APPLICATION OF EMPIRICAL AND ANALYTICAL USABILITY STUDIES TO THE USER-CENTRED DEVELOPMENT OF EXTENDED REALITY (XR) SOFTWARE.

Bálint Szabó[a], Rita Ladjánszky[a], Fruzsina Pataki-Bittó[a], Sarolta Tóvölgyi[a], Kata Kapusy[a]

a Budapest University of Technology and Economics, Department of Ergonomics and Psychology, Faculty of Economic and Social Sciences, Hungary
* szabo.balint@gtk.bme.hu

DOI: 10.63442/JUUK1300

KEYWORDS | EXTENDED REALITY (XR), WEB 4.0, USABILITY TESTING, HEURISTIC EVALUATION

ABSTRACT | The theoretical section of the paper outlines the progression of the Internet from the Web 1.0 era to the speculative notion of Web 4.0. The latter is referred to as the symbiotic web, wherein the prevailing solutions such as Artificial Intelligence (AI), Internet of Things (IoT), Virtual Reality (VR) and Augmented Reality (AR) developments serve as the framework for advancements that establish a harmonious connection between the user and the software product. Web 4.0 is expected to demonstrate unique characteristics, specifically proactive, content-exploring, self-learning, collaborative, and content-generating solutions (such as extended reality (XR)). XR is an umbrella term derived from the VR-AR continuum, and as an emerging concept, it can revolutionise how people perceive and engage with their environment. The paper evaluates a prototype XR software product designed for fashion model agencies through empirical and analytical usability studies. The research aims to thoroughly understand the user requirements for forthcoming immersive environments, encompassing insights from potential users and experts in user experience (UX) design. To achieve this goal, we conducted a Rapid Iterative Testing and Evaluation (RITE) usability study with a sample of potential users (n=13). We assessed the system's overall usability and gathered subjective feedback by administering a System Usability Scale (SUS) questionnaire and conducting a follow-up interview. Besides that, a usability inspection method was applied, where UX specialists (n=6) were engaged within the framework of heuristic evaluation. Similarly, in this instance, the heuristic evaluation was accompanied by a subsequent interview, in which the experts imparted their subjective experiences and professional insights to us. The user-centred studies have resulted in diverse forms of feedback. The RITE test identified several minor usability issues, such as the absence of software affordances and familiar interactions. On the other hand, the heuristic test has highlighted suggestions for long-term improvements, primarily focusing on the ergonomic design of the XR space across various sub-dimensions. We systematically compare and contrast usability observations from empirical and analytical studies, demonstrating the practical value of applying these methodologies in developing XR software products. Consequently, this paper provides a comprehensive analysis of the results derived from the different types of usability studies that could be useful for creating and developing future immersive spaces.

1. Introduction

Information technology (IT) has rapidly advanced, breaking barriers of time and space and enabling limitless communication. This transformation affects our lives, including work, leisure, thoughts, actions, and related media consumption or software usage (He et al., 2021; Megyeri & Szabó, 2024).

The Internet is the central force of today's digital revolution, acting as a powerful and unstoppable network that connects people across borders and time zones, serving as a worldwide hub for communication, broadcasts, and information sharing. It has evolved from a simple information-sharing tool to a dynamic social interface connecting individuals across changing physical and virtual environments (Tarnoff, 2022).

The evolution of the internet has been remarkable, advancing from the static nature of Web 1.0 to the promising potential of the upcoming Web 4.0 through significant progress. Every advancement has redefined software user interfaces, expanding the limits of what can be achieved in the digital domain (Aghaei et al., 2012).

The continuous evolution of information technology (IT) across the successive epochs of Web 1.0, Web 2.0, and Web 3.0 has profoundly influenced different software's user experience (UX). The transformations of actual UX trends have been driven by a gamut of technological innovations (e.g., radio frequency identification (RFID), mobile technology, virtual and augmented reality (VR/AR) devices, and artificial intelligence (AI)). Such advancements have heralded the arrival of novel and pioneering software across diverse industrial sectors (Khaleel Ibrahim, 2021).

The current path of advancement in IT suggests that several noticeable trends are expected to dominate the latest developments in the software market. Among these trends are the burgeoning domains of big data analytics, machine-to-machine (M2M) communication, cloud computing infrastructures, a variety of AI-driven products, and the burgeoning arena of extended reality (XR) innovations (Hendler, 2009). XR solutions combine augmented reality (AR), virtual reality (VR), and mixed reality (MR) to improve our interactions with the internet and the environment (Bharadiya, 2023).

