbara, Lau, Shun, Cheung, Sing-Hang, and Allen, Nicholas B. 'The impact of rumination on internal attention switching.' In: *Cognition & emotion* 26.2 (2012), pp. 209–223.

[117] Chun, Marvin M, Golomb, Julie D, and Turk-Browne, Nicholas B. 'A taxonomy of external and internal attention.' In: *Annual review of psychology* 62 (2011). DOI: 10.1146/annurev.psych.093008.100427.

[118] Cleworth, Taylor W, Horslen, Brian C, and Carpenter, Mark G. 'Influence of real and virtual heights on standing balance.' In: *Gait & posture* 36.2 (2012), pp. 172–176. DOI: 10.1016/j.gaitpost.2012.02.010.

[119] Collins, Jonny, Regenbrecht, Holger, Langlotz, Tobias, Can, Yekta Said, Ersoy, Cem, and Butson, Russell. 'Measuring Cognitive Load and Insight: A Methodology Exemplified in a Virtual Reality Learning Context.' In: *2019 IEEE International Symposium on Mixed and Augmented Reality*. ISMAR '19. New York, NY, USA: IEEE, 2019, pp. 351–362. DOI: 10.1109/ISMAR.2019.00033.

[120] Cona, Giorgia, Chiossi, Francesco, Di Tomasso, Silvia, Pellegrino, Giovanni, Piccione, Francesco, Bisiacchi, Patrizia, and Arcara, Giorgio. 'Theta and alpha oscillations as signatures of internal and external attention to delayed intentions: A magnetoencephalography (MEG) study.' In: *NeuroImage* 205 (2020), p. 116295. DOI: 10.1016/j.neuroimage.2019.116295.

[121] Conover, William Jay. Practical nonparametric statistics. Vol. 350. John Wiley & Sons, 1999. DOI: 10.2307/2981807.

[122] Cooper, Emily A. 'The Perceptual Science of Augmented Reality.' In: *Annual Review of Vision Science* 9 (2023). DOI: 10.1146/annurev-vision-111022-123758.

[123] Corbetta, Maurizio and Shulman, Gordon L. 'Control of goal-directed and stimulus-driven attention in the brain.' In: *Nature reviews neuroscience* 3.3 (2002), pp. 201–215.

[124] Corcoran, Andrew W, Alday, Phillip M, Schlesewsky, Matthias, and Bornkessel-Schlesewsky, Ina. 'Toward a reliable, automated method of individual alpha frequency (IAF) quantification.' In: *Psychophysiology* 55.7 (2018), e13064. DOI: 10.1111/psyp.13064.

[125] Cortese, Samuele, Ferrin, Maite, Brandeis, Daniel, Holtmann, Martin, Aggensteiner, Pascal, Daley, David, Santosh, Paramala, Simonoff, Emily, Stevenson, Jim, Stringaris, Argyris, et al. 'Neurofeedback for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials.' In: *Journal of the American Academy of Child & Adolescent Psychiatry* 55.6 (2016), pp. 444–455.

[126] Cruz-Neira, Carolina, Sandin, Daniel J, DeFanti, Thomas A, Kenyon, Robert V, and Hart, John C. 'The CAVE: Audio visual experience automatic virtual environment.' In: *Communications of the ACM* 35.6 (1992), pp. 64–73.

[127] Csikszentmihalyi, Mihaly, Abuhamdeh, Sami, and Nakamura, Jeanne. 'Flow.' In: *Handbook of competence and motivation* (2005), pp. 598–608.

[128] d'Agostino, Ralph B. 'An omnibus test of normality for moderate and large size samples.' In: *Biometrika* 58.2 (1971), pp. 341–348. DOI: https://doi.org/10.2307/2334522.

[129] Dahl, David B., Scott, David, Roosen, Charles, Magnusson, Arni, and Swinton, Jonathan. xtable: Export Tables to LaTeX or HTML. 2019.

[130] Darvas, F., Scherer, R., Ojemann, J.G., Rao, R.P., Miller, K.J., and Sorensen, L.B. 'High gamma mapping using EEG.' In: *NeuroImage* 49.1 (2010), pp. 930–938. DOI: 10.1016/j.neuroimage.2009.08.041.

[131] David-John, Brendan, Peacock, Candace, Zhang, Ting, Murdison, T. Scott, Benko, Hrvoje, and Jonker, Tanya R. 'Towards gaze-based prediction of the intent to interact in virtual reality.' In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '21 Short Papers. Virtual Event, Germany: Association for Computing Machinery, 2021. DOI: 10.1145/3448018.3458008.

[132] De Fockert, Jan W. 'Beyond perceptual load and dilution: a review of the role of working memory in selective attention.' In: *Frontiers in psychology* 4 (2013), p. 287.

[133] De Jong, Kenneth. ‘Adaptive system design: a genetic approach.’ In: *IEEE Transactions on Systems, Man, and Cybernetics* 10.9 (1980), pp. 566–574.

[134] De Rivecourt, Michel, Kuperus, MN, Post, WJ, and Mulder, LJM. ‘Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight.’ In: *Ergonomics* 51.9 (2008), pp. 1295–1319. DOI: 10.1080/00140130802120267.

[135] Deeb, Rasha, Ooms, Kristien, Brychtová, Alžběta, Van Eetvelde, Veerle, and De Maeyer, Philippe. ‘Background and foreground interaction: Influence of complementary colors on the search task.’ In: *Color Research & Application* 40.5 (2015), pp. 437–445. DOI: 10.1002/col.21920.

[136] Degno, Federica and Liversedge, Simon P. ‘Eye movements and fixation-related potentials in reading: a review.’ In: *Vision* 4.1 (2020), p. 11.

[137] Dehais, F., Causse, M., Vachon, F., Regis, N., Menant, E., and Tremblay, S. ‘Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness.’ In: *Human Factors* 61.5 (2019), pp. 753–765. DOI: 10.1177/0018720819834775.

[138] Dehais, F., Rida, I., Roy, R., El-Khoury, S., and Causse, M. ‘A Neuroergonomic Approach to Mental Workload, Engagement, and Human Performance.’ In: *Frontiers in Neuroscience* 14 (2020), p. 268. DOI: 10.3389/fnins.2020.00268.

[139] Deiber, Marie-Pierre, Missonnier, Pascal, Bertrand, Olivier, Gold, Gabriel, Fazio-Costa, Lara, Ibanez, Vicente, and Giannakopoulos, Panteleimon. ‘Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics.’ In: *Journal of cognitive neuroscience* 19.1 (2007), pp. 158–172. DOI: 10.1162/jocn.2007.19.1.158.

[140] Derby, Jessyca L and Chaparro, Barbara S. ‘Evaluating the Usability of Augmented Reality (AR) Applications: The Impact of The Device Used.’ In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 66. 1. SAGE Publications Sage CA: Los Angeles, CA. 2022, pp. 1355–1359. DOI: 10.1177/1071181322661431.

[141] Desor, Janet A. ‘Toward a psychological theory of crowding.’ In: *Journal of personality and social psychology* 21.1 (1972), p. 79. DOI: 10.1037/h0032112.

[142] Deubel, Heiner and Schneider, Werner X. ‘Saccade target selection and object recognition: Evidence for a common attentional mechanism.’ In: *Vision Research* 36.12 (1996), pp. 1827–1837.

[143] Deuschel, Tilman. ‘On the Influence of Human Factors in Adaptive User Interface Design.’ In: *Adjunct Publication of the 26th Conf. on User Modeling, Adaptation and Personalization*. UMAP ’18. Singapore, Singapore: ACM, 2018, pp. 187–190. DOI: 10.1145/3213586.3213587.

[144] Dey, Anind K. 'Providing Architectural Support for Building Context-Aware Applications.' PhD thesis. Georgia Institute of Technology, 2000.

[145] Dey, Arindam, Chatburn, Alex, and Billinghurst, Mark. 'Exploration of an EEG-Based Cognitively Adaptive Training System in Virtual Reality.' In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 220–226. DOI: 10.1109/VR.2019.8797840.

[146] Dhakal, Vivek, Feit, Anna Maria, Kristensson, Per Ola, and Oulasvirta, Antti. 'Observations on Typing from 136 Million Keystrokes.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–12. DOI: 10.1145/3173574.3174220.

[147] Di Lascio, Elena, Gashi, Shkurta, and Santini, Silvia. 'Unobtrusive Assessment of Students' Emotional Engagement during Lectures Using Electrodermal Activity Sensors.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2.3 (2018). DOI: 10.1145/3264913.

[148] Dietz, Dennis, Oechsner, Carl, Ou, Changkun, Chiossi, Francesco, Sarto, Fabio, Mayer, Sven, and Butz, Andreas. 'Walk This Beam: Impact of Different Balance Assistance Strategies and Height Exposure on Performance and Physiological Arousal in VR.' In: *28th ACM Symposium on Virtual Reality Software and Technology*. VRST '22. Tsukuba, Japan: Association for Computing Machinery, 2022. DOI: 10.1145/3562939.3567818.

[149] Ding, Ye, Kim, Myunghee, Kuindersma, Scott, and Walsh, Conor J. 'Human-in-the-loop optimization of hip assistance with a soft exosuit during walking.' In: *Science Robotics* 3.15 (2018), eaar5438. DOI: 10.1126/scirobotics.aar5438.

[150] Dinh, Huong Q, Walker, Neff, Hodges, Larry F, Song, Chang, and Kobayashi, Akira. 'Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments.' In: *Proceedings IEEE Virtual Reality*. IEEE. 1999, pp. 222–228. DOI: 10.1109/VR.1999.756955.

[151] Dippel, Gabriel, Mückschel, Moritz, Ziemssen, Tjalf, and Beste, Christian. 'Demands on response inhibition processes determine modulations of theta band activity in superior frontal areas and correlations with pupillometry–Implications for the norepinephrine system during inhibitory control.' In: *NeuroImage* 157 (2017), pp. 575–585. DOI: https://doi.org/10.1016/j.neuroimage.2017.06.037.

[152] Dix, Alan. Human-computer interaction. Pearson Education, 2003.

[153] Dixon, Benjamin J, Daly, Michael J, Chan, Harley, Vescan, Allan D, Witterick, Ian J, and Irish, Jonathan C. 'Surgeons blinded by enhanced navigation: the effect of augmented reality on attention.' In: *Surgical endoscopy* 27 (2013), pp. 454–461. DOI: 10.1007/s00464-012-2457-3.

[154] Doberenz, Sigrun, Roth, Walton T, Wollburg, Eileen, Maslowski, Nina I, and Kim, Sunyoung. 'Methodological considerations in ambulatory skin conductance monitoring.' In: *International Journal of Psychophysiology* 80.2 (2011), pp. 87–95. DOI: 10.1016/j.ijpsycho.2011.02.002.

[155] Doğan, Nurettin Özgür. 'Bland-Altman analysis: A paradigm to understand correlation and agreement.' In: *Turkish journal of emergency medicine* 18.4 (2018), pp. 139–141. DOI: 10.1016/j.tjem.2018.09.001.

[156] Dooren, Marieke van, Janssen, Joris H, et al. 'Emotional sweating across the body: Comparing 16 different skin conductance measurement locations.' In: *Physiology & behavior* 106.2 (2012), pp. 298–304. DOI: 10.1016/j.physbeh.2012.01.020.

[157] Doswell, Jayfus T and Skinner, Anna. 'Augmenting human cognition with adaptive augmented reality.' In: *Foundations of Augmented Cognition. Advancing Human Performance and Decision-Making through Adaptive Systems: 8th Int. Conference, AC 2014, Held as Part of HCI Int. 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings 8*. Springer. 2014, pp. 104–113. DOI: 10.1007/978-3-319-07527-3_10.

[158] Drewes, Jan, Feder, Sascha, and Einhäuser, Wolfgang. 'Gaze during locomotion in virtual reality and the real world.' In: *Frontiers in Neuroscience* 15 (2021), p. 656913. DOI: https://doi.org/10.3389/fnins.2021.656913.

[159] Duchowski, Andrew T. 'Gaze-Based Interaction: A 30 Year Retrospective.' In: *Computers & Graphics* 73 (2018), pp. 59–69. DOI: 10.1016/j.cag.2018.04.002.

[160] Duchowski, Andrew T., Krejtz, Krzysztof, Krejtz, Izabela, Biele, Cezary, Niedzielska, Anna, Kiefer, Peter, Raubal, Martin, and Giannopoulos, Ioannis. 'The Index of Pupillary Activity: Measuring Cognitive Load vis-à-vis Task Difficulty with Pupil Oscillation.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. DOI: 10.1145/3173574.3173856.

[161] Duncan, John and Humphreys, Glyn W. 'Visual search and stimulus similarity.' In: *Psychological review* 96.3 (1989), p. 433. DOI: 10.1037/0033-295X.96.3.433.

[162] Dussault, Caroline, Jouanin, Jean-Claude, Philippe, Matthieu, and Guezennec, Charles-Yannick. 'EEG and ECG changes during simulator operation reflect mental workload and vigilance.' In: *Aviation, space, and environmental medicine* (2005).

[163] Eckstein, Miguel P. 'Visual search: A retrospective.' In: *Journal of vision* 11.5 (2011), pp. 14–14. DOI: 10.1167/11.5.14. eprint: https://arvojournals.org/arvo/content_public/journal/jov/933487/jov-11-5-14.pdf.

[164] Elkin, Lisa A., Kay, Matthew, Higgins, James J., and Wobbrock, Jacob O. 'An Aligned Rank Transform Procedure for Multifactor Contrast Tests.' In: *The 34th Annual ACM*

Symposium on UIST. UIST ’21. Virtual Event, USA: ACM, 2021, pp. 754–768. DOI: 10.1145/3472749.3474784.

[165] ElKomy, Mai, Abdelrahman, Yomna, Funk, Markus, Dingler, Tilman, Schmidt, Albrecht, and Abdennadher, Slim. ‘ABBAS: an adaptive bio-sensors based assistive system.’ In: *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*. 2017, pp. 2543–2550.

[166] Enders, Leah R, Smith, Robert J, Gordon, Stephen M, Ries, Anthony J, and Touryan, Jonathan. ‘Gaze behavior during navigation and visual search of an open-world virtual environment.’ In: *Frontiers in Psychology* 12 (2021), p. 681042.

[167] Eng, Hing Yee, Chen, Diyu, and Jiang, Yuhong. ‘Visual working memory for simple and complex visual stimuli.’ In: *Psychonomic bulletin & review* 12.6 (2005), pp. 1127–1133. DOI: 10.3758/BF03206454.

[168] Engbert, Ralf and Kliegl, Reinhold. ‘Microsaccades uncover the orientation of covert attention.’ In: *Vision research* 43.9 (2003), pp. 1035–1045. DOI: 10.1016/S0042-6989(03)00084-1.

[169] Ens, Barrett, Lanir, Joel, Tang, Anthony, Bateman, Scott, Lee, Gun, Piumsomboon, Thammathip, and Billinghurst, Mark. ‘Revisiting collaboration through mixed reality: The evolution of groupware.’ In: *Int. J. Hum.-Comput. Stud.* 131.C (2019), pp. 81–98. DOI: 10.1016/j.ijhcs.2019.05.011.

[170] Ergenoglu, Tolgay, Demiralp, Tamer, Bayraktaroglu, Zubeyir, Ergen, Mehmet, Beydagi, Huseyin, and Uresin, Yagiz. ‘Alpha rhythm of the EEG modulates visual detection performance in humans.’ In: *Cognitive brain research* 20.3 (2004), pp. 376–383. DOI: 10.1016/j.cogbrainres.2004.03.009.

[171] Ernst, Marc O. and Bülthoff, Heinrich H. ‘Merging the senses into a robust percept.’ In: *Trends in Cognitive Sciences* 8.4 (2004), pp. 162–169. DOI: 10.1016/j.tics.2004.02.002.

[172] Estupiñán, Sergio, Rebelo, Francisco, Noriega, Paulo, Ferreira, Carlos, and Duarte, Emília. ‘Can virtual reality increase emotional responses (Arousal and Valence)? A pilot study.’ In: *Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part II 3*. Cham, Switzerland: Springer, 2014, pp. 541–549. DOI: 10.1007/978-3-319-07626-3_51.

[173] Euser, Anne M, Dekker, Friedo W, and Le Cessie, Saskia. ‘A practical approach to Bland-Altman plots and variation coefficients for log transformed variables.’ In: *Journal of clinical epidemiology* 61.10 (2008), pp. 978–982. DOI: 10.1016/j.jclinepi.2007.11.003.

[174] Evangelista Belo, João Marcelo, Feit, Anna Maria, Feuchtner, Tiare, and Grønbæk, Kaj. 'XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–11.

[175] Ewing, Kate C, Fairclough, Stephen H, and Gilleade, Kiel. 'Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop.' In: *Frontiers in human neuroscience* 10 (2016), p. 223. DOI: 10.3389/fnhum.2016.00223.

[176] Eyraud, Robin, Zibetti, Elisabetta, and Baccino, Thierry. 'Allocation of visual attention while driving with simulated augmented reality.' In: *Transportation research part F: traffic psychology and behaviour* 32 (2015), pp. 46–55. DOI: 10.1016/j.trf.2015.04.011.

[177] Fairclough, Stephen. 'A closed-loop perspective on symbiotic human-computer interaction.' In: *International Workshop on Symbiotic Interaction*. Springer. 2015, pp. 57–67. DOI: 10.1007/978-3-319-24917-9_6.

[178] Fairclough, Stephen H. 'Fundamentals of physiological computing.' In: *Interacting with computers* 21.1-2 (2009), pp. 133–145. DOI: 10.1016/j.intcom.2008.10.011..

[179] Fairclough, Stephen H and Venables, Louise. 'Prediction of subjective states from psychophysiology: A multivariate approach.' In: *Biological psychology* 71.1 (2006), pp. 100–110. DOI: 10.1016/j.biopsycho.2005.03.007.

[180] Fairclough, Stephen H, Venables, Louise, and Tattersall, Andrew. 'The influence of task demand and learning on the psychophysiological response.' In: *International Journal of Psychophysiology* 56.2 (2005), pp. 171–184. DOI: 10.1016/j.ijpsycho.2004.11.003.

[181] Feick, Martin, Kleer, Niko, Tang, Anthony, and Krüger, Antonio. 'The Virtual Reality Questionnaire Toolkit.' In: UIST '20 Adjunct. Virtual Event, USA: ACM, 2020. DOI: 10.1145/3379350.3416188.

[182] Feld, Nico, Bimberg, Pauline, Weyers, Benjamin, and Zielasko, Daniel. 'Simple and Efficient? Evaluation of Transitions for Task-Driven Cross-Reality Experiences.' In: *IEEE Transactions on Visualization and Computer Graphics* (2024), pp. 1–18. DOI: 10.1109/TVCG.2024.3356949.

[183] Feng, Ying Xing, Tang, Tong Boon, and Ho, Eric Tatt Wei. 'Phasic Electrodermal Activity Indicates Changes in Workload and Affective States.' In: *2021 Int. Conf. on Intelligent Cybernetics Technology & Applications (ICICyTA)*. NY, NY, USA: IEEE, 2021, pp. 133–137. DOI: 10.1109/ICICyTA53712.2021.9689112.

[184] Fiedler, Patrique, Fonseca, Carlos, Supriyanto, Eko, Zanow, Frank, and Haueisen, Jens. 'A high-density 256-channel cap for dry electroencephalography.' In: *Human brain mapping* 43.4 (2022), pp. 1295–1308. DOI: 10.1002/hbm.25721.

[185] Findlater, Leah and Wobbrock, Jacob. 'Personalized Input: Improving Ten-Finger Touchscreen Typing through Automatic Adaptation.' In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (2012), pp. 815–824. DOI: 10.1145/2207676.2208520.

[186] Findlay, John M, Brown, Valerie, and Gilchrist, Iain D. 'Saccade target selection in visual search: The effect of information from the previous fixation.' In: *Vision research* 41.1 (2001), pp. 87–95.

[187] Finke, Andrea, Essig, Kai, Marchioro, Giuseppe, and Ritter, Helge. 'Toward FRP-based brain-machine interfaces—single-trial classification of fixation-related potentials.' In: *PloS one* 11.1 (2016), e0146848. DOI: 10.1371/journal.pone.0146848.

[188] Finseth, Tor, Dorneich, Michael C., Keren, Nir, Franke, Warren, Vardeman, Stephen, Segal, Jonathan, Deick, Andrew, Cavanah, Elizabeth, and Thompson, Kelly. 'The effectiveness of adaptive training for stress inoculation in a simulated astronaut task.' In: *Proc. of HFES* (2021). DOI: 10.1177/1071181321651241.

[189] Fitousi, Daniel and Wenger, Michael J. 'Processing capacity under perceptual and cognitive load: a closer look at load theory.' In: *Journal of experimental psychology: human perception and performance* 37.3 (2011), p. 781. DOI: 10.1037/a0020675.

[190] Fleck, Philipp, Sousa Calepso, Aimee, Hubenschmid, Sebastian, Sedlmair, Michael, and Schmalstieg, Dieter. 'RagRug: A Toolkit for Situated Analytics.' In: *IEEE Transactions on Visualization and Computer Graphics* (2022), pp. 1–1. DOI: 10.1109/TVCG.2022.3157058.

[191] Forschack, Norman, Gundlach, Christopher, Hillyard, Steven, and Müller, Matthias M. 'Dynamics of attentional allocation to targets and distractors during visual search.' In: *NeuroImage* 264 (2022), p. 119759. DOI: 10.1016/j.neuroimage.2022.119759.

[192] Frederiksen, Joakim Grant, Sørensen, Stine Maya Dreier, Konge, Lars, Svendsen, Morten Bo Søndergaard, Nobel-Jørgensen, Morten, Bjerrum, Flemming, and Andersen, Steven Arild Wuyts. 'Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: a randomized trial.' In: *Surgical endoscopy* 34 (2020), pp. 1244–1252.

[193] Friedman, Doron. 'Brain-computer interfacing and virtual reality.' In: *Handbook of digital games and entertainment technologies* (2015), pp. 1–22. DOI: 10.1007/978-981-4560-52-8_2-1.

[194] Fukuda, Keisuke and Vogel, Edward K. 'Individual differences in recovery time from attentional capture.' In: *Psychological science* 22.3 (2011), pp. 361–368. DOI: 10.1177/0956797611398493.

[195] Gal, Ran, Shapira, Lior, Ofek, Eyal, and Kohli, Pushmeet. 'FLARE: Fast layout for augmented reality applications.' In: *2014 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE. 2014, pp. 207–212.

[196] Gardony, Aaron L, Lindeman, Robert W, and Brunyé, Tad T. 'Eye-tracking for human-centered mixed reality: promises and challenges.' In: *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR)*. Vol. 11310. SPIE. 2020, pp. 230–247.

[197] Gaspelin, Nicholas, Lamy, Dominique, Egeth, Howard E., Liesefeld, Heinrich R., Kerzel, Dirk, Mandal, Ananya, Müller, Matthias M., Schall, Jeffrey D., Schubö, Anna, Slagter, Heleen A., Stilwell, Brad T., and Moorselaar, Dirk van. 'The Distractor Positivity Component and the Inhibition of Distracting Stimuli.' In: *Journal of Cognitive Neuroscience* 35.11 (2023), pp. 1693–1715. DOI: 10.1162/jocn_a_02051.

[198] Gavazzeni, Joachim, Wiens, Stefan, and Fischer, Håkan. 'Age effects to negative arousal differ for self-report and electrodermal activity.' In: *Psychophysiology* 45.1 (2008), pp. 148–151. DOI: 10.1111/j.1469-8986.2007.00596.x.

[199] Gavgani, Alireza Mazloumi, Nesbitt, Keith V, Blackmore, Karen L, and Nalivaiko, Eugene. 'Profiling subjective symptoms and autonomic changes associated with cybersickness.' In: *Autonomic Neuroscience* 203 (2017), pp. 41–50. DOI: 10.1016/j.autneu.2016.12.004.

[200] Gebhardt, Christoph, Hecox, Brian, Opheusden, Bas van, Wigdor, Daniel, Hillis, James, Hilliges, Otmar, and Benko, Hrvoje. 'Learning cooperative personalized policies from gaze data.' In: *Proceedings of the 32nd annual ACM symposium on user interface software and technology*. 2019, pp. 197–208.

[201] Gerup, Jaris, Soerensen, Camilla B, and Dieckmann, Peter. 'Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review.' In: *Int. journal of medical education* 11 (2020), p. 1.

[202] Gevins, Alan, Smith, Michael E, McEvoy, Linda, and Yu, Daphne. 'High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice.' In: *Cerebral cortex (New York, NY: 1991)* 7.4 (1997), pp. 374–385. DOI: 10.1093/cercor/7.4.374.

[203] Ghaffari, Roozbeh, Yang, Da Som, Kim, Joohee, Mansour, Amer, Wright Jr, John A, Model, Jeffrey B, Wright, Donald E, Rogers, John A, and Ray, Tyler R. 'State of Sweat: Emerging Wearable Systems for Real-Time, Noninvasive Sweat Sensing and Analytics.' In: *ACS sensors* (2021).

[204] Gherri, Elena and Eimer, Martin. 'Active listening impairs visual perception and selectivity: An ERP study of auditory dual-task costs on visual attention.' In: *Journal of cognitive neuroscience* 23.4 (2011), pp. 832–844. DOI: 10.1162/jocn.2010.21468.

[205] Giannakakis, Giorgos, Grigoriadis, Dimitris, Giannakaki, Katerina, Simantiraki, Olympia, Roniotis, Alexandros, and Tsiknakis, Manolis. 'Review on Psychological Stress Detection Using Biosignals.' In: *IEEE Transactions on Affective Computing* 13.1 (2022), pp. 440–460. DOI: 10.1109/TAFFC.2019.2927337.

[206] Giavarina, Davide. 'Understanding bland altman analysis.' In: *Biochemia medica* 25.2 (2015), pp. 141–151.

[207] Giles, Grace E., Mahoney, Caroline R., Brunyé, Tad T., Taylor, Holly A., and Kanarek, Robin B. 'Stress effects on mood, HPA axis, and autonomic response: comparison of three psychosocial stress paradigms.' In: *PloS one* 9.12 (2014), e113618. DOI: 10.1371/journal.pone.0113618.

[208] Gilroy, Stephen, Porteous, Julie, Charles, Fred, and Cavazza, Marc. 'Exploring passive user interaction for adaptive narratives.' In: *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*. 2012, pp. 119–128.

[209] Gimhae, Gyeong-Nam. 'Six human factors to acceptability of wearable computers.' In: *International Journal of Multimedia and Ubiquitous Engineering* 8.3 (2013), pp. 103–114.

[210] Gnacek, Michal, Broulidakis, John, Mavridou, Ifigeneia, Fatoorechi, Mohsen, Seiss, Ellen, Kostoulas, Theodoros, Balaguer-Ballester, Emili, Kiprijanovska, Ivana, Rosten, Claire, and Nduka, Charles. 'Emteqpro—fully integrated biometric sensing array for non-invasive biomedical research in virtual reality.' In: *Frontiers in virtual reality* 3 (2022), p. 781218.

[211] Gonçalves, Guilherme, Coelho, Hugo, Monteiro, Pedro, Melo, Miguel, and Bessa, Maximino. 'Systematic Review of Comparative Studies of the Impact of Realism in Immersive Virtual Experiences.' In: *ACM Comput. Surv.* 55.6 (2022), pp. 1–36. DOI: 10.1145/3533377.

[212] Gonzalez-Franco, Mar, Abdlkarim, Diar, Bhatia, Arpit, Macgregor, Stuart, Fotso-Puepi, Jason Alexander, Gonzalez, Eric J, Seifi, Hasti, Di Luca, Massimiliano, and Ahuja, Karan. 'Hovering Over the Key to Text Input in XR.' In: *arXiv preprint arXiv:2406.09579* (2024).

[213] Govaart, Gisela, Schettino, Antonio, Helbling, Saskia, Mehler, David, Ngiam, William Xiang Quan, Moreau, David, Chiossi, Francesco, Zanesco, Anthony, Yang, Yu-Fang, Gau, Remi, et al. 'EEG ERP preregistration template.' In: (2022).

[214] Gradl, Stefan, Wirth, Markus, Mächtlinger, Nico, Poguntke, Romina, Wonner, Andrea, Rohleder, Nicolas, and Eskofier, Bjoern M. 'The Stroop Room: A Virtual Reality-

Enhanced Stroop Test.' In: *25th ACM Symposium on Virtual Reality Software and Technology*. VRST '19. Parramatta, NSW, Australia: Association for Computing Machinery, 2019. DOI: 10.1145/3359996.3364247.

[215] Gramann, Klaus, Mckendrick, Ryan, Baldwin, Carryl, Roy, Raphaëlle, Jeunet, Camille, Mehta, Ranjana, and Vecchiato, Giovanni. 'Grand field challenges for cognitive neuroergonomics in the coming decade.' In: *Frontiers in Neuroergonomics* 2 (2021). DOI: 10.3389/fnrgo.2021.643969.

[216] Gramfort, Alexandre, Luessi, Martin, Larson, Eric, Engemann, Denis A, Strohmeier, Daniel, Brodbeck, Christian, Goj, Roman, Jas, Mainak, Brooks, Teon, and Parkkonen, Lauri. 'MEG and EEG data analysis with MNE-Python.' In: *Frontiers in neuroscience* 7 (2013), p. 267. DOI: 10.3389/fnins.2013.00267.

[217] Grant, Catharina C., Rensburg, Dina C. J. van, Strydom, Nina, and Viljoen, Margaretha. 'Importance of tachogram length and period of recording during noninvasive investigation of the autonomic nervous system.' In: *Annals of Noninvasive Electrocardiology* 16.2 (2011), pp. 131–139. DOI: 10.1111/j.1542-474X.2011.00422.x.

[218] Grasset, Raphael, Langlotz, Tobias, Kalkofen, Denis, Tatzgern, Markus, and Schmalstieg, Dieter. 'Image-driven view management for augmented reality browsers.' In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2012, pp. 177–186.

[219] Graves, Lee EF, Ridgers, Nicola D, Williams, Karen, Stratton, Gareth, Atkinson, Greg, and Cable, Nigel T. 'The physiological cost and enjoyment of Wii Fit in adolescents, young adults, and older adults.' In: *Journal of Physical Activity and Health* 7.3 (2010), pp. 393–401. DOI: 10.1123/jpah.7.3.393.

[220] Greco, Alberto, Valenza, Gaetano, Lanata, Antonio, Scilingo, Enzo Pasquale, and Citi, Luca. 'cvxEDA: A convex optimization approach to electrodermal activity processing.' In: *IEEE Transactions on Biomedical Engineering* 63.4 (2015), pp. 797–804.

[221] Greco, Alberto, Valenza, Gaetano, and Scilingo, Enzo Pasquale. 'Modeling for the Analysis of the EDA.' In: *Advances in Electrodermal Activity Processing with Applications for Mental Health: From Heuristic Methods to Convex Optimization*. Cham: Springer Int. Publishing, 2016, pp. 19–33. DOI: 10.1007/978-3-319-46705-4_2.

[222] Gruberger, Michal, Ben-Simon, Eti, Levkovitz, Yechiel, Zangen, Abraham, and Hendler, Talma. 'Towards a neuroscience of mind-wandering.' In: *Frontiers in human neuroscience* 5 (2011), p. 56.

[223] Grubert, Jens, Witzani, Lukas, Ofek, Eyal, Pahud, Michel, Kranz, Matthias, and Kristensson, Per Ola. 'Effects of Hand Representations for Typing in Virtual Reality.' In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 151–158. DOI: 10.1109/VR.2018.8446250.

[224] Gruenefeld, Uwe, Auda, Jonas, Mathis, Florian, Schneegass, Stefan, Khamis, Mohamed, Gugenheimer, Jan, and Mayer, Sven. 'VRception: Rapid Prototyping of Cross-Reality Systems in Virtual Reality.' In: *CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–15. DOI: 10.1145/3491102.3501821.

[225] Gruenefeld, Uwe, Stratmann, Tim Claudius, Ali, Abdallah El, Boll, Susanne, and Heuten, Wilko. 'RadialLight: Exploring Radial Peripheral LEDs for Directional Cues in Head-Mounted Displays.' In: *Proceedings of the 20th Int. Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '18. Barcelona, Spain: ACM, 2018. DOI: 10.1145/3229434.3229437.

[226] Guna, Jože, Geršak, Gregor, Humar, Iztok, Krebl, Maja, Orel, Marko, Lu, Huimin, and Pogačnik, Matevž. 'Virtual reality sickness and challenges behind different technology and content settings.' In: *Mobile Networks and Applications* 25.4 (2020), pp. 1436–1445. DOI: 10.1007/s11036-019-01373-w.

[227] Gürkök, Hayrettin and Nijholt, Anton. 'Brain–computer interfaces for multimodal interaction: a survey and principles.' In: *International Journal of Human-Computer Interaction* 28.5 (2012), pp. 292–307. DOI: 10.1080/10447318.2011.582022.

[228] Hadnett-Hunter, Jacob, O'Neill, Eamonn, and Proulx, Michael J. 'Contributed session ii: Visual search in virtual reality (vsvr): A visual search toolbox for virtual reality.' In: *Journal of Vision* 22.3 (2022), pp. 19–19. DOI: 10.1167/jov.22.3.19.

[229] Halem, Sjoerd van, Van Roekel, Eeske, Kroencke, Lara, Kuper, Niclas, and Denissen, Jaap. 'Moments that matter? On the complexity of using triggers based on skin conductance to sample arousing events within an experience sampling framework.' In: *European Journal of Personality* 34.5 (2020), pp. 794–807. DOI: 10.1002/per.2252.

[230] Hamilton, Patrick. 'Open source ECG analysis.' In: *Computers in Cardiology*. Memphis, TN, USA: IEEE, 2002, pp. 101–104. DOI: 10.1109/CIC.2002.1166717.

[231] Hammond, Hugo, Thomas, Graham, and Gilchrist, Iain D. 'Dynamically changing attention in complex visual stimuli.' In: *Journal of Vision* 22.14 (2022), pp. 3602–3602. DOI: https://doi.org/10.1167/jov.22.14.3602.

[232] Hancock, PA and Szalma, JL. 'The future of neuroergonomics.' In: *Theoretical Issues in Ergonomics Science* 4.1-2 (2003), pp. 238–249. DOI: 10.1080/1463922021000020927.

[233] Hancock, Peter A. and Warm, Joel S. 'A Dynamic Model of Stress and Sustained Attention.' In: *Human Factors Society Annual Meeting Proceedings*. Vol. 33. 1989, pp. 76–80.

[234] Hansen, Wilfred J. 'User Engineering Principles for Interactive Systems.' In: *Proceedings of the November 16-18, 1971, Fall Joint Computer Conference*. AFIPS '71 (Fall). Las Vegas,

Nevada: Association for Computing Machinery, 1972, pp. 523–532. DOI: 10.1145/1479064.1479159.

[235] Harmony, Thalía. 'The functional significance of delta oscillations in cognitive processing.' In: *Frontiers in integrative neuroscience* 7 (2013), p. 83.

[236] Harmony, Thalía, Fernández, Thalía, Silva, Juan, Bernal, Jorge, Díaz-Comas, Lourdes, Reyes, Alfonso, Marosi, Erzsébet, Rodríguez, Mario, and Rodríguez, Miguel. 'EEG delta activity: an indicator of attention to internal processing during performance of mental tasks.' In: *Int. journal of psychophysiology* 24.1-2 (1996), pp. 161–171. DOI: 10.1016/S0167-8760(96)00053-0.

[237] Hart, Sandra G. 'NASA-task load index (NASA-TLX); 20 years later.' In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage publications Sage CA: Los Angeles, CA. Los Angeles, CA: Sage publications Sage CA, 2006, pp. 904–908. DOI: 10.1177/154193120605000909.

[238] Hart, Sandra G and Staveland, Lowell E. 'Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research.' In: *Advances in psychology*. Vol. 52. Elsevier, 1988, pp. 139–183. DOI: 10.1016/S0166-4115(08)62386-9.