These developments increasingly constitute integral facets of cutting-edge IT initiatives that contribute to the emergence of the fourth-generation Internet era, Web 4.0. Web 4.0 symbolises an envisaged frontier wherein the fusion between cyberspace and the tangible realm undergoes further augmentation. With the support of numerous emerging IT solutions, this synergistic convergence is ready to give users more customised, intelligent, and immersive experiences, offering designers new UX challenges and opportunities (Pfeiffer et al., 2016).

This paper presents a user-centred approach employed to enhance XR software pertaining to Web 4.0 by comprehending user and UX experts' requirements and insights.

2. Theoretical Background

2.1 From Static Web to Metaverse

Tim Berners-Lee conceptualised the World Wide Web in 1989 as a continuously developing system for organising information (Machado et al., 2019). In the past thirty years, the Internet has evolved significantly, progressing through different stages known as Web 1.0, Web 2.0, and Web 3.0, with the first signs of Web 4.0 currently in progress. Various eras can be distinguished by unique characteristics that have impacted the UX of the associated software (Eftekhari et al., 2011; Hung et al., 2010; Khaleel Ibrahim, 2021).

In the Web 1.0 era (1989-2004), known as the cognition web, the online space resembled a static, read-only environment. Information dissemination dominated, and user interaction was limited, marked by one-way information flows. Developers wielded significant influence on the design of webpages, focusing on the logical structure and visual meaning, thus creating a cohesive UX (Benford et al., 2009; Choudhury, 2014; Fleming & Koman, 1998).

Transitioning into Web 2.0 from 2004, dubbed the communication web, user participation became the focal point. Coined as the read-write web, this phase facilitated user-generated content through social networking, blogging, and interactive web applications. Web 2.0 spurred advancements in responsive design and dynamic content, enhancing user interactivity and shaping a more engaging, multi-model UX (Law & Van Schaik, 2010; Pranay & Chhetri, 2015). Web 3.0, the Web of Collaboration, was introduced in 2013 and inspired by John Markoff's vision in 2006. The fundamental concept of executive or intelligent web involves improving the ability to understand and process information effectively. Machine-readable data is integrated with metadata to allow software to interpret and analyse web content, resulting in an intelligent and contextual UX. This evolution aimed to enhance machine systems' ability to understand data and aid in decision-making, thereby enhancing the creation of immersive software (Hiremath & Kenchakkanavar, 2016).

The rise of the Web of Integration (Web 4.0) signifies a new paradigm in the IT sector, marked by the integration of various models, technologies, and social interactions. This new version of the internet represents a significant departure from previous versions, introducing a more extensive and interconnected structure. Web 4.0 is a vision of the internet that involves increased personalisation and contextual awareness. The discussed idea envisions a seamless integration of physical and digital realms, where interconnected devices, sophisticated analytics, and intelligent algorithms collaborate to offer customised and intuitive metaverse services (Almeida, 2017).

The metaverse is a concept associated with Web 4.0, representing a globally interconnected virtual 3D environment accessed through the internet, enabling individuals from different locations to engage in shared social experiences. A metaverse integrates the physical and digital realms, allowing users to engage with virtual content in their surroundings. Digital information and virtual objects are tied to particular physical locations, enabling users to access and engage with them in real-world situations (Pangilinan et al., 2019).

For example, utilising XR technologies and incorporating intelligent algorithms (e.g. AI) in a virtual setting can provide consumers with immersive and realistic experiences (Ahuja et al., 2023). It allows digital content to be superimposed onto the real world, creating immersive experiences and new opportunities for communication, collaboration, and information access. Users can navigate and engage with digital content in real time in their actual surroundings, allowing interactive UX in fields like gaming, education, design, and retail (Kuhn & Ballatore, 2015).

2.2 Definition of Extended Reality (XR)

Milgram & Kishino (1994) introduced a fundamental visual representation to demonstrate the range of immersive experiences, encompassing the concept of XR (Figure 1). The AR-VR continuum, illustrated in their diagram, provides an elementary basis for comprehending the different levels of reality-virtuality integration. The virtuality continuum represents a range of actual and virtual environments at opposite extremes (Ribeiro et al., 2021).

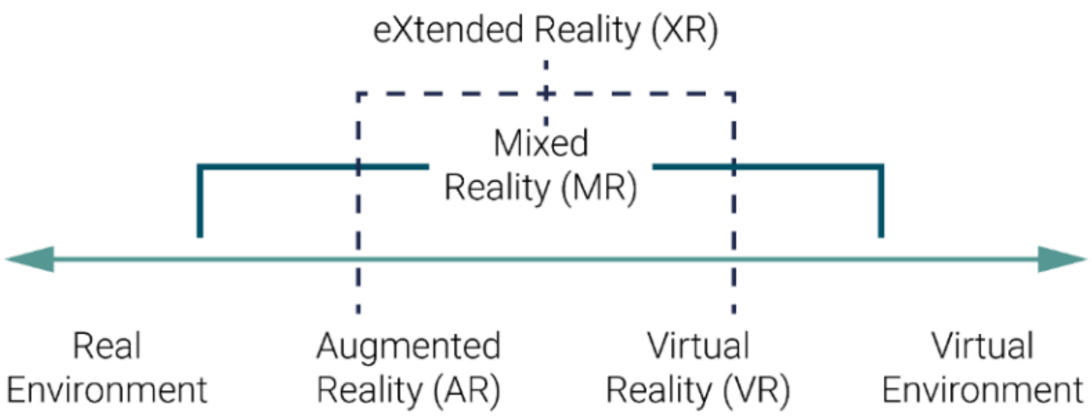

Figure 1. XR adaptation of Milgram and Kishino's virtuality continuum.

AR has a historical lineage dating back to the 1950s; nevertheless, the term gained prominence primarily in the 1990s (Berryman, 2012). AR enhances the user's perception of the environment by superimposing digital information, focusing on a more realistic experience. In subsequent years, AR has undergone multifaceted characterisations, with its central essence commonly construed as a dynamic and real-time amalgamation of digital information and the physical environment (Billinghurst et al., 2015; Javornik, 2016). The evolving nature of AR's conceptualisation underscores its capacity to seamlessly integrate digital content with the perceptible world, indicative of its continual evolution as an interdisciplinary field bridging computer science and human-computer interaction (Fjeld, 2004).

Since World War II, the military has utilised simulators to train combat pilots. The Link Trainer application, used for pilot training before World War II, was one of the earliest instances of a technology similar to modern VR (Jeon, 2015). Virtual reality (VR) encompasses an extensive collection of perceptual experiences generated via the application of various technological devices (Boas, 2013). VR leans towards the virtual reality end, completely involving humans in computer-generated settings and disconnecting them from their immediate physical surroundings. The concept pertains to a synthetic domain in which individuals strive for complete immersion, thus facilitating their active engagement and exposure to many occurrences (Sveistrup, 2004).

Defining the concept of MR is challenging (Rokhsaritalemi et al., 2020). Scientific studies have shown that even experts in the field need help to determine what defines MR. Speicher, Hall, & Nebeling (2019) conducted expert interviews with various academic and industry stakeholders to elucidate the main characteristics of the concept of MR. Based on the research findings, it can be stated that MR is commonly associated with the terms AR and VR. Sala (2021) defines MR as a medium that combines physical and virtual environments through immersive, computer-generated settings. This concept can be easily grasped when considering Milgram and Kishino's (1994) perspective.

As an overarching term, XR encapsulates the entire spectrum that Milgram & Kishino (1994) delineated. XR transcends the confines of AR and VR, encompassing a continuum of immersive experiences from augmented real-world environments to entirely virtual ones (Ziker et al., 2021).

XR usually acts as an inclusive umbrella term, acknowledging the interplay and convergence of virtual and real-world elements. Nonetheless, researchers have discussed the significance and relevance of the letter X in terms of XR. One perspective argues that the X is inaccurately interpreted as extended, as VR does not extend reality but substitutes it. X in the proposed interpretation can be replaced with the actuality utilised in the development, similar to an algebraic variable. It can be interpreted as future developments, which refer to the unknown reality that lies ahead (Rauschnabel et al., 2022)

Lee, Moon, Ko, Lee, & Yoo (2020) state that the concept of XR, which combines VR and AR, is becoming increasingly popular as it revolutionises how humans interact with the digital and physical realms, providing users with novel perspectives and methods of engaging with their surroundings. As a result of its radical redesign of the software user interface, XR introduced unknown challenges, which were noted in the discourse of UX experts.

The primary focus of UX for screen-based software was optimising the user interface to meet the users' needs and expectations and essential business key performance indicators (KPIs), such as conversion rates. Meanwhile, in XR environments, where all senses are fully exposed to artificial and immersive stimuli, the directness and impressive potency of the UX impose an entirely new set of obligations on designers. Exploring the user's emotions and gaining their trust is essential when designing the user interface of XR software (Hillmann, 2021).