[239] Haslwanter, Thomas. 'Mathematics of three-dimensional eye rotations.' In: *Vision Research* 35.12 (1995), pp. 1727–1739. DOI: 10.1016/0042-6989(94)00257-M.

[240] Hassib, Mariam, Schneegass, Stefan, Eiglsperger, Philipp, Henze, Niels, Schmidt, Albrecht, and Alt, Florian. 'EngageMeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography.' In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 5114–5119. DOI: 10.1145/3025453.3025669.

[241] Haufe, Stefan, Kim, Jeong-Woo, Kim, Il-Hwa, Sonnleitner, Andreas, Schrauf, Michael, Curio, Gabriel, and Blankertz, Benjamin. 'Electrophysiology-based detection of emergency braking intention in real-world driving.' In: *Journal of Neural Engineering* 11.5 (2014), p. 056011.

[242] Hayashi, Takuto, Okamoto, Eika, Nishimura, Haruhiko, Mizuno-Matsumoto, Yuko, Ishii, Ryouhei, and Ukai, Satoshi. 'Beta activities in EEG associated with emotional stress.' In: *International Journal of Intelligent Computing in Medical Sciences & Image Processing* 3.1 (2009), pp. 57–68. DOI: 10.1080/1931308X.2009.10644171.

[243] Healey, Jennifer, Seger, Justin, and Picard, Rosalind. 'Quantifying driver stress: Developing a system for collecting and processing bio-metric signals in natural situations.' In: *Biomedical sciences instrumentation* 35 (1999), pp. 193–198.

[244] Hedges, Larry V. 'Effect sizes in cluster-randomized designs.' In: (2007). DOI: 10.3102/1076998606298.

[245] Herrmann, Christoph S, Strüber, Daniel, Helfrich, Randolph F, and Engel, Andreas K. 'EEG oscillations: from correlation to causality.' In: *International Journal of Psychophysiology* 103 (2016), pp. 12–21.

[246] Heuer, Anna and Schubö, Anna. 'Cueing distraction: Electrophysiological evidence for anticipatory active suppression of distractor location.' In: *Psychological research* 84.8 (2020), pp. 2111–2121.

[247] Hillebrand, Arjan, Tewarie, Prejaas, Van Dellen, Edwin, Yu, Meichen, Carbo, Ellen WS, Douw, Linda, Gouw, Alida A, Van Straaten, Elisabeth CW, and Stam, Cornelis J. 'Direction of information flow in large-scale resting-state networks is frequency-dependent.' In: *Proceedings of the National Academy of Sciences* 113.14 (2016), pp. 3867–3872.

[248] Hillstrom, Anne P. 'Repetition effects in visual search.' In: *Perception & psychophysics* 62 (2000), pp. 800–817. DOI: 10.3758/BF03206924.

[249] Hincapié, Mauricio, Caponio, Andrea, Rios, Horacio, and Mendívil, Eduardo González. 'An introduction to Augmented Reality with applications in aeronautical maintenance.' In: *2011 13th international conference on transparent optical networks*. New York, NY, USA: IEEE, 2011, pp. 1–4. DOI: 10.1109/ICTON.2011.5970856.

[250] Hincapié-Ramos, Juan David, Guo, Xiang, Moghadasian, Paymahn, and Irani, Pourang. 'Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pp. 1063–1072. DOI: 10.1145/2556288.2557130.

[251] Hirobe, Daichi, Uranishi, Yuki, Orlosky, Jason, Shirai, Shizuka, Ratsamee, Photchara, and Takemura, Haruo. 'Objective Measurements of Background Color Shifts Caused by Optical See-Through Head-Mounted Displays.' In: *2022 IEEE Int. Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. New York, NY, USA: IEEE, 2022, pp. 389–390. DOI: 10.1109/ISMAR-Adjunct57072.2022.00084.

[252] Hirsch, Linda, Müller, Florian, Chiossi, Francesco, Benga, Theodor, and Butz, Andreas Martin. 'My Heart Will Go On: Implicitly Increasing Social Connectedness by Visualizing Asynchronous Players' Heartbeats in VR Games.' In: *Proc. ACM Hum.-Comput. Interact.* 7.CHI PLAY (2023). DOI: 10.1145/3611057.

[253] Hoffman, James E and Subramaniam, Baskaran. 'The role of visual attention in saccadic eye movements.' In: *Perception & Psychophysics* 57.6 (1995), pp. 787–795.

[254] Hollingworth, Andrew and Bahle, Brett. 'Eye Tracking in Visual Search Experiments.' In: *Spatial Learning and Attention Guidance*. New York, NY: Springer US, 2020, pp. 23–35. DOI: 10.1007/7657_2019_30.

[255] Holmqvist, Kenneth, Nyström, Marcus, Andersson, Richard, Dewhurst, Richard, Jarodzka, Halszka, and Van de Weijer, Joost. Eye tracking: A comprehensive guide to methods and measures. oup Oxford, 2011.

[256] Höök, Kristina. 'Steps to take before intelligent user interfaces become real.' In: *Interacting with computers* 12.4 (2000), pp. 409–426.

[257] Hopfinger, Joseph B and West, Vicki M. 'Interactions between endogenous and exogenous attention on cortical visual processing.' In: *NeuroImage* 31.2 (2006), pp. 774–789.

[258] Hsieh, Liang-Tien and Ranganath, Charan. 'Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval.' In: *Neuroimage* 85 (2014), pp. 721–729. DOI: 10.1016/j.neuroimage.2013.08.003.

[259] Hu, Mandy X., Lamers, Femke, Geus, Eco J. C. de, and Penninx, Brenda WJH. 'Differential Autonomic Nervous System Reactivity in Depression and Anxiety During Stress Depending on Type of Stressor.' In: *Psychosomatic medicine* 78.5 (2016), pp. 562–572. DOI: 10.1097/PSY.0000000000000313.

[260] Huang, Ann, Knierim, Pascal, Chiossi, Francesco, Chuang, Lewis L, and Welsch, Robin. 'Proxemics for Human-Agent Interaction in Augmented Reality.' In: *CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022, pp. 1–13. DOI: 10.1145/3491102.3517593.

[261] Huang, Ann, Knierim, Pascal, Chiossi, Francesco, Chuang, Lewis L, and Welsch, Robin. 'Proxemics for Human-Agent Interaction in Augmented Reality.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3517593.

[262] Huang, Yi-Jheng, Liu, Kang-Yi, Lee, Suiang-Shyan, and Yeh, I-Cheng. 'Evaluation of a hybrid of hand gesture and controller inputs in virtual reality.' In: *International Journal of Human–Computer Interaction* 37.2 (2021), pp. 169–180. DOI: 10.1080/10447318.2020.1809248.

[263] Huang, Wen, Roscoe, Rod D, Johnson-Glenberg, Mina C, and Craig, Scotty D. 'Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion.' In: *Journal of Computer Assisted Learning* 37.3 (2021), pp. 745–758. DOI: 10.1111/jcal.12520.

[264] Hubenschmid, Sebastian, Wieland, Jonathan, Fink, Daniel Immanuel, Batch, Andrea, Zagermann, Johannes, Elmqvist, Niklas, and Reiterer, Harald. 'ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: ACM, 2022, pp. 1–20. DOI: 10.1145/3491102.3517550.

[265] Hubenschmid, Sebastian, Zagermann, Johannes, Butscher, Simon, and Reiterer, Harald. 'STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. DOI: 10.1145/3411764.3445298.

[266] Hughes, Charles E, Stapleton, Christopher B, Hughes, Darin E, and Smith, Eileen M. 'Mixed reality in education, entertainment, and training.' In: *IEEE computer graphics and applications* 25.6 (2005), pp. 24–30.

[267] Hulleman, Johan. 'Quantitative and qualitative differences in the top-down guiding attributes of visual search.' In: *Journal of Experimental Psychology: Human Perception and Performance* 46.9 (2020), p. 942. DOI: 10.1037/xhp0000764.

[268] Hussain, Muhammad and Park, Jaehyun. 'Effect of Transparency Levels and Real-World Backgrounds on the User Interface in Augmented Reality Environments.' In: *Int. Journal of Human–Computer Interaction* (2023), pp. 1–10. DOI: 10.1080/10447318.2023.2212218.

[269] Hutmacher, Fabian. 'Why is there so much more research on vision than on any other sensory modality?' In: *Frontiers in psychology* 10 (2019), p. 2246. DOI: 10.3389/fpsyg.2019.02246.

[270] IJsselsteijn, Wijnand A, Kort, Yvonne AW de, and Poels, Karolien. 'The game experience questionnaire.' In: *Eindhoven: Technische Universiteit Eindhoven* 46.1 (2013).

[271] Ionescu, Radu Tudor, Alexe, Bogdan, Leordeanu, Marius, Popescu, Marius, Papadopoulos, Dim P., and Ferrari, Vittorio. 'How Hard Can It Be? Estimating the Difficulty of Visual Search in an Image.' In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 2157–2166. DOI: 10.1109/CVPR.2016.237.

[272] Itti, Laurent. 'Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes.' In: *Visual Cognition* 12.6 (2005), pp. 1093–1123.

[273] Jackson, Kenneth M, Shaw, Tyler H, and Helton, William S. 'The effects of dual-task interference on visual search and verbal memory.' In: *Ergonomics* 66.1 (2023), pp. 125–135.

[274] Jacob, Robert J.K. and Karn, Keith S. 'Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises.' In: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (2003).

[275] Jaeggi, Susanne M., Buschkuehl, Martin, Perrig, Walter J, and Meier, Beat. 'The concurrent validity of the N-back task as a working memory measure.' In: *Memory* 18.4 (2010), pp. 394–412. DOI: 10.1080/09658211003702171.

[276] Jaimes, Alejandro and Sebe, Nicu. ‘Multimodal human–computer interaction: A survey.’ In: *Computer vision and image understanding* 108.1-2 (2007), pp. 116–134. DOI: 10.1016/j.cviu.2006.10.019.

[277] Jannati, Ali, Gaspar, John M, and McDonald, John J. ‘Tracking target and distractor processing in fixed-feature visual search: evidence from human electrophysiology.’ In: *Journal of Experimental Psychology: Human Perception and Performance* 39.6 (2013), p. 1713.

[278] Janssen, Tieme WP, Grammer, Jennie K, Bleichner, Martin G, Bulgarelli, Chiara, Davidesco, Ido, Dikker, Suzanne, Jasińska, Kaja K, Siugzdaite, Roma, Vassena, Eliana, Vatakis, Argiro, et al. ‘Opportunities and limitations of mobile neuroimaging technologies in educational neuroscience.’ In: *Mind, Brain, and Education* 15.4 (2021), pp. 354–370. DOI: 10.1111/mbe.12302.

[279] Jarvis, Jan, Putze, Felix, Heger, Dominic, and Schultz, Tanja. ‘Multimodal person independent recognition of workload related biosignal patterns.’ In: *Proceedings of the 13th International Conference on Multimodal Interfaces*. ICMI ’11. Alicante, Spain: Association for Computing Machinery, 2011, pp. 205–208. DOI: 10.1145/2070481.2070516.

[280] 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: j.neuroimage.2017.06.030.

[281] Jayawardena, Gavindya. ‘Raemap: Real-time advanced eye movements analysis pipeline.’ In: *ACM ETRA*. 2020, pp. 1–4.

[282] Jensen, Ole, Gelfand, Jack, Kounios, John, and Lisman, John E. ‘Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task.’ In: *Cerebral cortex* 12.8 (2002), pp. 877–882. DOI: 10.1093/cercor/12.8.877.

[283] Jensen, Ole and Tesche, Claudia D. ‘Frontal theta activity in humans increases with memory load in a working memory task.’ In: *European journal of Neuroscience* 15.8 (2002), pp. 1395–1399.

[284] Jetter, Hans-Christian, Schröder, Jan-Henrik, Gugenheimer, Jan, Billinghurst, Mark, Anthes, Christoph, Khamis, Mohamed, and Feuchtner, Tiare. ‘Transitional Interfaces in Mixed and Cross-Reality: A New Frontier?’ In: *Proc. of ISS*. ISS ’21. ACM, 2021. DOI: 10.1145/3447932.3487940.

[285] Jha, Gouri, Sharma, Lavanya shm, and Gupta, Shailja. ‘Future of augmented reality in healthcare department.’ In: *Proceedings of Second Int. Conference on Computing, Communications, and Cyber-Security: IC4S 2020*. Springer. 2021, pp. 667–678. DOI: 10.1007/978-981-16-0733-2_47.

[286] Jiang, Weiwei, Yang, Kangning, Windl, Maximiliane, Chiossi, Francesco, Tag, Benjamin, Mayer, Sven, and Sarsenbayeva, Zhanna. 'Current Challenges of Using Wearable Devices for Online Emotion Sensing.' In: *arXiv preprint arXiv:2208.05206* (30, 2022). DOI: 10.48550/arXiv.2208.05206.

[287] Jiang, Yuying, Zhang, Haoran, and Yu, Shan. 'Changes in Delta and Theta Oscillations in the Brain Indicate Dynamic Switching of Attention between Internal and External Processing.' In: *4th International Conference on Biometric Engineering and Applications*. ICBEA '21. Taiyuan, China: ACM, 2021, pp. 25–31. DOI: 10.1145/3476779.3476784.

[288] Jr Hosmer, David W, Lemeshow, Stanley, and Sturdivant, RX. 'Applied logistic regression.' In: *Hoboken: John Wiley & Sons* (2004). DOI: 10.1002/9781118548387.

[289] Kajimura, Shogo and Nomura, Michio. 'When we cannot speak: Eye contact disrupts resources available to cognitive control processes during verb generation.' In: *Cognition* 157 (2016), pp. 352–357.

[290] Kamienkowski, Juan E, Ison, Matias J, Quiroga, Rodrigo Quian, and Sigman, Mariano. 'Fixation-related potentials in visual search: A combined EEG and eye tracking study.' In: *Journal of vision* 12.7 (2012), pp. 4–4. DOI: https://doi.org/10.1167/12.7.4.

[291] Kamkuimo K., Sorelle Audrey, Girard, Benoît, and Menelas, Bob-Antoine J. 'Dynamic Difficulty Adjustment Through Real-Time Physiological Feedback for a More Adapted Virtual Reality Exposure Therapy.' In: *Games and Learning Alliance*. Cham: Springer International Publishing, 2020, pp. 102–111. DOI: 10.1007/978-3-030-63464-3_10.

[292] Kappeler-Setz, Cornelia, Gravenhorst, Franz, Schumm, Johannes, Arnrich, Bert, and Tröster, Gerhard. 'Towards long term monitoring of electrodermal activity in daily life.' In: *Personal and ubiquitous computing* 17.2 (2013), pp. 261–271. DOI: 10.1007/s00779-011-0463-4.

[293] Karahanoğlu, Armağan and Erbuğ, Çiğdem. 'Perceived Qualities of Smart Wearables: Determinants of User Acceptance.' In: *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces*. DPPI '11. Milano, Italy: Association for Computing Machinery, 2011. DOI: 10.1145/2347504.2347533.

[294] Karakaş, Sirel. 'A review of theta oscillation and its functional correlates.' In: *International Journal of Psychophysiology* 157 (2020), pp. 82–99.

[295] Karbach, Julia and Verhaeghen, Paul. 'Making working memory work: A meta-analysis of executive-control and working memory training in older adults.' In: *Psychological science* 25.11 (2014), pp. 2027–2037.

[296] Karolus, Jakob, Schuff, Hendrik, Kosch, Thomas, Wozniak, Paweł W., and Schmidt, Albrecht. 'EMGuitar: Assisting Guitar Playing with Electromyography.' In: *Proceedings of*

the 2018 Designing Interactive Systems Conference. DIS '18. Hong Kong, China: Association for Computing Machinery, 2018, pp. 651–655. DOI: 10.1145/3196709.3196803.

[297] Karolus, Jakob, Wozniak, Paweł W., Chuang, Lewis L., and Schmidt, Albrecht. 'Robust Gaze Features for Enabling Language Proficiency Awareness.' In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 2998–3010. DOI: 10.1145/3025453.3025601.

[298] Karrach, Ladislav, Pivarčiová, Elena, and Božek, Pavol. 'Identification of QR code perspective distortion based on edge directions and edge projections analysis.' In: *Journal of imaging* 6.7 (2020), p. 67. DOI: 10.3390/jimaging6070067.

[299] Karwowski, Waldemar. 'A Review of Human Factors Challenges of Complex Adaptive Systems: Discovering and Understanding Chaos in Human Performance.' In: *Human Factors* 54.6 (2012), pp. 983–995. DOI: 10.1177/0018720812467459. eprint: https://doi.org/10.1177/0018720812467459.

[300] Kaunitz, Lisandro N, Kamienkowski, Juan E, Varatharajah, Alexander, Sigman, Mariano, Quiroga, Rodrigo Quian, and Ison, Matias J. 'Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task.' In: *NeuroImage* 89.7 (2014), pp. 297–305. DOI: 10.1167/12.7.4. eprint: https://arvojournals.org/arvo/content_public/journal/jov/933492/i1534-7362-12-7-4.pdf.

[301] Kavakli, Manolya, Kartiko, Iwan, Porte, John, and Bigoin, Nolwenn. 'Effects of digital content on motion sickness in immersive virtual environments.' In: *3rd Int. Conf. on Computer Science & Information Systems, July*. 2008, pp. 23–24.

[302] Keim, Daniel, Andrienko, Gennady, Fekete, Jean-Daniel, Gorg, Carsten, Kohlhammer, Jorn, and Melançon, Guy. 'Visual analytics: Definition, process, and challenges.' In: *Lecture notes in computer science* 4950 (2008), pp. 154–176.

[303] Kenyon, Robert V and Ellis, Stephen R. 'Vision, perception, and object manipulation in virtual environments.' In: *Virtual reality for physical and motor rehabilitation*. Springer, 2014, pp. 47–70.

[304] Keshavarz, Behrang and Hecht, Heiko. 'Validating an efficient method to quantify motion sickness.' In: *Human factors* 53.4 (2011), pp. 415–426.