3. Aim of the Research and Methodology

The research aims to improve an XR software program prototype for modelling agencies through empirical usability testing and analytical assessment.

The research plan was designed to answer the following research questions:

- What are the user needs for designing XR spaces?
- What are the similarities and differences between empirical and analytical user feedback?

3.1 Introduction of the Prototype of the XR Software

The software prototype is a high-fidelity product developed by a Hungarian IT company specifically for the Meta Quest 2 device. The software enables fashion companies to view models from an existing model agency, assisting decision-makers in selecting future collaboration models in an XR environment. Using a head-mounted display, fashion professionals can pre-screen and select models without travelling, saving them considerable time and effort (Figure 2).

Figure 2. The Meta Quest 2 head-mounted display and the XR fashion model selection software handles.

The main innovation of the development is that users can interact with UI elements using their hands instead of relying on a control tool to navigate in an immersive environment. This is made possible by the hand-tracking feature, which utilises software to track and map hands (Figure 3).

Users can view the models' presentations and catwalks in various environments ("Rooftop" and "Studio"). Besides that, models can be selected as favourites and then emailed to the decision-makers at the model agency.

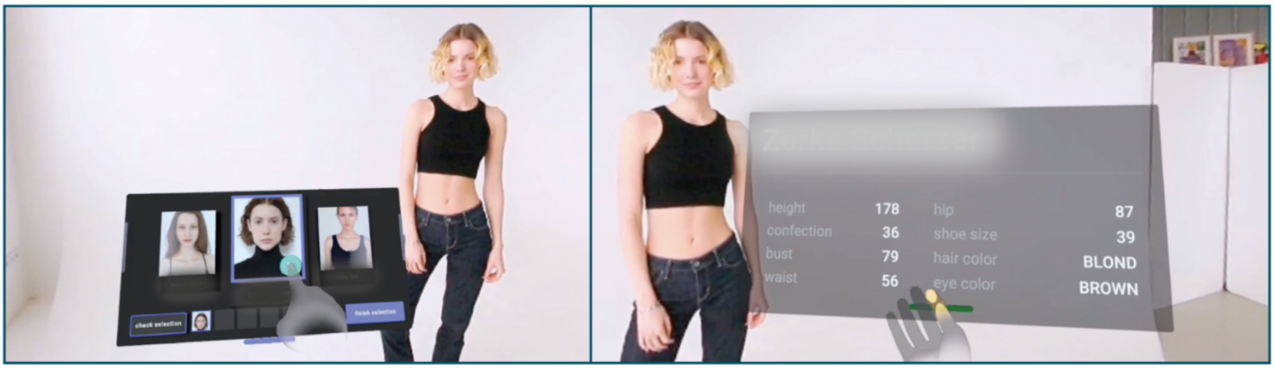

Figure 3. Hand-tracking control and parts of the XR fashion model XR software's user interface.

3.2 Research Methods

The study was consecutively divided into two qualitative sections in the spring of 2023. The first was an empirical usability study (n=13), using a pre-designed study to identify usability issues in the XR environment (Svedberg, 2020). Besides that, we supplemented the usability testing methodology with a System Usability Scale (SUS) questionnaire to assess the software quantitively by the participants (Grier et al., 2013). To gain a deeper understanding of user expectations, the experiment concluded with a follow-up interview in each case.

In the second phase, we used an analytical approach as a heuristic evaluation (n=6), in which digital product design experts offered comments on the specific software product and the XR environment (Becker & Freitas, 2023). The tests were conducted this way because, when designing an innovative product, it was vital to get feedback from potential consumers and experts to steer the prototype in the right direction.