[305] Keynan, Jackob N, Cohen, Avihay, Jackont, Gilan, Green, Nili, Goldway, Noam, Davidov, Alexander, Meir-Hasson, Yehudit, Raz, Gal, Intrator, Nathan, Fruchter, Eyal, et al. 'Electrical fingerprint of the amygdala guides neurofeedback training for stress resilience.' In: *Nature human behaviour* 3.1 (2019), pp. 63–73. DOI: 10.1038/s41562-018-0484-3.

[306] Khalid, Muhammad Bilal, Rao, Naveed Iqbal, Rizwan-i-Haque, Intisar, Munir, Sarmad, and Tahir, Farhan. 'Towards a brain computer interface using wavelet transform with averaged and time segmented adapted wavelets.' In: *2009 2nd international conference on computer, control and communication*. IEEE. 2009, pp. 1–4.

[307] Kiefer, Peter, Giannopoulos, Ioannis, Duchowski, Andrew, and Raubal, Martin. 'Measuring Cognitive Load for Map Tasks Through Pupil Diameter.' In: *Geographic Information Science*. Ed. by Jennifer A. Miller, David O'Sullivan, and Nancy Wiegand. Cham: Springer Int. Publishing, 2016, pp. 323–337. DOI: 10.1007/978-3-319-45738-3_21.

[308] Kirchner, Wayne K. 'Age differences in short-term retention of rapidly changing information.' In: *Journal of experimental psychology* 55.4 (1958), p. 352.

[309] Klarkowski, Madison, Johnson, Daniel, Wyeth, Peta, Phillips, Cody, and Smith, Simon. 'Don't Sweat the Small Stuff: The Effect of Challenge-Skill Manipulation on Electrodermal Activity.' In: *Proceedings of the 2018 Annual Symp. on Computer-Human Interaction in Play*. CHI PLAY '18. Melbourne, VIC, Australia: ACM, 2018, pp. 231–242. DOI: 10.1145/3242671.3242714.

[310] Kleckner, Ian, Wormwood, Jolie B, Jones, Rebecca M, Siegel, Erika, Culakova, Eva, Heathers, James, Barrett, Lisa Feldman, Lord, Catherine, Quigley, Karen, and Goodwin, Matthew. 'Adaptive thresholding increases ability to detect changes in rate of skin conductance responses to psychologically arousing stimuli.' In: (2021). DOI: 10.31234/osf.io/b4agz.

[311] Klimesch, Wolfgang. 'Alpha-band oscillations, attention, and controlled access to stored information.' In: *Trends in cognitive sciences* 16.12 (2012), pp. 606–617. DOI: 10.1016/j.tics.2012.10.007.

[312] Klimesch, Wolfgang. 'EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis.' In: *Brain research reviews* 29.2-3 (1999), pp. 169–195.

[313] Knierim, Pascal, Kosch, Thomas, and Schmidt, Albrecht. 'The Nomadic Office: A Location Independent Workspace Through Mixed Reality.' In: *IEEE Pervasive Computing* 20.4 (2021), pp. 71–78. DOI: 10.1109/MPRV.2021.3119378.

[314] Knyazev, Gennady G. 'EEG delta oscillations as a correlate of basic homeostatic and motivational processes.' In: *Neuroscience & Biobehavioral Reviews* 36.1 (2012), pp. 677–695.

[315] Kohlmorgen, Jens, Dornhege, Guido, Braun, Mikio, Blankertz, Benjamin, Müller, Klaus-Robert, Curio, Gabriel, Hagemann, Konrad, Bruns, Andreas, Schrauf, Michael, Kincses, Wilhelm, et al. 'Improving human performance in a real operating envi-

ronment through real-time mental workload detection.' In: *Toward brain-computer interfacing* 409422 (2007), pp. 409–422.

[316] Kosch, Thomas, Funk, Markus, Schmidt, Albrecht, and Chuang, Lewis L. 'Identifying Cognitive Assistance with Mobile Electroencephalography: A Case Study with In-Situ Projections for Manual Assembly.' In: *Proc. ACM Hum.-Comput. Interact.* 2.EICS (2018). DOI: 10.1145/3229093.

[317] Kosch, Thomas, Karolus, Jakob, Ha, Havy, and Schmidt, Albrecht. 'Your Skin Resists: Exploring Electrodermal Activity As Workload Indicator During Manual Assembly.' In: *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS '19. Valencia, Spain: ACM, 2019, 8:1–8:5. DOI: 10.1145/3319499.3328230.

[318] 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.

[319] Kosch, Thomas, Schmidt, Albrecht, Thanheiser, Simon, and Chuang, Lewis L. 'One Does Not Simply RSVP: Mental Workload to Select Speed Reading Parameters Using Electroencephalography.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, pp. 1–13. DOI: 10.1145/3313831.3376766.

[320] Kosch, Thomas, Welsch, Robin, Chuang, Lewis, and Schmidt, Albrecht. 'The Placebo Effect of Artificial Intelligence in Human-Computer Interaction.' In: *ACM Trans. Comput.-Hum. Interact.* (2022). DOI: 10.1145/3529225.

[321] Koulieris, George Alex, Akşit, Kaan, Stengel, Michael, Mantiuk, Rafał K, Mania, Katerina, and Richardt, Christian. 'Near-eye display and tracking technologies for virtual and augmented reality.' In: *Computer Graphics Forum*. Vol. 38. Hoboken, New Jersey, USA: Wiley Online Library, 2019, pp. 493–519. DOI: 10.1111/cgf.13654.

[322] Kovacs, Balint Istvan, Erb, Ingrid, Kaufmann, Hannes, and Ferschin, Peter. 'MR. Sketch. Immediate 3D Sketching via Mixed Reality Drawing Canvases.' In: *2023 IEEE ISMAR*. IEEE. 2023, pp. 10–19.

[323] Krakowczyk, Daniel G., Reich, David R., Chwastek, Jakob, Jakobi, Deborah N., Prasse, Paul, Süss, Assunta, Turuta, Oleksii, Kasprowski, Paweł, and Jäger, Lena A. 'Pymovements: A Python Package for Eye Movement Data Processing.' In: *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. ETRA '23. Tubingen, Germany: ACM, 2023. DOI: 10.1145/3588015.3590134.

[324] Kramer, Daniel. 'Predictions of performance by EEG and skin conductance.' In: *Indiana undergraduate journal of cognitive science* 2 (2007), pp. 3–13.

[325] Krämer, Julia, Chiossi, Francesco, and Kosch, Thomas. 'Escaping the Filter Bubble: Evaluating Electroencephalographic Theta Band Synchronization as Indicator for Selective Exposure in Online News Reading.' In: *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '25)*. Yokohama, Japan: ACM, 2025. DOI: 10.1145/3706599.3720078.

[326] Kristensson, Per Ola and Vertanen, Keith. 'Performance comparisons of phrase sets and presentation styles for text entry evaluations.' In: *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*. IUI '12. Lisbon, Portugal: Association for Computing Machinery, 2012, pp. 29–32. DOI: 10.1145/2166966.2166972.

[327] Kristjánsson, Arni. 'Simultaneous priming along multiple feature dimensions in a visual search task.' In: *Vision research* 46.16 (2006), pp. 2554–2570. DOI: 10.1016/j.visres.2006.01.015.

[328] Kuutti, Kari and Bannon, Liam J. 'The turn to practice in HCI: towards a research agenda.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, pp. 3543–3552. DOI: 10.1145/2556288.2557111.

[329] Kuznetsova, Alexandra, Bruun Brockhoff, Per, and Haubo Bojesen Christensen, Rune. lmerTest: Tests in Linear Mixed Effects Models. 2020.

[330] Labonte-Lemoyne, Elise, Courtemanche, François, Louis, Victoire, Fredette, Marc, Sénécal, Sylvain, and Léger, Pierre-Majorique. 'Dynamic threshold selection for a biocybernetic loop in an adaptive video game context.' In: *Frontiers in human neuroscience* 12 (2018), p. 282. DOI: 10.3389/fnhum.2018.00282.

[331] Laeng, Bruno and Alnaes, Dag. 'Pupillometry.' In: *Eye Movement Research: An Introduction to its Scientific Foundations and Applications* (2019), pp. 449–502.

[332] Lages, Wallace S and Bowman, Doug A. 'Walking with adaptive augmented reality workspaces: design and usage patterns.' In: *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 2019, pp. 356–366.

[333] Laufer, László and Németh, Bottyán. 'Predicting User Action from Skin Conductance.' In: *Proceedings of the 13th International Conference on Intelligent User Interfaces*. IUI '08. Gran Canaria, Spain: Association for Computing Machinery, 2008, pp. 357–360. DOI: 10.1145/1378773.1378829.

[334] Lavie, Nilli. 'Perceptual load as a necessary condition for selective attention.' In: *Journal of Experimental Psychology: Human perception and performance* 21.3 (1995), p. 451. DOI: 10.1037/0096-1523.21.3.451.

[335] Lavie, Nilli, Ro, Tony, and Russell, Charlotte. 'The role of perceptual load in processing distractor faces.' In: *Psychological science* 14.5 (2003), pp. 510–515. DOI: 10.1111/1467-9280.03453.

[336] Law, Effie L.-C., Brühlmann, Florian, and Mekler, Elisa D. 'Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice.' In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '18. Melbourne, VIC, Australia: Association for Computing Machinery, 2018, pp. 257–270. DOI: 10.1145/3242671.3242683.

[337] Law, Effie L.-C., Brühlmann, Florian, and Mekler, Elisa D. 'Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice.' In: *Proc. of the 2018 CHIPLAY*. CHI PLAY '18. 2018, pp. 257–270. DOI: 10.1145/3242671.3242683.

[338] Le, Huy Viet, Mayer, Sven, Wolf, Katrin, and Henze, Niels. 'Finger Placement and Hand Grasp during Smartphone Interaction.' In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 2576–2584. DOI: 10.1145/2851581.2892462.

[339] Lee, Cha, Rincon, Gustavo A, Meyer, Greg, Höllerer, Tobias, and Bowman, Doug A. 'The effects of visual realism on search tasks in mixed reality simulation.' In: *IEEE TVCG* 19.4 (2013), pp. 547–556.

[340] Lee, Jaewook, Jin, Fanjie, Kim, Younsoo, and Lindlbauer, David. 'User Preference for Navigation Instructions in Mixed Reality.' In: *2022 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR)*. NY, NY, USA: IEEE, 2022, pp. 802–811. DOI: 10.1109/VR51125.2022.00102.

[341] Lee, Lik-Hang, Braud, Tristan, Zhou, Pengyuan, Wang, Lin, Xu, Dianlei, Lin, Zijun, Kumar, Abhishek, Bermejo, Carlos, and Hui, Pan. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. 2021.

[342] Lee, Te-Won, Girolami, Mark, and Sejnowski, Terrence J. 'Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources.' In: *Neural computation* 11.2 (1999), pp. 417–441.

[343] Lee, YB, Yoon, SW, Lee, CK, and Lee, MH. 'Wearable EDA sensor gloves using conducting fabric and embedded system.' In: *World Congress on Medical Physics and Biomedical Engineering 2006*. Springer. 2007, pp. 883–888.

[344] Lewis, James R. 'The system usability scale: past, present, and future.' In: *International Journal of Human–Computer Interaction* 34.7 (2018), pp. 577–590. DOI: 10.1080/10447318.2018.1455307.

[345] Li, Jie, Vinayagamoorthy, Vinoba, Williamson, Julie, Shamma, David A., and Cesar, Pablo. 'Social VR: A New Medium for Remote Communication and Collaboration.' In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411763.3441346.

[346] Li, Yue, Zheng, Lu, and Wang, Xuewen. 'Flexible and wearable healthcare sensors for visual reality health-monitoring.' In: *Virtual Reality & Intelligent Hardware* 1.4 (2019), pp. 411–427. DOI: 10.1016/j.vrih.2019.08.001.

[347] Liesefeld, Heinrich René, Liesefeld, Anna Marie, Töllner, Thomas, and Müller, Hermann J. 'Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG.' In: *NeuroImage* 156 (2017), pp. 166–173.

[348] Lietz, Petra. 'Research into questionnaire design: A summary of the literature.' In: *Int. journal of market research* 52.2 (2010), pp. 249–272. DOI: 10.2501/S147078530920120.

[349] Lim, CL, Barry, RJ, Gordon, E, Sawant, A, Rennie, C, and Yiannikas, C. 'The relationship between quantified EEG and skin conductance level.' In: *Int. Journal of Psychophysiology* 21.2-3 (1996), pp. 151–162. DOI: 10.1016/0167-8760(95)00049-6.

[350] Lin, Wan-Hua, Wu, Dan, Li, Chunyue, Zhang, Heye, and Zhang, Yuan-Ting. 'Comparison of heart rate variability from PPG with that from ECG.' In: *The international conference on health informatics*. Cham, Switzerland: Springer, 2014, pp. 213–215. DOI: 10.1007/978-3-319-03005-0_54.

[351] Lindlbauer, David. 'The future of mixed reality is adaptive.' In: *XRDS: Crossroads, The ACM Magazine for Students* 29.1 (2022), pp. 26–31.

[352] Lindlbauer, David, Feit, Anna Maria, and Hilliges, Otmar. 'Context-aware online adaptation of mixed reality interfaces.' In: *Proceedings of the 32nd annual ACM symposium on UIST*. 2019, pp. 147–160. DOI: 10.1145/3332165.3347945.

[353] Llobera, Joan, Spanlang, Bernhard, Ruffini, Giulio, and Slater, Mel. 'Proxemics with Multiple Dynamic Characters in an Immersive Virtual Environment.' In: *ACM Trans. Appl. Percept.* 8.1 (2010), Article 3. DOI: 10.1145/1857893.1857896.

[354] Lo, Steson and Andrews, Sally. 'To transform or not to transform: Using generalized linear mixed models to analyse reaction time data.' In: *Frontiers in psychology* 6 (2015), p. 1171. DOI: 10.3389/fpsyg.2015.01171.

[355] Long, Xingyu, Mayer, Sven, and Chiossi, Francesco. 'Multimodal Detection of External and Internal Attention in Virtual Reality using EEG and Eye Tracking Features.' In: *Proceedings of Mensch Und Computer 2024*. MuC '24. Karlsruhe, Germany: Association for Computing Machinery, 2024, pp. 29–43. DOI: 10.1145/3670653.3670657.

[356] Lorenc, Elizabeth S, Mallett, Remington, and Lewis-Peacock, Jarrod A. 'Distraction in visual working memory: Resistance is not futile.' In: *Trends in cognitive sciences* 25.3 (2021), pp. 228–239. DOI: 10.1016/j.tics.2020.12.004.

[357] Lotte, Fabien, Bougrain, Laurent, Cichocki, Andrzej, Clerc, Maureen, Congedo, Marco, Rakotomamonjy, Alain, and Yger, Florian. 'A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update.' In: *Journal of neural engineering* 15.3 (2018), p. 031005.

[358] Luck, Steven J. An introduction to the event-related potential technique. MIT press, 2014.

[359] Luo, Weizhou, Lehmann, Anke, Widengren, Hjalmar, and Dachselt, Raimund. 'Where should we put it? layout and placement strategies of documents in augmented reality for collaborative sensemaking.' In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–16.

[360] Luong, Tiffany, Argelaguet, Ferran, Martin, Nicolas, and Lécuyer, Anatole. 'Introducing mental workload assessment for the design of virtual reality training scenarios.' In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. New York, NY, USA: IEEE, 2020, pp. 662–671. DOI: 10.1109/VR46266.2020.00089.

[361] Luong, Tiffany, Martin, Nicolas, Raison, Anais, Argelaguet, Ferran, Diverrez, Jean-Marc, and Lécuyer, Anatole. 'Towards real-time recognition of users mental workload using integrated physiological sensors into a VR HMD.' In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. New York, NY, USA: IEEE, 2020, pp. 425–437. DOI: 10.1109/ISMAR50242.2020.00068.

[362] Lykken, David T and Venables, Peter H. 'Direct measurement of skin conductance: A proposal for standardization.' In: *Psychophysiology* 8.5 (1971), pp. 656–672. DOI: 10.1111/j.1469-8986.1971.tb00501.x.

[363] MacInnes, W Joseph, Jóhannesson, Ómar I, Chetverikov, Andrey, and Kristjánsson, Árni. 'No advantage for separating overt and covert attention in visual search.' In: *Vision* 4.2 (2020), p. 28.

[364] Magosso, Elisa, De Crescenzio, Francesca, Ricci, Giulia, Piastra, Sergio, and Ursino, Mauro. 'EEG alpha power is modulated by attentional changes during cognitive tasks and virtual reality immersion.' In: *Comp. Intelligence and Neuroscience* 2019 (2019). DOI: 10.1155/2019/7051079.

[365] Magosso, Elisa, Ricci, Giulia, and Ursino, Mauro. 'Alpha and theta mechanisms operating in internal-external attention competition.' In: *Journal of Integrative Neuroscience* 20.1 (2021), pp. 1–19. DOI: 10.31083/j.jin.2021.01.422.

[366] Makowski, Dominique, Lüdecke, Daniel, Ben-Shachar, Mattan S., and Patil, Indrajeet. report: Automated Reporting of Results and Statistical Models. 2021.

[367] Makowski, Dominique, Pham, Tam, Lau, Zen J., Brammer, Jan C., Lespinasse, François, Pham, Hung, Schölzel, Christopher, and Chen, SH. 'NeuroKit2: A Python toolbox for neurophysiological signal processing.' In: *Behavior research methods* 53.4 (2021), pp. 1689–1696. DOI: 10.3758/s13428-020-01516-y.

[368] Makransky, Guido, Terkildsen, Thomas S, and Mayer, Richard E. 'Adding immersive virtual reality to a science lab simulation causes more presence but less learning.' In: *Learning and Instruction* 60 (2019), pp. 225–236.

[369] Maljkovic, Vera and Nakayama, Ken. 'Priming of pop-out: I. Role of features.' In: *Memory & cognition* 22.6 (1994), pp. 657–672. DOI: 10.3758/BF03209251.

[370] Mantha, Srinivas, Roizen, Michael F, Fleisher, Lee A, Thisted, Ronald, and Foss, Joseph. 'Comparing methods of clinical measurement: reporting standards for Bland and Altman analysis.' In: *Anesthesia & Analgesia* 90.3 (2000), pp. 593–602. DOI: 10.1097/00000539-200003000-00018.

[371] Marciano, Hadas and Yeshurun, Yaffa. 'Large inter-individual and intra-individual variability in the effect of perceptual load.' In: *PLoS One* 12.4 (2017), e0175060. DOI: 10.1371/journal.pone.0175060.

[372] Marín-Morales, Javier, Llinares, Carmen, Guixeres, Jaime, and Alcañiz, Mariano. 'Emotion recognition in immersive virtual reality: From statistics to affective computing.' In: *Sensors* 20.18 (2020), p. 5163. DOI: 10.3390/s20185163.

[373] Marini, Francesco, Breeding, Katherine A, and Snow, Jacqueline C. 'Distinct visuomotor brain dynamics for real-world objects versus planar images.' In: *Neuroimage* 195 (2019), pp. 232–242. DOI: 10.1016/j.neuroimage.2019.02.026.

[374] Marucci, Matteo, Di Flumeri, Gianluca, Borghini, Gianluca, Sciaraffa, Nicolina, Scandola, Michele, Pavone, Enea Francesco, Babiloni, Fabio, Betti, Viviana, and Aricò, Pietro. 'The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence.' In: *Scientific Reports* 11.1 (2021), pp. 1–15.

[375] Matsuda, Nathan, Chapiro, Alex, Zhao, Yang, Smith, Clinton, Bachy, Romain, and Lanman, Douglas. 'Realistic luminance in VR.' In: *SIGGRAPH Asia 2022 Conference Papers*. 2022, pp. 1–8. DOI: 10.1145/3550469.3555427.

[376] Matthews, G., Davies, D.R., Westerman, S.J., and Stammers, R.B. Human performance: Cognition, stress, and individual differences. Psychology Press, 2002.

[377] Mayer, Sven, Schwind, Valentin, Schweigert, Robin, and Henze, Niels. 'The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. DOI: 10.1145/3173574.3174227.

[378] McAtamney, Lynn and Corlett, E Nigel. 'RULA: a survey method for the investigation of work-related upper limb disorders.' In: *Applied Ergonomics* 24.2 (1993), pp. 91–99.

[379] McCarthy, Cameron, Pradhan, Nikhilesh, Redpath, Calum, and Adler, Andy. 'Validation of the Empatica E4 wristband.' In: *2016 IEEE EMBS International Student Conference (ISC)*. 2016, pp. 1–4. DOI: 10.1109/EMBSISC.2016.7508621.

[380] McGill, Mark, Boland, Daniel, Murray-Smith, Roderick, and Brewster, Stephen. 'A dose of reality: Overcoming usability challenges in vr head-mounted displays.' In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015, pp. 2143–2152. DOI: 10.1145/2702123.2702382.

[381] McGill, Mark, Brewster, Stephen, De Sa Medeiros, Daniel Pires, Bovet, Sidney, Gutierrez, Mario, and Kehoe, Aidan. 'Creating and Augmenting Keyboards for Extended Reality with the Keyboard Augmentation Toolkit.' In: *ACM Trans. Comput.-Hum. Interact.* 29.2 (2022). DOI: 10.1145/3490495.

[382] McMillan, Kathryn M, Laird, Angela R, Witt, Suzanne T, and Meyerand, M Elizabeth. 'Self-paced working memory: Validation of verbal variations of the n-back paradigm.' In: *Brain research* 1139 (2007), pp. 133–142. DOI: 10.1016/j.brainres.2006.12.058.

[383] McNab, Fiona and Dolan, Raymond J. 'Dissociating distractor-filtering at encoding and during maintenance.' In: *Journal of Experimental Psychology: Human Perception and Performance* 40.3 (2014), p. 960. DOI: 10.1037/a0036013.

[384] McPeek, Robert M, Skavenski, Alexander A, and Nakayama, Ken. 'Concurrent processing of saccades in visual search.' In: *Vision research* 40.18 (2000), pp. 2499–2516.

[385] Mehler, Bruce, Reimer, Bryan, Coughlin, Joseph F, and Dusek, Jeffery A. 'Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers.' In: *Transportation Research Record* 2138.1 (2009), pp. 6–12.

[386] Menghini, Luca, Cellini, Nicola, Goldstone, Aimee, Baker, Fiona C, and Zambotti, Massimiliano de. 'A standardized framework for testing the performance of sleep-tracking technology: step-by-step guidelines and open-source code.' In: *Sleep* 44.2 (2020). DOI: 10.1093/sleep/zsaa170. eprint: https://academic.oup.com/sleep/article-pdf/44/2/zsaa170/36244924/zsaa170.pdf.

[387] Menghini, Luca, Gianfranchi, Evelyn, Cellini, Nicola, Patron, Elisabetta, Tagliabue, Mariaelena, and Sarlo, Michela. 'Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions.' In: *Psychophysiology* 56.11 (2019), e13441. DOI: 10.1111/psyp.13441.

[388] Mentler, Tilo and Wessel, Daniel. 'A Toolkit for Studying Attention and Reaction Times to Smartglass Messages in Conditions of Different Perceptual Load.' In: *Proceedings of the 20th Congress of the Int. Ergonomics Association (IEA 2018) Volume V: Human Simulation and Virtual Environments, Work With Computing Systems (WWCS), Process Control 20*. Springer. 2019, pp. 595–604. DOI: 10.1007/978-3-319-96077-7_64.

[389] Merletti, Roberto and Farina, Dario. Surface Electromyography: Physiology, Engineering and Applications. IEEE Press Series on Biomedical Engineering. John Wiley & Sons, 2016.

[390] Milgram, Paul and Kishino, Fumio. 'A taxonomy of mixed reality visual displays.' In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), pp. 1321–1329. DOI: 10.1.1.102.4646.

[391] Milgram, Paul, Takemura, Haruo, Utsumi, Akira, and Kishino, Fumio. 'Augmented reality: a class of displays on the reality-virtuality continuum.' In: *Telemanipulator and Telepresence Technologies*. Vol. 2351. International Society for Optics and Photonics. SPIE, 1995, pp. 282–292. DOI: 10.1117/12.197321.

[392] Milstein, Nir and Gordon, Ilanit. 'Validating measures of electrodermal activity and heart rate variability derived from the empatica e4 utilized in research settings that involve interactive dyadic states.' In: *Frontiers in Behavioral Neuroscience* (2020), p. 148. DOI: 10.3389/fnbeh.2020.00148.

[393] Miskovic, Vladimir and Schmidt, Louis A. 'Cross-regional cortical synchronization during affective image viewing.' In: *Brain research* 1362 (2010), pp. 102–111. DOI: 10.1016/j.brainres.2010.09.102.

[394] Mitrevska, Teodora, Tag, Benjamin, Perusquía-Hernández, Monica, Niijima, Arinobu, Sidenmark, Ludwig, Solovey, Erin, Ali, Abdallah El, Mayer, Sven, and Chiossi, Francesco. 'SIG PhysioCHI: Human-Centered Physiological Computing in Practice.' In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Yokohama, Japan: ACM, 2025. DOI: 10.1145/3706599.3716289.

[395] Miyake, Shinji, Yamada, Shimpei, Shoji, Takuro, Takae, Yasuhiko, Kuge, Nobuyuki, and Yamamura, Tomohiro. 'Physiological responses to workload change. A test/retest examination.' In: *Applied ergonomics* 40.6 (2009), pp. 987–996. DOI: 10.1016/j.apergo.2009.02.005.

[396] Mnookin, Jennifer, Kellman, Philip J, Dror, Itiel, Erlikhman, Gennady, Garrigan, Patrick, Ghose, Tandra, Metler, Everett, and Charlton, Dave. 'Error rates for latent

fingerprinting as a function of visual complexity and cognitive difficulty.' In: *National Institute of Justice, US Department of Justice: Washington, DC* (2016).

[397] Momin, Abdul and Sanyal, Sudip. 'Analysis of Electrodermal Activity Signal Collected During Visual Attention Oriented Tasks.' In: *IEEE Access* 7 (2019), pp. 88186–88195. DOI: 10.1109/ACCESS.2019.2925933.

[398] Mühl, Christian, Allison, Brendan, Nijholt, Anton, and Chanel, Guillaume. 'A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges.' In: *Brain-Computer Interfaces* 1.2 (2014), pp. 66–84.

[399] Mühlberger, Andreas, Herrmann, Martin J, Wiedemann, Georg, Ellgring, Heiner, and Pauli, Paul. 'Repeated exposure of flight phobics to flights in virtual reality.' In: *Behaviour research and therapy* 39.9 (2001), pp. 1033–1050. DOI: 10.1016/S0005-7967(00)00076-0.

[400] Müller, Klaus-Robert, Tangermann, Michael, Dornhege, Guido, Krauledat, Matthias, Curio, Gabriel, and Blankertz, Benjamin. 'Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring.' In: *Journal of Neuroscience Methods* 167.1 (2008), pp. 82–90.

[401] Müller, Philipp, Staal, Sander, Bâce, Mihai, and Bulling, Andreas. 'Designing for Noticeability: Understanding the Impact of Visual Importance on Desktop Notifications.' In: *CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3501954.

[402] Muñoz, John E, Quintero, Luis, Stephens, Chad L, and Pope, Alan T. 'A psychophysiological model of firearms training in police officers: a virtual reality experiment for biocybernetic adaptation.' In: *Frontiers in psychology* 11 (2020), p. 683. DOI: 10.3389/fpsyg.2020.00683.

[403] Muñoz, John E., Cameirão, M., Bermúdez i Badia, S., and Gouveia, E. Rubio. 'Closing the Loop in Exergaming - Health Benefits of Biocybernetic Adaptation in Senior Adults.' In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '18. Melbourne, VIC, Australia: Association for Computing Machinery, 2018, pp. 329–339. DOI: 10.1145/3242671.3242673.

[404] Muñoz, John E., Cao, Shi, and Boger, Jennifer. 'Kinematically Adaptive Exergames: Personalizing Exercise Therapy Through Closed-Loop Systems.' In: *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 2019, pp. 118–1187. DOI: 10.1109/AIVR46125.2019.00026.

[405] Myers, Nicholas E, Stokes, Mark G, and Nobre, Anna C. 'Prioritizing information during working memory: beyond sustained internal attention.' In: *Trends in cognitive sciences* 21.6 (2017), pp. 449–461. DOI: 10.1016/j.tics.2017.03.010.

[406] Naderi, Mehrdad, Pladere, Tatjana, Alksnis, Reinis, and Krumina, Gunta. 'Brain activity underlying visual search in depth when viewing volumetric multiplanar images.' In: *Scientific Reports* 13.1 (2023), p. 7672. DOI: 10.1038/s41598-023-34758-9.

[407] Nalivaiko, Eugene, Davis, Simon L, Blackmore, Karen L, Vakulin, Andrew, and Nesbitt, Keith V. 'Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time.' In: *Physiology & behavior* 151 (2015), pp. 583–590.

[408] Navarro, Karla Felix. 'Wearable, wireless brain computer interfaces in augmented reality environments.' In: *International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004.* Vol. 2. IEEE. 2004, pp. 643–647. DOI: 10.1109/ITCC.2004.1286726.

[409] Negi, Shivsevak and Mitra, Ritayan. 'EEG Metrics to Determine Cognitive Load and Affective States: A Pilot Study.' In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. UbiComp '18. Singapore, Singapore: Association for Computing Machinery, 2018, pp. 182–185. DOI: 10.1145/3267305.3267618.

[410] Nikula, Reiner. 'Psychological correlates of nonspecific skin conductance responses.' In: *Psychophysiology* 28.1 (1991), pp. 86–90. DOI: https://doi.org/10.1111/j.1469-8986.1991.tb03392.x.

[411] Novak, Domen, Mihelj, Matjaž, and Munih, Marko. 'Dual-task performance in multimodal human-computer interaction: a psychophysiological perspective.' In: *Multimedia Tools and Applications* 56.3 (2012), pp. 553–567.

[412] Nwagu, Chukwuemeka, AlSlaity, Alaa, and Orji, Rita. 'EEG-Based Brain-Computer Interactions in Immersive Virtual and Augmented Reality: A Systematic Review.' In: *Proc. ACM Hum.-Comput. Interact.* 7.EICS (2023). DOI: 10.1145/3593226.

[413] O'Brien, Heather L. and Toms, Elaine G. 'What is user engagement? A conceptual framework for defining user engagement with technology.' In: *Journal of the American Society for Information Science and Technology* 59.6 (2008), pp. 938–955. DOI: 10.1002/asi.20801.

[414] O'Connell, Redmond G, Dockree, Paul M, Robertson, Ian H, Bellgrove, Mark A, Foxe, John J, and Kelly, Simon P. 'Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur.' In: *Journal of Neuroscience* 29.26 (2009), pp. 8604–8611. DOI: 10.1523/JNEUROSCI.5967-08.2009.

[415] O'Donnell, R. D. and Eggemeier, F. T. 'Workload assessment methodology.' In: *Handbook of perception and human performance. Volume 2: Cognitive processes and performance* (1986), pp. 42-1 - 42–49.

[416] Oever, Floris van den, Gorobets, Valentina, Saetrevik, Bjørn, Fjeld, Morten, and Kunz, Andreas. 'Comparing Visual Search between Physical Environments and VR.' In: *2022 IEEE Int. Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2022, pp. 411–416. DOI: 10.1109/ISMAR-Adjunct57072.2022.00089.

[417] Oliva, Manuel. 'Pupil size and search performance in low and high perceptual load.' In: *Cognitive, Affective, & Behavioral Neuroscience* 19.2 (2019), pp. 366–376. DOI: 10.3758/s13415-018-00677-w.

[418] Olivia, Aude, Mack, Michael L, Shrestha, Mochan, and Peeper, Angela. 'Identifying the perceptual dimensions of visual complexity of scenes.' In: *Proceedings of the annual meeting of the cognitive science society*. Vol. 26. escholarship, 2004.

[419] Olk, Bettina, Dinu, Alina, Zielinski, David J, and Kopper, Regis. 'Measuring visual search and distraction in immersive virtual reality.' In: *Royal Society open science* 5.5 (2018), p. 172331. DOI: 10.1098/rsos.172331.

[420] Onton, Julie, Delorme, Arnaud, and Makeig, Scott. 'Frontal midline EEG dynamics during working memory.' In: *Neuroimage* 27.2 (2005), pp. 341–356. DOI: 10.1016/j.neuroimage.2005.04.014.

[421] Ou, Changkun, Buschek, Daniel, Mayer, Sven, and Butz, Andreas. 'The Human in the Infinite Loop: A Case Study on Revealing and Explaining Human-AI Interaction Loop Failures.' In: *Proceedings of Mensch Und Computer 2022*. MuC '22. Darmstadt, Germany: Association for Computing Machinery, 2022, pp. 158–168. DOI: 10.1145/3543758.3543761.

[422] Ou, Changkun, Mayer, Sven, and Butz, Andreas. 'The Impact of Expertise in the Loop for Exploring Machine Rationality.' In: *28th ACM Symposium on Intelligent User Interface*. IUI '23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 158–173. DOI: 10.1145/3581641.3584040.

[423] Palinko, Oskar, Kun, Andrew L., Shyrokov, Alexander, and Heeman, Peter. 'Estimating Cognitive Load Using Remote Eye Tracking in a Driving Simulator.' In: *Proc. of the 2010 ETRA*. ETRA '10. Austin, Texas: ACM, 2010, pp. 141–144. DOI: 10.1145/1743666.1743701.

[424] Palinko, Oskar, Kun, Andrew L., Shyrokov, Alexander, and Heeman, Peter. 'Estimating cognitive load using remote eye tracking in a driving simulator.' In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. ETRA '10. Austin, Texas: Association for Computing Machinery, 2010, pp. 141–144. DOI: 10.1145/1743666.1743701.

[425] Palmer, Evan M, Horowitz, Todd S, Torralba, Antonio, and Wolfe, Jeremy M. 'What are the shapes of response time distributions in visual search?' In: *Journal of experi-*

mental psychology: human perception and performance 37.1 (2011), p. 58. DOI: 10.1037/a0020747.

[426] Pan, Zhigeng, Cheok, Adrian David, Yang, Hongwei, Zhu, Jiejie, and Shi, Jiaoying. 'Virtual reality and mixed reality for virtual learning environments.' In: *Computers & Graphics* 30.1 (2006), pp. 20–28. DOI: 10.1016/j.cag.2005.10.004.

[427] Papaioannou, Orestis and Luck, Steven J. 'Effects of eccentricity on the attention-related N2pc component of the event-related potential waveform.' In: *Psychophysiology* 57.5 (2020), e13532. DOI: 10.1111/psyp.13532.

[428] Parent, Mark, Peysakhovich, Vsevolod, Mandrick, Kevin, Tremblay, Sébastien, and Causse, Mickaël. 'The diagnosticity of psychophysiological signatures: Can we disentangle mental workload from acute stress with ECG and fNIRS?' In: *Int. J. of Psychophysiology* (2019). DOI: https://doi.org/10.1016/j.ijpsycho.2019.09.005.

[429] Parnandi, Avinash and Gutierrez-Osuna, Ricardo. 'A comparative study of game mechanics and control laws for an adaptive physiological game.' In: *Journal on Multimodal User Interfaces* 9.1 (2015), pp. 31–42. DOI: 10.1007/s12193-014-0159-y.

[430] Parsons, Thomas D and Courtney, Christopher G. 'Interactions between threat and executive control in a virtual reality stroop task.' In: *IEEE Transactions on Affective Computing* 9.1 (2016), pp. 66–75.

[431] Payne, Andrew F. H., Schell, Anne M., and Dawson, Michael E. 'Lapses in skin conductance responding across anatomical sites: Comparison of fingers, feet, forehead, and wrist.' In: *Psychophysiology* 53.7 (2016), pp. 1084–1092. DOI: 10.1111/psyp.12643.