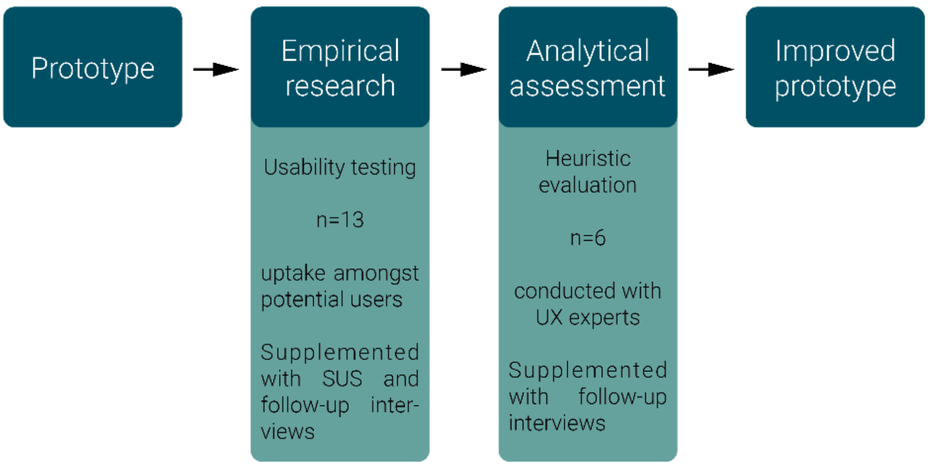

Figure 4. Research plan.

Methodological details of the empirical research:

- Usability testing

Usability testing is a qualitative approach that utilises user testing to assess various software products (Issa & Isaias, 2022). The method examines how well a product can be used for its intended purpose and evaluates whether specific product functions are met. The methodology assesses if an interface is usable by its intended users to carry out the tasks it was created for (Rubin et al., 2008).

We conducted the test using Rapid Iterative Testing and Evaluation (RITE). RITE is a usability test variation that emphasises quickly identifying and implementing solutions rather than pinpointing problems (Medlock et al., 2005; Szabó & Hercegfi, 2023). After conducting usability testing with a small group of participants, we refined the software based on user feedback for the subsequent testing phase. This ensured that modifications were made using a user-centred and iterative approach.

The usability study was carried out in three phases with varying numbers of participants: the first with four, the second with five, and the third with four individuals. Most participants were university students at the BA/BSc, MA/MSc, or PhD level, aged between 20 and 34 years (Table 1).

The development team had a two-week period to implement the modifications identified during the iterative phase before the next round of testing.

Table 1. Demographic information on empirical research participants grouped by iterative phase.

Iterative phase	1	2	3
Number of participants	4	5	4
Identification numbers	ID1, ID2, ID3, ID4	ID5, ID6, ID7, ID8, ID9	ID10, ID11, ID12, ID13
Age range	20-23	20-24	23-34
Gender distribution	Male (100%)	Male (20%), Female (80%)	Male (75%), Female (25%)
Studies	mechatronics engineer, electrical engineer, computer engineer, international management	industrial designer, psychology	computer vision, oriental languages and cultures, molecular bionics, building services engineer

During the usability test, participants were required to solve model selection tasks and scenarios outlined in predetermined descriptions. The scenarios outlined specific business challenges faced by a senior employee at a fashion agency, allowing participants to easily relate to them and creating a clear decision-making situation. To obtain a more comprehensive understanding of users' requirements, the think-aloud protocol was applied to understand the user's cognitive processes during their interaction with the product. (Jääskeläinen, 2010). Continuously using this moderation technique throughout the usability testing would likely introduce bias in the immersive environment. We created the protocol with concurrent probing (CP), a moderation technique where participants respond to questions from the facilitator at different stages of the task solution (Usability.GOV, 2013).

- System Usability Scale (SUS)

After usability testing, the research progressed by administering a System Usability Scale (SUS) questionnaire to quantitatively assess the software's usability (Bangor et al., 2008). SUS is a widely used user-centred method in scientific research and industrial practice to measure subjective usability. The questionnaire was primarily designed to assess the overall usability of complex systems and is also suitable for evaluating XR software products. The questionnaire evaluates the user's views on the system's usability using a five-point Likert scale and ten statements. The System Usability Scale (SUS) consists of an equal number of positively and negatively worded statements that are used to derive a percentage score ranging from 0 to 100 (Lewis, 2018).

- Follow-up interviews

After finishing the SUS questionnaire, participants could express their opinions about the XR software in follow-up interviews. We inquired about their cognitive processes during the tasks, enjoyment of the software, and motivation to use it. The interviews explored the participants' experiences with hand tracking as they used it to operate the software programme.

Aside from the visual aspects, the interview questions focused on immersion, precisely the arrangement of software elements in the XR environment and evaluating the auditory features (model presentations and the audiovisual effects) of XR fashion shows.