[432] Pecchinenda, Anna. 'The affective significance of skin conductance activity during a difficult problem-solving task.' In: *Cognition & Emotion* 10.5 (1996), pp. 481–504.

[433] Peifer, Corinna, Schulz, André, Schächinger, Hartmut, Baumann, Nicola, and Antoni, Conny H. 'The relation of flow-experience and physiological arousal under stress—can u shape it?' In: *Journal of Experimental Social Psychology* 53 (2014), pp. 62–69.

[434] Peppet, Scott R. 'Regulating the internet of things: first steps toward managing discrimination, privacy, security and consent.' In: *Tex. L. Rev.* 93 (2014), p. 85. DOI: https://heinonline.org/HOL/P?h=hein.journals/tlr93&i=95.

[435] Pernice, Riccardo, Javorka, Michal, Krohova, Jana, Czippelova, Barbora, Turianikova, Zuzana, Busacca, Alessandro, and Faes, Luca. 'Reliability of short-term heart rate variability indexes assessed through photoplethysmography.' In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. EMBC'18. New York, NY, USA: IEEE, 2018, pp. 5610–5513. DOI: 10.1109/EMBC.2018.8513634.

[436] Perrin, François, Pernier, Jacques, Bertrand, Olivier, and Echallier, Jean Francois. 'Spherical splines for scalp potential and current density mapping.' In: *Electroencephalography and clinical neurophysiology* 72.2 (1989), pp. 184–187. DOI: 10.1016/0013-4694(89)90180-6.

[437] Peterson, Steven M, Furuichi, Emily, and Ferris, Daniel P. 'Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading.' In: *PloS one* 13.7 (2018), e0200306.

[438] Petrescu, Livia, Petrescu, Cătălin, Mitruț, Oana, Moise, Gabriela, Moldoveanu, Alin, Moldoveanu, Florica, and Leordeanu, Marius. 'Integrating biosignals measurement in virtual reality environments for anxiety detection.' In: *Sensors* 20.24 (2020), p. 7088. DOI: 10.3390/s20247088.

[439] Pfurtscheller, Gert, Stancak Jr, A, and Neuper, Ch. 'Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review.' In: *International journal of psychophysiology* 24.1-2 (1996), pp. 39–46. DOI: 10.1016/S0167-8760(96)00066-9.

[440] Pham, Duc-Minh and Stuerzlinger, Wolfgang. 'HawKEY: Efficient and Versatile Text Entry for Virtual Reality.' In: *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. VRST '19. Parramatta, NSW, Australia: Association for Computing Machinery, 2019. DOI: 10.1145/3359996.3364265.

[441] Pierno, Andrea C, Caria, Andrea, Glover, Scott, and Castiello, Umberto. 'Effects of increasing visual load on aurally and visually guided target acquisition in a virtual environment.' In: *Applied ergonomics* 36.3 (2005), pp. 335–343.

[442] Pietilä, Julia, Mehrang, Saeed, Tolonen, Johanna, Helander, Elina, Jimison, Holly, Pavel, Misha, and Korhonen, Ilkka. 'Evaluation of the accuracy and reliability for photoplethysmography based heart rate and beat-to-beat detection during daily activities.' In: *EMBEC & NBC 2017: Joint Conference of the European Medical and Biological Engineering Conference (EMBEC) and the Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC), Tampere, Finland, June 2017*. Springer. 2018, pp. 145–148. DOI: 10.1007/978-981-10-5122-7_37.

[443] Pion-Tonachini, Luca, Kreutz-Delgado, Ken, and Makeig, Scott. 'ICLabel: An automated electroencephalographic independent component classifier, dataset, and website.' In: *NeuroImage* 198.76 (2019), pp. 181–197. DOI: 10.21105/joss.04484.

[444] Pointecker, Fabian, Friedl-Knirsch, Judith, Jetter, Hans-Christian, and Anthes, Christoph. 'From Real to Virtual: Exploring Replica-Enhanced Environment Transitions along the Reality-Virtuality Continuum.' In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '24. Honolulu, HI, USA: Association for Computing Machinery, 2024. DOI: 10.1145/3613904.3642844.

[445] Polich, John. 'Updating P300: an integrative theory of P3a and P3b.' In: *Clinical neurophysiology* 118.10 (2007), pp. 2128–2148. DOI: 10.1016/j.clinph.2007.04.019.

[446] Poon, C.C.Y. and Zhang, Y.T. 'Cuff-less and Noninvasive Measurements of Arterial Blood Pressure by Pulse Transit Time.' In: *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. New York, NY, USA: IEEE, 2005, pp. 5877–5880. DOI: 10.1109/IEMBS.2005.1615827.

[447] Pope, Alan T, Bogart, Edward H, and Bartolome, Debbie S. 'Biocybernetic system evaluates indices of operator engagement in automated task.' In: *Biological psychology* 40.1-2 (1995), pp. 187–195. DOI: 10.1016/0301-0511(95)05116-3.

[448] Pornpattananangkul, Narun, Grogans, Shannon, Yu, Rongjun, and Nusslock, Robin. 'Single-trial EEG dissociates motivation and conflict processes during decision-making under risk.' In: *Neuroimage* 188 (2019), pp. 483–501.

[449] Posada-Quintero, Hugo F and Bolkhovsky, Jeffrey B. 'Machine learning models for the identification of cognitive tasks using autonomic reactions from heart rate variability and electrodermal activity.' In: *Behavioral Sciences* 9.4 (2019), p. 45. DOI: 10.3390/bs9040045.

[450] Posada-Quintero, Hugo F. and Chon, Ki H. 'Phasic Component of Electrodermal Activity is more Correlated to Brain Activity than Tonic Component.' In: *2019 IEEE EMBS International Conference on Biomedical & Health Informatics*. BHI. New York, NY, USA: IEEE, 2019, pp. 1–4. DOI: 10.1109/BHI.2019.8834567.

[451] Poston, Alan. 'Human engineering design data digest.' In: *Washington, DC: Department of Defense Human Factors Engineering Technical Advisory Group* (2000), pp. 61–75.

[452] Pratt, Nikki, Willoughby, Adrian, and Swick, Diane. 'Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence.' In: *Frontiers in human neuroscience* 5 (2011), p. 57. DOI: 10.3389/fnhum.2011.00057.

[453] Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, Society for, Boucsein, Wolfram, Fowles, Don C., Grimnes, Sverre, Ben-Shakhar, Gershon, Roth, Walton T., Dawson, Michael E., and Filion, Diane L. 'Publication Recommendations for Electrodermal Measurements.' In: *Psychophysiology* 49.8 (2012), pp. 1017–1034. DOI: 10.1111/j.1469-8986.1981.tb03024.x.

[454] Puma, Sébastien, Matton, Nadine, Paubel, Pierre-V, Raufaste, Éric, and El-Yagoubi, Radouane. 'Using theta and alpha band power to assess cognitive workload in multitasking environments.' In: *International Journal of Psychophysiology* 123 (2018), pp. 111–120.

[455] Putman, Peter, Verkuil, Bart, Arias-Garcia, Elsa, Pantazi, Ioanna, and Schie, Charlotte van. 'EEG theta/beta ratio as a potential biomarker for attentional control and

resilience against deleterious effects of stress on attention.' In: *Cognitive, Affective, & Behavioral Neuroscience* 14 (2014), pp. 782–791. DOI: 10.3758/s13415-013-0238-7.

[456] Putze, Felix, Popp, Johannes, Hild, Jutta, Beyerer, Jürgen, and Schultz, Tanja. 'Intervention-Free Selection Using EEG and Eye Tracking.' In: *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ICMI '16. Tokyo, Japan: Association for Computing Machinery, 2016, pp. 153–160. DOI: 10.1145/2993148.2993199.

[457] Putze, Felix, Scherer, Maximilian, and Schultz, Tanja. 'Starring into the Void? Classifying Internal vs. External Attention from EEG.' In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. NordiCHI '16. Gothenburg, Sweden: ACM, 2016, pp. 1–4. DOI: 10.1145/2971485.2971555.

[458] Querino, Emanuel, Dos Santos, Lafaiete, Ginani, Giuliano, Nicolau, Eduardo, Miranda, Débora, Romano-Silva, Marco, and Malloy-Diniz, Leandro. 'Cognitive effort and pupil dilation in controlled and automatic processes.' In: *Translational Neuroscience* 6.1 (2015), pp. 168–173.

[459] Quintero, Luis, Papapetrou, Panagiotis, and Muñoz, John E. 'Open-source physiological computing framework using heart rate variability in mobile virtual reality applications.' In: *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE. 2019, pp. 126–1267.

[460] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2021.

[461] Radu, Iulian, Joy, Tugce, Bowman, Yiran, Bott, Ian, and Schneider, Bertrand. 'A Survey of Needs and Features for Augmented Reality Collaborations in Collocated Spaces.' In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW1 (2021). DOI: 10.1145/3449243.

[462] Ragan, Eric D., Bowman, Doug A., Kopper, Regis, Stinson, Cheryl, Scerbo, Siroberto, and McMahan, Ryan P. 'Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task.' In: *IEEE TVCG* 21.7 (2015), pp. 794–807. DOI: 10.1109/TVCG.2015.2403312.

[463] Rajan, Rahul, Selker, Ted, and Lane, Ian. 'Task Load Estimation and Mediation Using Psycho-Physiological Measures.' In: *Proceedings of the 21st International Conference on Intelligent User Interfaces*. IUI '16. Sonoma, California, USA: Association for Computing Machinery, 2016, pp. 48–59. DOI: 10.1145/2856767.2856769.

[464] Rämä, PIA and Baccino, Thierry. 'Eye fixation–related potentials (EFRPs) during object identification.' In: *Visual Neuroscience* 27.5-6 (2010), pp. 187–192. DOI: 10.1017/S0952523810000283.

[465] Ramakrishnan, Ramesh Kumar, Gavas, Rahul, Viraraghavan, Venkata Subramaninan, Hissaria, Lalit Kumar, Pal, Arpan, and Balamuralidhar, P. 'Monitoring The Early Physiological Effects of Fitness Regimes Using PPG Signals.' In: *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. New York, NY, USA: IEEE, 2020, pp. 1–5. DOI: 10.1109/PerComWorkshops48775.2020.9156258.

[466] Rappa, Natasha Anne, Ledger, Susan, Teo, Timothy, Wai Wong, Kok, Power, Brad, and Hilliard, Bruce. 'The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: a scoping review.' In: *Interactive Learning Environments* 30.7 (2022), pp. 1338–1350. DOI: 10.1080/10494820.2019.1702560.

[467] Rasch, Julian, Wilhalm, Matthias, Müller, Florian, and Chiossi, Francesco. 'AR You on Track? Investigating Effects of Augmented Reality Anchoring on Dual-Task Performance While Walking.' In: *CHI Conference on Human Factors in Computing Systems (CHI '25)*. Yokohama, Japan: ACM, 2025. DOI: 10.1145/3706598.3714258.

[468] Raufi, Bujar and Longo, Luca. 'An Evaluation of the EEG alpha-to-theta and theta-to-alpha band Ratios as Indexes of Mental Workload.' In: *Frontiers in Neuroinformatics* 16 (2022), p. 44. DOI: https://doi.org/10.3389/fninf.2022.861967.

[469] Raugh, Ian M., Chapman, Hannah C., Bartolomeo, Lisa A., Gonzalez, Cristina, and Strauss, Gregory P. 'A comprehensive review of psychophysiological applications for ecological momentary assessment in psychiatric populations.' In: *Psychological assessment* 31.3 (2019), p. 304. DOI: 10.1037/pas0000651.

[470] Ravaja, Niklas. 'Contributions of psychophysiology to media research: Review and recommendations.' In: *Media Psychology* 6.2 (2004), pp. 193–235. DOI: 10.1207/s1532785xmep0602_4.

[471] Ravenzwaaij, Don van, Monden, Rei, Tendeiro, Jorge N, and Ioannidis, John. 'Bayes factors for superiority, non-inferiority, and equivalence designs.' In: *BMC medical research methodology* 19.1 (2019), pp. 1–12. DOI: 10.1186/s12874-019-0699-7.

[472] Rayner, Keith. 'Eye movements in reading and information processing: 20 years of research.' In: *Psychological bulletin* 124.3 (1998), p. 372. DOI: 10.1037/0033-2909.124.3.372.

[473] Rees, Geraint, Frith, Chris, and Lavie, Nilli. 'Processing of irrelevant visual motion during performance of an auditory attention task.' In: *Neuropsychologia* 39.9 (2001), pp. 937–949. DOI: 10.1016/S0028-3932(01)00016-1.

[474] Regenbrecht, Holger, Lum, Tim, Kohler, Petra, Ott, Claudia, Wagner, M, Wilke, Wilhelm, and Mueller, Erich. 'Using augmented virtuality for remote collaboration.' In: *Presence* 13.3 (2004), pp. 338–354.

[475] Rehman, Umair and Cao, Shi. 'Augmented-reality-based indoor navigation: A comparative analysis of handheld devices versus google glass.' In: *IEEE Transactions on Human-Machine Systems* 47.1 (2016), pp. 140–151.

[476] Richter, M, Gendolla, GHE, and Wright, RA. 'Three decades of research on motivational intensity theory: What we have learned about effort and what we still don't know.' In: *Advances in motivation science* 3 (2016), pp. 149–186. DOI: `10.1016/bs.adms.2016.02.001`.

[477] Richter, M., Gendolla, G.H.E., and Wright, R.A. 'Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know.' In: *Advances in Motivation Science*. Advances in Motivation Science (2016). Ed. by Andrew J. Elliot. DOI: `https://doi.org/10.1016/bs.adms.2016.02.001`.

[478] Riddle, Justin, Scimeca, Jason M, Cellier, Dillan, Dhanani, Sofia, and D'Esposito, Mark. 'Causal evidence for a role of theta and alpha oscillations in the control of working memory.' In: *Current Biology* 30.9 (2020), pp. 1748–1754. DOI: `10.1016/j.cub.2020.02.065`.

[479] Ries, Anthony J, Touryan, Jon, Ahrens, Barry, and Connolly, Patrick. 'The impact of task demands on fixation-related brain potentials during guided search.' In: *PloS one* 11.6 (2016), e0157260. DOI: `10.1371/journal.pone.0157260`.

[480] Righi, Luana L, Xavier, Gilberto F, Baldo, Marcus VC, and Haddad, Hamilton. 'Responses to auditory distractors during a class depend on participant's enjoyment: An eye-tracking and skin conductance study.' In: *Psychology & Neuroscience* 12.4 (2019), p. 473. DOI: `https://doi.org/10.1037/pne0000178`.

[481] Rissman, Jesse, Gazzaley, Adam, and D'Esposito, Mark. 'The effect of non-visual working memory load on top-down modulation of visual processing.' In: *Neuropsychologia* 47.7 (2009), pp. 1637–1646.

[482] Rivu, Radiah, Mäkelä, Ville, Hassib, Mariam, Abdelrahman, Yomna, and Alt, Florian. 'Exploring How Saliency Affects Attention in Virtual Reality.' In: *IFIP Conference on Human-Computer Interaction*. Springer. 2021, pp. 147–155. DOI: `https://doi.org/10.1007/978-3-030-85607-6_10`.

[483] Rivu, Radiah, Zhou, Yumeng, Welsch, Robin, Mäkelä, Ville, and Alt, Florian. 'When friends become strangers: Understanding the influence of avatar gender on interpersonal distance in virtual reality.' In: *IFIP Conf. on Human-Computer Interaction*. Springer, 2021, pp. 234–250. DOI: `10.1007/978-3-030-85607-6_16`.

[484] Roche, Richard AP, Garavan, Hugh, Foxe, John J, and O'Mara, Shane M. 'Individual differences discriminate event-related potentials but not performance during response inhibition.' In: *Experimental brain research* 160 (2005), pp. 60–70.

[485] Rodrigues, Nils, Shao, Lin, Yan, Jia Jun, Schreck, Tobias, and Weiskopf. 'Eye Gaze on Scatterplot: Concept and First Results of Recommendations for Exploration of SPLOMs Using Implicit Data Selection.' In: *Proceedings of the 14th ACM Symposium on Eye Tracking Research & Applications*. ETVIS '22. Seattle, Washington: Association for Computing Machinery, 2022.

[486] Roo, Joan Sol and Hachet, Martin. 'One Reality: Augmenting How the Physical World is Experienced by Combining Multiple Mixed Reality Modalities.' In: *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. UIST '17. Québec City, QC, Canada: ACM, 2017, pp. 787–795. DOI: 10.1145/3126594.3126638.

[487] Roper, Zachary JJ, Cosman, Joshua D, and Vecera, Shaun P. 'Perceptual load corresponds with factors known to influence visual search.' In: *Journal of experimental psychology: human perception and performance* 39.5 (2013), p. 1340.

[488] Rowe, Gillian, Hirsh, Jacob B, and Anderson, Adam K. 'Positive affect increases the breadth of attentional selection.' In: *Proceedings of the National Academy of Sciences* 104.1 (2007), pp. 383–388.

[489] Rushby, Jacqueline A, Barry, Robert J, Clarke, Adam R, and VaezMousavi, Mohammad. 'Arousal and activation effects on physiological and behavioral responding during a continuous performance task.' In: (2007).

[490] Rzayev, Rufat, Mayer, Sven, Krauter, Christian, and Henze, Niels. 'Notification in VR: The Effect of Notification Placement, Task and Environment.' In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '19. Barcelona, Spain: Association for Computing Machinery, 2019, pp. 199–211. DOI: 10.1145/3311350.3347190.

[491] Sakai, Tsugunosuke, Tamaki, Harunya, Ota, Yosuke, Egusa, Ryohei, Imagaki, Shigenori, Kusunoki, Fusako, Sugimoto, Masanori, and Mizoguchi, Hiroshi. 'Eda-based estimation of visual attention by observation of eye blink frequency.' In: *Int. Journal on Smart Sensing and Intelligent Systems* 10.2 (2017), pp. 1–12. DOI: 10.21307/ijssis-2017-212.

[492] Salselas, Inês, Penha, Rui, and Bernardes, Gilberto. 'Sound design inducing attention in the context of audiovisual immersive environments.' In: *Personal and Ubiquitous Computing* 25 (2021), pp. 737–748. DOI: 10.1007/s00779-020-01386-3.

[493] Salvucci, Dario D and Goldberg, Joseph H. 'Identifying fixations and saccades in eye-tracking protocols.' In: *Proceedings of the 2000 symposium on Eye tracking research & applications*. 2000, pp. 71–78. DOI: 10.1145/355017.355028.

[494] Samini, Ali, Palmerius, Karljohan Lundin, and Ljung, Patric. 'A review of current, complete augmented reality solutions.' In: *2021 Int. Conference on Cyberworlds (CW)*. IEEE. 2021, pp. 49–56.

[495] Sandhu, Rajwant and Dyson, Benjamin James. 'Cross-modal perceptual load: the impact of modality and individual differences.' In: *Experimental Brain Research* 234 (2016), pp. 1279–1291. DOI: 10.1007/s00221-015-4517-0.

[496] Saponas, T Scott, Tan, Desney S., Morris, Dan, and Balakrishnan, Ravin. 'Demonstrating the Feasibility of Using Forearm Electromyography for Muscle-Computer Interfaces.' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. New York, NY, USA: ACM, 2008, pp. 515–524. DOI: 10.1145/1357054.1357138.

[497] Sauseng, Paul, Griesmayr, Birgit, Freunberger, Roman, and Klimesch, Wolfgang. 'Control mechanisms in working memory: a possible function of EEG theta oscillations.' In: *Neuroscience & Biobehavioral Reviews* 34.7 (2010), pp. 1015–1022. DOI: 10.1016/j.neubiorev.2009.12.006.

[498] Sauseng, Paul, Klimesch, Wolfgang, Schabus, Manuel, and Doppelmayr, Michael. 'Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory.' In: *International journal of Psychophysiology* 57.2 (2005), pp. 97–103. DOI: https://doi.org/10.1016/j.ijpsycho.2005.03.018.

[499] Sauseng, Paul, Klimesch, Wolfgang, Stadler, Waltraud, Schabus, Manuel, Doppelmayr, Michael, Hanslmayr, Simon, Gruber, Walter R, and Birbaumer, Niels. 'A shift of visual spatial attention is selectively associated with human EEG alpha activity.' In: *European journal of neuroscience* 22.11 (2005), pp. 2917–2926.

[500] Sawaki, Risa, Geng, Joy J, and Luck, Steven J. 'A common neural mechanism for preventing and terminating the allocation of attention.' In: *Journal of Neuroscience* 32.31 (2012), pp. 10725–10736.

[501] Schäfer, Axel and Vagedes, Jan. 'How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram.' In: *International journal of cardiology* 166.1 (2013), pp. 15–29. DOI: 10.1016/j.ijcard.2012.03.119.

[502] Schilit, B., Adams, N., and Want, R. 'Context-Aware Computing Applications.' In: *1994 First Workshop on Mobile Computing Systems and Applications*. IEEE. 1994, pp. 85–90. DOI: 10.1109/WMCSA.1994.16.

[503] Schilit, B., Adams, N., and Want, R. 'Context-Aware Computing Applications.' In: *1994 First Workshop on Mobile Computing Systems and Applications*. 1994, pp. 85–90. DOI: 10.1109/WMCSA.1994.16.

[504] Schmalstieg, Dieter and Hollerer, Tobias. Augmented reality: principles and practice. Addison-Wesley Professional, 2016.

[505] Schmidt, Albrecht, Beigl, Michael, and Gellersen, Hans-W. 'There is more to context than location.' In: *Computers & Graphics* 23.6 (1999), pp. 893–901.

[506] 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.

[507] 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.

[508] Schneider, RL. 'A mathematical-model of human-skin conductance.' In: *Psychophysiology*. Vol. 24. 5. SOC PSYCHOPHYSIOL RES 1010 VERMONT AVE NW SUITE 1100, WASHINGTON, DC 20005, 1987, pp. 610–610.

[509] Schwerdtfeger, Bjorn and Klinker, Gudrun. 'Supporting order picking with augmented reality.' In: *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. IEEE. 2008, pp. 91–94.

[510] Sedano, Carolina I., Tikka, Päivi, and Nieminen, Jari. 'Towards Analyzing the Cognitive, Emotional, and Social Engagement in Game-Based Learning.' In: *Journal of Interactive Learning Research* 24.3 (2013), pp. 289–316.

[511] Selzer, Matias N, Gazcon, Nicolas F, and Larrea, Martin L. 'Effects of virtual presence and learning outcome using low-end virtual reality systems.' In: *Displays* 59 (2019), pp. 9–15. DOI: 10.1016/j.displa.2019.04.002.

[512] Setz, Cornelia, Arnrich, Bert, Schumm, Johannes, La Marca, Roberto, Tröster, Gerhard, and Ehlert, Ulrike. 'Discriminating Stress From Cognitive Load Using a Wearable EDA Device.' In: *IEEE Transactions on Information Technology in Biomedicine* 14.2 (2010), pp. 410–417. DOI: 10.1109/TITB.2009.2036164.

[513] Sharma, Mansi, Rekrut, Maurice, Alexandersson, Jan, and Krüger, Antonio. 'Towards Improving EEG-Based Intent Recognition in Visual Search Tasks.' In: *Int. Conference on Neural Information Processing*. Springer. 2022, pp. 604–615. DOI: 10.1007/978-3-031-30111-7_51.

[514] Shenoy, Pradeep, Krauledat, Matthias, Blankertz, Benjamin, Rao, Rajesh PN, and Müller, Klaus-Robert. 'Towards adaptive classification for BCI.' In: *Journal of neural engineering* (2006).

[515] Shimomura, Tomonari and Kumada, Takatsune. 'Spatial working memory load affects counting but not subitizing in enumeration.' In: *Attention, Perception, & Psychophysics* 73 (2011), pp. 1694–1709. DOI: 10.3758/s13414-011-0135-5.

[516] Shin, Jae-eun, Yoon, Boram, Kim, Dooyoung, and Woo, Woontack. 'A User-Oriented Approach to Space-Adaptive Augmentation: The Effects of Spatial Affordance on Narrative Experience in an Augmented Reality Detective Game.' In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, 2021, pp. 1–13. DOI: 10.1145/3411764.3445675.

[517] Shishikura, Motofumi, Tamura, Hiroshi, and Sakai, Ko. 'Correlation between neural responses and human perception in figure-ground segregation.' In: *Frontiers in systems neuroscience* 16 (2023), p. 999575. DOI: 10.3389/fnsys.2022.999575.

[518] Sidner, Candace L., Dzikovska, Myroslava, Michaelis, Jim, and Rich, Charles. 'The Role of Engagement in Dialogue Systems for Virtual Humans.' In: *International Conference on Intelligent Virtual Agents*. 2004, pp. 123–140.

[519] Silva, Nelson, Blascheck, Tanja, Jianu, Radu, Rodrigues, Nils, Weiskopf, Daniel, Raubal, Martin, and Schreck, Tobias. 'Eye Tracking Support for Visual Analytics Systems: Foundations, Current Applications, and Research Challenges.' In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*. ETRA '19. Denver, Colorado: Association for Computing Machinery, 2019. DOI: 10.1145/3314111.3319919.

[520] Skarbez, Richard, Smith, Missie, and Whitton, Mary C. 'Revisiting Milgram and Kishino's reality-virtuality continuum.' In: *Frontiers in Virtual Reality* 2 (2021), p. 647997. DOI: 10.3389/frvir.2021.647997.

[521] Skulmowski, Alexander, Nebel, Steve, Remmele, Martin, and Rey, Günter Daniel. 'Is a preference for realism really naive after all? A cognitive model of learning with realistic visualizations.' In: *Educational Psychology Review* (2021), pp. 1–27. DOI: 10.1007/s10648-021-09638-1.

[522] Slater, Mel. 'A note on presence terminology.' In: *Presence connect* 3.3 (2003), pp. 1–5.

[523] Slater, Mel. 'Immersion and the illusion of presence in virtual reality.' In: *British journal of psychology* 109.3 (2018), pp. 431–433. DOI: 10.1111/bjop.12305.

[524] Slater, Mel and Sanchez-Vives, Maria V. 'Enhancing our lives with immersive virtual reality.' In: *Frontiers in Robotics and AI* 3 (2016), p. 74.

[525] Smallwood, J. and Schooler, J.W. 'The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness.' In: *Annual Review of Psychology* 66 (2015), pp. 487–518. DOI: 10.1146/annurev-psych-010814-015331.

[526] Sokolov, Evgeny N, Spinks, John A, Näätänen, Risto, and Lyytinen, Heikki. The orienting response in information processing. Lawrence Erlbaum Associates Publishers, 2002. DOI: 10.4324/9781410601490.

[527] Souchet, Alexis D., Diallo, Mamadou Lamarana, and Lourdeaux, Domitile. 'Cognitive load Classification with a Stroop task in Virtual Reality based on Physiological data.' In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2022, pp. 656–666. DOI: 10.1109/ISMAR55827.2022.00083.

[528] Soukoreff, R William and MacKenzie, I Scott. 'Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric.' In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, pp. 113–120. DOI: https://doi.org/10.1145/642611.642632.

[529] Souza, Rhaíra Helena Caetano e and Naves, Eduardo Lázaro Martins. 'Attention detection in virtual environments using EEG signals: A scoping review.' In: *frontiers in physiology* 12 (2021), p. 727840. DOI: 10.3389/fphys.2021.727840.

[530] Soveri, Anna, Antfolk, Jan, Karlsson, Linda, Salo, Benny, and Laine, Matti. 'Working memory training revisited: A multi-level meta-analysis of n-back training studies.' In: *Psychonomic bulletin & review* 24 (2017), pp. 1077–1096. DOI: 10.3758/s13423-016-1217-0.

[531] Speicher, Marco, Feit, Anna Maria, Ziegler, Pascal, and Krüger, Antonio. 'Selection-Based Text Entry in Virtual Reality.' In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. , Montreal QC, Canada, Association for Computing Machinery, 2018, pp. 1–13. DOI: 10.1145/3173574.3174221.

[532] Splawn, Joshua M and Miller, Michael E. 'Prediction of perceived workload from task performance and heart rate measures.' In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 57. SAGE Publications Sage CA: Los Angeles, CA. 2013, pp. 778–782. DOI: 10.1177/1541931213571170.

[533] Staib, Matthias, Castegnetti, Giuseppe, and Bach, Dominik R. 'Optimising a model-based approach to inferring fear learning from skin conductance responses.' In: *Journal of neuroscience methods* 255 (2015), pp. 131–138.

[534] Stauffert, Jan-Philipp, Niebling, Florian, and Latoschik, Marc Erich. 'Latency and cybersickness: impact, causes, and measures. A review.' In: *Frontiers in Virtual Reality* 1 (2020), p. 582204. DOI: 10.3389/frvir.2020.582204.

[535] Steichen, Ben and Fu, Bo. 'Towards adaptive information visualization-a study of information visualization aids and the role of user cognitive style.' In: *Frontiers in Artificial Intelligence* (2019), p. 22.

[536] Stein, Phyllis K, Ehsani, Ali A, Domitrovich, Peter P, Kleiger, Robert E, and Rottman, Jeffrey N. 'Effect of exercise training on heart rate variability in healthy older adults.' In: *American heart journal* 138.3 (1999), pp. 567–576. DOI: 10.1007/s00421-009-1089-3.

[537] Stepanova, Ekaterina R., Haynes, Alice C., Turmo Vidal, Laia, Chiossi, Francesco, El Ali, Abdallah, Quintero, Luis, Luft, Yoav, Campo Woytuk, Nadia, and Mayer, Sven. 'Envisioning Ubiquitous Biosignal Interaction with Multimedia.' In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*. MUM '24. Association for Computing Machinery, 2024, pp. 495–500. DOI: 10.1145/3701571.3701609.

[538] Streli, Paul, Armani, Rayan, Cheng, Yi Fei, and Holz, Christian. 'HOOV: Hand Out-Of-View Tracking for Proprioceptive Interaction Using Inertial Sensing.' In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. DOI: 10.1145/3544548.3581468.

[539] Stuldreher, Ivo V, Thammasan, Nattapong, Van Erp, Jan BF, and Brouwer, Anne-Marie. 'Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time.' In: *Frontiers in Neuroscience* 14 (2020), p. 575521. DOI: 10.3389/fnins.2020.575521.

[540] Sturman, D.J. and Zeltzer, D. 'A survey of glove-based input.' In: *IEEE Computer Graphics and Applications* 14.1 (1994), pp. 30–39. DOI: 10.1109/38.250916.

[541] Sutherland, Ivan E. 'A head-mounted three dimensional display.' In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. 1968, pp. 757–764.

[542] Sutherland, Ivan E et al. 'The ultimate display.' In: *Proceedings of the IFIP Congress*. Vol. 2. 506-508. New York. 1965, pp. 506–508.

[543] Syed, Toqeer Ali, Siddiqui, Muhammad Shoaib, Abdullah, Hurria Binte, Jan, Salman, Namoun, Abdallah, Alzahrani, Ali, Nadeem, Adnan, and Alkhodre, Ahmad B. 'In-depth review of augmented reality: Tracking technologies, development tools, AR displays, collaborative AR, and security concerns.' In: *Sensors* 23.1 (2022), p. 146.

[544] Talsma, Durk, Doty, Tracy J, and Woldorff, Marty G. 'Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration?' In: *Cerebral cortex* 17.3 (2007), pp. 679–690.

[545] Tang, Arthur, Owen, Charles, Biocca, Frank, and Mou, Weimin. 'Experimental evaluation of augmented reality in object assembly task.' In: *Proceedings. Int. symposium on mixed and augmented reality*. IEEE. 2002, pp. 265–266.

[546] Tao, Da, Tan, Haibo, Wang, Hailiang, Zhang, Xu, Qu, Xingda, and Zhang, Tingru. 'A systematic review of physiological measures of mental workload.' In: *International journal of environmental research and public health* 16.15 (2019), p. 2716.

[547] Tavares, Gabriela, Perona, Pietro, and Rangel, Antonio. 'The attentional drift diffusion model of simple perceptual decision-making.' In: *Frontiers in neuroscience* 11 (2017), p. 468. DOI: 10.3389/fnins.2017.00468.

[548] Thomas, P CAUDELL and David, WM. 'Augmented reality: An application of heads-up display technology to manual manufacturing processes.' In: *Hawaii international conference on system sciences*. Vol. 2. ACM SIGCHI Bulletin. 1992, pp. 659–669.

[549] Toker, Dereck, Conati, Cristina, Carenini, Giuseppe, and Haraty, Mona. 'Towards adaptive information visualization: on the influence of user characteristics.' In: *International Conference on User Modeling, Adaptation, and Personalization*. Springer. 2012, pp. 274–285.

[550] Tran, Yvonne, Craig, Ashley, Craig, Rachel, Chai, Rifai, and Nguyen, Hung. 'The influence of mental fatigue on brain activity: Evidence from a systematic review with meta-analyses.' In: *Psychophysiology* 57.5 (2020), e13554. DOI: https://doi.org/10.1111/psyp.13554.

[551] Treacy Solovey, Erin, Afergan, Daniel, Peck, Evan M., Hincks, Samuel W., and Jacob, Robert J. K. 'Designing Implicit Interfaces for Physiological Computing: Guidelines and Lessons Learned Using FNIRS.' In: *ACM Trans. Comput.-Hum. Interact.* 21.6 (2015). DOI: 10.1145/2687926.

[552] Treisman, Anne M. 'Selective attention in man.' In: *British Medical Bulletin* 20.1 (1964), pp. 12–16.

[553] Treisman, Anne M. and Gelade, Garry. 'A feature-integration theory of attention.' In: *Cognitive Psychology* 12.1 (1980), pp. 97–136. DOI: 10.1016/0010-0285(80)90005-5.

[554] Trejo, Leonard J, Kubitz, Karla, Rosipal, Roman, Kochavi, Rebekah L, Montgomery, Leslie D, et al. 'EEG-based estimation and classification of mental fatigue.' In: *Psychology* (2015). DOI: 10.4236/psych.2015.65055.

[555] Tremmel, Christoph, Herff, Christian, Sato, Tetsuya, Rechowicz, Krzysztof, Yamani, Yusuke, and Krusienski, Dean J. 'Estimating cognitive workload in an interactive virtual reality environment using EEG.' In: *Frontiers in human neuroscience* 13 (2019), p. 401. DOI: 10.3389/fnhum.2019.00401.

[556] Trick, Lana M and Pylyshyn, Zenon W. 'Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision.' In: *Psychological review* 101.1 (1994), p. 80. DOI: 10.1037/0033-295X.101.1.80.

[557] Troiano, Alexandra. 'Wearables and personal health data: putting a premium on your privacy.' In: *Brook. L. Rev.* 82 (2016), p. 1715. DOI: https://brooklynworks.brooklaw.edu/blr/vol82/iss4/6.

[558] Udovičić, Goran, Topić, Ante, and Russo, Mladen. 'Wearable technologies for smart environments: A review with emphasis on BCI.' In: *2016 24th International Conference*

on Software, Telecommunications and Computer Networks (SoftCOM). 2016, pp. 1–9. DOI: 10.1109/SOFTCOM.2016.7772186.

[559] Umer, Waleed, Yu, Yantao, Antwi-Afari, Maxwell Fordjour, Jue, Li, Siddiqui, Mohsin K, and Li, Heng. 'Heart rate variability based physical exertion monitoring for manual material handling tasks.' In: *International Journal of Industrial Ergonomics* 89 (2022), p. 103301. DOI: 10.1016/j.ergon.2022.103301.

[560] Ušćumlić, Marija and Blankertz, Benjamin. 'Active visual search in non-stationary scenes: coping with temporal variability and uncertainty.' In: *Journal of Neural Engineering* 13.1 (2016), p. 016015. DOI: 10.1088/1741-2560/13/1/016015.

[561] Van Berkel, Niels and Hornbæk, Kasper. 'Implications of Human-Computer Interaction Research.' In: *interactions* 30.4 (2023), pp. 50–55. DOI: 10.1145/3600103.

[562] Van den Berg, AV and Van Loon, EM. 'An invariant for timing of saccades during visual search.' In: *Vision Research* 45.12 (2005), pp. 1543–1555. DOI: 10.1016/j.visres.2004.12.018.

[563] Varela-Aldás, José, Fuentes, Esteban M, Palacios-Navarro, Guillermo, and García-Magariño, Iván. 'A comparison of heart rate in normal physical activity vs. immersive virtual reality exergames.' In: *International Conference on Human Systems Engineering and Design: Future Trends and Applications*. Cham, Switzerland: Springer, 2019, pp. 684–689. DOI: 10.1007/978-3-030-27928-8_104.

[564] Veltman, JA and Gaillard, AWK. 'Physiological workload reactions to increasing levels of task difficulty.' In: *Ergonomics* 41.5 (1998), pp. 656–669. DOI: 10.1080/001401398186829.

[565] Venables, Peter H and Christie, Margaret J. 'Electrodermal activity.' In: *Techniques in psychophysiology* 54.3 (1980).

[566] Venkatesh, Viswanath, Morris, Michael G., Davis, Gordon B., and Davis, Fred D. 'User acceptance of information technology: Toward a unified view.' In: *MIS quarterly* (2003), pp. 425–478. DOI: 10.2307/30036540.

[567] Vidaurre, Carmen, Kawanabe, Motoaki, Bünau, Paul von, Blankertz, Benjamin, and Müller, Klaus-Robert. 'Toward unsupervised adaptation of LDA for brain–computer interfaces.' In: *IEEE Transactions on Biomedical Engineering* 58.3 (2010), pp. 587–597.

[568] Vilardaga, Roger, Bricker, Jonathan B., and McDonell, Michael G. 'The promise of mobile technologies and single case designs for the study of individuals in their natural environment.' In: *Journal of Contextual Behavioral Science* 3.2 (2014), pp. 148–153. DOI: 10.1016/j.jcbs.2014.03.003.

[569] Vitali, Francesca, Tarperi, Cantor, Cristini, Jacopo, Rinaldi, Andrea, Zelli, Arnaldo, Lucidi, Fabio, Schena, Federico, Bortoli, Laura, and Robazza, Claudio. 'Action monitoring through external or internal focus of attention does not impair endurance performance.' In: *Frontiers in Psychology* 10 (2019), p. 535. DOI: 10.3389/fpsyg.2019.00535.

[570] Vlahovic, Sara, Suznjevic, Mirko, Pavlin-Bernardic, Nina, and Skorin-Kapov, Lea. 'The effect of VR gaming on discomfort, cybersickness, and reaction time.' In: *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2021, pp. 163–168.

[571] Vortmann, Lisa-Marie, Kroll, Felix, and Putze, Felix. 'EEG-based classification of internally-and externally-directed attention in an augmented reality paradigm.' In: *Frontiers in human neuroscience* 13 (2019), p. 348. DOI: 10.3389/fnhum.2019.00348.

[572] Vortmann, Lisa-Marie and Putze, Felix. 'Attention-aware brain computer interface to avoid distractions in augmented reality.' In: *Extended abstracts of the 2020 chi conference on human factors in computing systems*. 2020, pp. 1–8.

[573] Vortmann, Lisa-Marie and Putze, Felix. 'Exploration of Person-Independent BCIs for Internal and External Attention-Detection in Augmented Reality.' In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5.2 (2021). DOI: 10.1145/3463507.

[574] Vortmann, Lisa-Marie, Schult, Moritz, Benedek, Mathias, Walcher, Sonja, and Putze, Felix. 'Real-time multimodal classification of internal and external attention.' In: *Adjunct of the 2019 Int. Conference on Multimodal Interaction*. 2019, pp. 1–7.

[575] Vozzi, Alessia, Ronca, Vincenzo, Aricò, Pietro, Borghini, Gianluca, Sciaraffa, Nicolina, Cherubino, Patrizia, Trettel, Arianna, Babiloni, Fabio, and Di Flumeri, Gianluca. 'The sample size matters: to what extent the participant reduction affects the outcomes of a neuroscientific research. A case-study in neuromarketing field.' In: *Sensors* 21.18 (2021), p. 6088. DOI: 10.3390/s21186088.

[576] Vrana, Scott R. 'The Psychophysiology of Disgust: Differentiating Negative Emotional Contexts with Facial EMG.' In: *Psychophysiology* 30.3 (1993), pp. 279–286. DOI: 10.1111/j.1469-8986.1993.tb03354.x.

[577] Vries, Ingmar EJ de, Slagter, Heleen A, and Olivers, Christian NL. 'Oscillatory control over representational states in working memory.' In: *Trends in cognitive sciences* 24.2 (2020), pp. 150–162. DOI: 10.1016/j.tics.2020.12.004.

[578] Wacharamanotham, Chat, Eisenring, Lukas, Haroz, Steve, and Echtler, Florian. 'Transparency of CHI research artifacts: Results of a self-reported survey.' In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–14. DOI: 10.1145/3313831.3376448.

[579] Wagener, Nadine, Bentvelzen, Marit, Dänekas, Bastian, Woźniak, Paweł W., and Niess, Jasmin. 'VeatherReflect: Employing Weather as Qualitative Representation of Stress Data in Virtual Reality.' In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. DIS '23. Pittsburgh, PA, USA: Association for Computing Machinery, 2023, pp. 446–458. DOI: 10.1145/3563657.3596125.

[580] Walker, Francesco, Dey, Debargha, Martens, Marieke, Pfleging, Bastian, Eggen, Berry, and Terken, Jacques. 'Feeling-of-Safety Slider: Measuring Pedestrian Willingness to Cross Roads in Field Interactions with Vehicles.' In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–6. DOI: 10.1145/3290607.3312880.

[581] Wall, Conor, Celik, Yunus, Hetherington, Victoria, McMeekin, Peter, Walker, Richard, Graham, Lisa, Vitorio, Rodrigo, and Godfrey, Alan. 'Considering and understanding developmental and deployment barriers for wearable technologies in neurosciences.' In: *Frontiers in Neuroscience* 18 (2024), p. 1379619. DOI: 10.3389/fnins.2024.1379619.

[582] Wang, Bing-Shun, Wang, Xiao-Jin, and Gong, Li-Kun. 'The construction of a Williams design and randomization in cross-over clinical trials using SAS.' In: *Journal of statistical software* 29 (2009), pp. 1–10. DOI: 10.18637/jss.v029.c01.

[583] Wang, Chiu-Hsuan, Chen, Bing-Yu, and Chan, Liwei. 'RealityLens: A User Interface for Blending Customized Physical World View into Virtual Reality.' In: *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. UIST '22. Bend, OR, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3526113.3545686.

[584] Wang, Jue, Yan, Nan, Liu, Hailong, Liu, Mingyu, and Tai, Changfeng. 'Brain-computer interfaces based on attention and complex mental tasks.' In: *Digital Human Modeling: First Inter. Conference on Digital Human Modeling*. Springer. 2007, pp. 467–473. DOI: 10.1007/978-3-540-73321-8_54.

[585] Wang, Nanjia and Maurer, Frank. 'A Design Space for Single-User Cross-Reality Applications.' In: *Proceedings of the 2022 Int. Conference on Advanced Visual Interfaces*. AVI 2022. Frascati, Rome, Italy: ACM, 2022. DOI: 10.1145/3531073.3531116.

[586] Wang, Rui I., Pelfrey, Brandon, Duchowski, Andrew T., and House, Donald H. 'Online Gaze Disparity via Bioncular Eye Tracking on Stereoscopic Displays.' In: *2012 Second Int. Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*. 2012, pp. 184–191. DOI: 10.1109/3DIMPVT.2012.37.

[587] Wang, Yijun, Gao, Xiaorong, Hong, Bo, Jia, Chuan, and Gao, Shangkai. 'Brain-computer interfaces based on visual evoked potentials.' In: *IEEE Engineering in medicine and biology magazine* 27.5 (2008), pp. 64–71.

[588] Wascher, Edmund, Hoffmann, Sven, Sänger, Jessica, and Grosjean, Marc. 'Visuo-spatial processing and the N1 component of the ERP.' In: *Psychophysiology* 46.6 (2009), pp. 1270–1277.

[589] Wascher, Edmund, Rasch, Björn, Sänger, Jessica, Hoffmann, Sven, Schneider, Daniel, Rinkenauer, Gerhard, Heuer, Herbert, and Gutberlet, Ingmar. 'Frontal theta activity reflects distinct aspects of mental fatigue.' In: *Biological psychology* (2014). DOI: `https://doi.org/10.1016/j.biopsycho.2013.11.010`.

[590] Welch, Peter. 'The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms.' In: *IEEE Transactions on audio and electroacoustics* 15.2 (1967), pp. 70–73. DOI: `10.1109/TAU.1967.1161901`.

[591] Wenzel, Markus A, Golenia, Jan-Eike, and Blankertz, Benjamin. 'Classification of eye fixation related potentials for variable stimulus saliency.' In: *Frontiers in neuroscience* 10 (2016), p. 23. DOI: `10.3389/fnins.2016.00023`.

[592] Wickens, Christopher D. 'Multiple Resources and Performance Prediction.' In: *Theoretical Issues in Ergonomics Science*. Vol. 3. 2. 2002, pp. 159–177.

[593] Wickham, Hadley. tidyverse: Easily Install and Load the Tidyverse. 2021.

[594] Wilhelm, Frank H, Pfaltz, Monique C, Gross, James J, Mauss, Iris B, Kim, Sun I, and Wiederhold, Brenda K. 'Mechanisms of virtual reality exposure therapy: The role of the behavioral activation and behavioral inhibition systems.' In: *Applied psychophysiology and biofeedback* 30.3 (2005), pp. 271–284. DOI: `10.1007/s10484-005-6383-1`.

[595] Wilkinson, Michael, Pugh, Zachary H., Crowson, Aaron, Feng, Jing, Mayhorn, Christopher B., and Gillan, Douglas J. 'Seeing in slow motion: manipulating arousal in virtual reality.' In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 63. Los Angeles, CA: SAGE Publications Sage CA, 2019, pp. 1649–1653. DOI: `10.1177/1071181319631512`.

[596] Williams, Mark G., Mathews, Andrew, and MacLeod, Colin. 'The emotional Stroop task and psychopathology.' In: *Psychological bulletin* 120.1 (1996), p. 3. DOI: `10.1037/0033-2909.120.1.3`.

[597] Winslow, Brent, Carpenter, Angela, Flint, Jesse, Wang, Xuezhong, Tomasetti, David, Johnston, Matthew, and Hale, Kelly. 'Combining EEG and eye tracking: using fixation-locked potentials in visual search.' In: *Journal of Eye Movement Research* 6.4 (2013). DOI: `10.16910/jemr.6.4.5`.

[598] Witmer, Bob G and Singer, Michael J. 'Measuring presence in virtual environments: A presence questionnaire.' In: *Presence* 7.3 (1998), pp. 225–240.

[599] 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. Vancouver, BC, Canada: ACM, 2011, pp. 143–146. DOI: 10.1145/1978942.1978963.

[600] Wobrock, Dennis, Finke, Andrea, Schack, Thomas, and Ritter, Helge. 'Using fixation-related potentials for inspecting natural interactions.' In: *Frontiers in human neuroscience* 14 (2020), p. 447.

[601] Wolfe, Jeremy M. 'Approaches to visual search: Feature integration theory and guided search.' In: *The Oxford handbook of attention* 11 (2014), pp. 35–44.

[602] Wolfe, Jeremy M. 'Visual search: How do we find what we are looking for?' In: *Annual review of vision science* 6 (2020), pp. 539–562. DOI: 10.1146/annurev-vision-091718-015048.

[603] Wolfe, Jeremy M and Horowitz, Todd S. 'Five factors that guide attention in visual search.' In: *Nature Human Behaviour* 1.3 (2017), p. 0058. DOI: 10.1038/s41562-017-0058.

[604] Wolfe, Jeremy M and Van Wert, Michael J. 'Varying target prevalence reveals two dissociable decision criteria in visual search.' In: *Current biology* 20.2 (2010), pp. 121–124. DOI: 10.1016/j.cub.2009.11.066.

[605] Wolpert, Daniel M. and Landy, Michael S. 'Motor control is decision-making.' In: *Current Opinion in Neurobiology* 22.6 (2012), pp. 996–1003. DOI: 10.1016/j.conb.2012.05.003.

[606] Woodman, Geoffrey F and Luck, Steven J. 'Electrophysiological measurement of rapid shifts of attention during visual search.' In: *Nature* 400.6747 (1999), pp. 867–869. DOI: 10.1038/23698.

[607] Wright, Rex A. 'Refining the prediction of effort: Brehm's distinction between potential motivation and motivation intensity.' In: *Social and Personality Psychology Compass* 2.2 (2008), pp. 682–701. DOI: 10.1111/j.1751-9004.2008.00093.x.

[608] Wulf, Gabriele and Lewthwaite, Rebecca. 'Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning.' In: *Psychonomic bulletin & review* 23 (2016), pp. 1382–1414.

[609] Xu, Qi, Zhou, Hui, Wang, Yongji, and Huang, Jian. 'Fuzzy support vector machine for classification of EEG signals using wavelet-based features.' In: *Medical engineering & physics* 31.7 (2009), pp. 858–865.

[610] Yang, Chenguang, Ye, Yuhang, Li, Xinyang, and Wang, Ruowei. 'Development of a neuro-feedback game based on motor imagery EEG.' In: *Multimedia Tools and Applications* 77 (2018), pp. 15929–15949.

[611] Yin, Jie, Arfaei, Nastaran, MacNaughton, Piers, Catalano, Paul J, Allen, Joseph G, and Spengler, John D. 'Effects of biophilic interventions in office on stress reaction and cognitive function: A randomized crossover study in virtual reality.' In: *Indoor air* 29.6 (2019), pp. 1028–1039. DOI: https://doi.org/10.1111/ina.12593.

[612] Yin, Ruiyang, Wang, Depeng, Zhao, Shufang, Lou, Zheng, and Shen, Guozhen. 'Wearable sensors-enabled human–machine interaction systems: from design to application.' In: *Advanced Functional Materials* 31.11 (2021), p. 2008936. DOI: 10.1002/adfm.202008936.

[613] Ysehak Abay, Tomas, Shafqat, Kamran, and Kyriacou, Panayiotis A. 'Perfusion changes at the forehead measured by photoplethysmography during a head-down tilt protocol.' In: *Biosensors* 9.2 (2019), p. 71. DOI: 10.3390/bios9020071.

[614] Zagermann, Johannes, Pfeil, Ulrike, and Reiterer, Harald. 'Measuring cognitive load using eye tracking technology in visual computing.' In: *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization*. 2016, pp. 78–85.

[615] Zakopoulos, Nikos A, Kotsis, Vasilios T, Pitiriga, Vassiliki Ch, Toumanidis, Savvas T, Lekakis, John P, Nanas, Serafim N, Vemmos, Kostas N, Stamatelopoulos, Stamatios F, and Moulopoulos, Spyridon D. 'White-coat effect in normotension and hypertension.' In: *Blood pressure monitoring* 7.5 (2002), pp. 271–276. DOI: https://doi.org/10.1097/00126097200210000-00004.

[616] Zeagler, Clint. 'Where to Wear It: Functional, Technical, and Social Considerations in on-Body Location for Wearable Technology 20 Years of Designing for Wearability.' In: *Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ISWC '17. Maui, Hawaii: Association for Computing Machinery, 2017, pp. 150–157. DOI: 10.1145/3123021.3123042.

[617] Zhang, Lili and Murdoch, Michael J. 'Color matching criteria in augmented reality.' In: *Journal of Perceptual Imaging* 1.1 (2018), pp. 10506–1. DOI: 10.2352/J.Percept.lmaging.2018.1.1.010506.

[618] 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. New Orleans, LA, USA: Association for Computing Machinery, 2019, pp. 1151–1159. DOI: 10.1145/3332165.3347869.

List of Figures

List of Tables

List of Acronyms

AI	Artificial Intelligence
ANOVA	Analysis of Variance
AR	Augmented Reality
ART	Aligned Rank Transform
AV	Augmented Virtuality
BCI	Brain Computer Interface
CAR	common average reference
CAVE	Cave Automatic Virtual Environment
CER	Corrected Error Rate
CPI	Personal Informatics
ECG	Electrocardiography
EDA	Electrodermal Activity
EEG	Electroencephalography
EMG	Electromyogram
ERP	Event-Related Potential
ET	Eye Tracking
FIR	Finite Impulse Response
FKA	First Key Accuracy
FRP	Fixation-Related Potentials
GEQ	Game Experience Questionnaire
GLMM	Generalized Linear Mixed Model
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HR	Heart Rate
HRV	Heart Rate Variability
IAF	Individual Alpha Frequency
ICA	Independent Component Analysis
IP	Internet Protocol
IPA	Index of Pupillary Activity
ISI	Inter-Stimulus Interval
LDA	Linear Discriminant Analysis
LEEP	Head-Mounted Displays
LSL	Lab Streaming Layer
ML	Machine Learning
MR	Mixed Reality
NPC	Not-Player Charachter
NSSCR	Not-Specific Skin Conductance Response
NTP	Network Time Protocol

OSF	Open Science Framework
PD	Distractor Positivity
PPG	Photoplethysmography
RANSAC	RANdom SAmple Consensus
REML	Restricted maximum likelihood
RT	Reaction Time
SC	Skin Conductance
SCL	Skin Conductance Level
SCR	Skin Conductance Response
SNR	Signal-To-Noise-Ratio
T2LF	Time to Last Fixation
TCP	Transmission Control Protocol
TER	Total Error Rate
UDP	User Datagram Protocol
UI	User Interface
UX	User Experience
VR	Virtual Reality
WM	Working Memory
WPM	Words per Minute

A
APPENDIX

This thesis is composed of mine and my co-authors original thoughts an comments. As writing aids, I employed DeepL [1] to translate English text into German. I used ChatGPT[2] primarily for ideation.

[1] https://www.deepl.com/en/

[2] https://chatgpt.com/

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 15. Juli 2024

Francesco Chiossi