Methodological details of the analytical assessment:

- Heuristic evaluation

Heuristic evaluation is a commonly used method in industrial UX practice where a small group of experts assess the user interface's suitability for the design problem based on established ergonomic guidelines (Delice & Güngör, 2009). When evaluating screen-based software products, evaluators frequently employ established guidelines (e.g., Shneiderman, Nielsen) to assess the user interface (Delice & Güngör, 2009). Given the unique features of the XR environment, experts were asked to conduct a comprehensive and expert evaluation of the XR environment, openly addressing all relevant UX factors. In the absence of predetermined guidelines and checklists, evaluators possess the capacity to approach interface evaluation with a flexible mindset (Nielsen, 1994). This approach facilitates the identification of novel issues or opportunities that might remain overlooked when employing rigid criteria. Consequently, evaluators are empowered to leverage their creativity and domain expertise, enabling a holistic assessment of the interface encompassing aspects such as aesthetics, overall functionality, and other pertinent factors influencing UX. Heuristic evaluation thus prioritises the collection of comprehensive feedback concerning the efficacy of an interactive system, eschewing the mere enumeration of issues based on specific usability principles (Hertzum & Jacobsen, 2001). During the heuristic evaluation, we also implemented the think-aloud technique to enable participants to rapidly articulate their professional thoughts (Alshammari et al., 2015).

Six individuals with extensive experience in UX design or previous VR/AR projects participated in the study (Table 2).

Table 2. Demographic information on analytical assessment participants.

Identification number	ID1	ID2	ID3	ID4	ID5	ID6
Age	43	30	31	48	43	40
Gender	Male	Female	Female	Female	Male	Male
Experience in years regarding UX	8	1	9	7	20	24
Current work	Researcher on VR development in human resources (HR)	Junior researcher of an interaction and immersive research group	Senior researcher in the field of human-computer interaction (HCI)	Freelance web designer and developer	Mentor of designers and product design managers	UX consultant and design team mentor

- Follow-up interviews

After the heuristic evaluation, experts participated in a follow-up interview to discuss their subjective experiences and professional insights. We also sought to understand the participants' perceptions of the XR environment, its influence on them, and the emotions it elicited. We also requested their input on the professional suitability of the hand-tracking feature and any ergonomic challenges they faced. We asked the experts to provide information about themselves, including their current activities, areas of interest, research field, and experiences.

4. Results

4.1 Empirical Research Results

The results of the empirical research are presented in three iterative phases, following the characteristics of usability testing, which was applied to the RITE methodology.

- Usability testing results (iteration phase 1)

During iteration phase 1, the model selection process proceeded without any issues. When asked to select a particular model, comparable results were noted. The majority of the participants found this task easy to solve. One considered the "Add to favourites" option not the default choice for a final selection but a preliminary selection for further consideration. The text "Add to favourites" confused some of the participants as they expected a different function: "I thought it would lead to a separate list of favourites, like on clothes shop websites." (ID4)

Multiple users experienced a sudden limitation in their field of view at the periphery of their visual boundaries while using the software, leading to discomfort: “I feel confined and restricted, like being in a matchbox.” (ID1)

The initial pop-up window provided instructions on hand-tracking control, but the model selection area appeared in the background, leading to confusion among the participants during the onboarding process. During usage, participants found it beneficial to observe the outline of their hand while manipulating and moving through the XR environment using hand tracking, eliminating the need for extra information. They claimed that onboarding not, but affordances and past physical life experiences aided their intuitive understanding of control.

The hand-tracking technology was criticised solely for its sensitivity. The participants found it frustrating when the software failed to accurately detect the users' intentions. A moderator had to help a participant activate the button he wanted to press during an assignment. The issue arose from the user's lack of perceived tactile feedback in the software, leading to uncertainty in interacting with the button: “I struggled because I had only lightly touched the button.” (ID3)

ID3 was also unsure about interacting with different elements during the usability testing due to his lack of familiarity with the hand-tracking mode of operation. Despite its challenges, the hand-tracking control was seen as a positive feature of the XR software.

Participants encountered difficulties multiple times while in the XR environment, failing to find tutorial videos or additional onboarding support during usage to overcome the problem.

The models' proximity particularly unsettled participants. Therefore, they looked for traits that would allow them to stop the models in space. During the "Introduction" mode, they struggled to understand and analyse the models' speech due to the audio recording quality.

The participants reported that the user experience was greatly improved by the highly realistic way the models moved and the presence of various locations. The rooftop view was described as "dizzying," suggesting it was authentic and made the participants feel like they were there.

Participants expressed a positive impression of the XR software during follow-up interviews, describing it as "futuristic" and "very exciting." They verified the recognised usability issues, such as onboarding difficulties and hand-tracking limitations. They expected a bright future for XR software in different sectors, including home decoration, driver's licence acquisition, long-distance travel, virtual amusement parks, and online conference participation, with possible cost reductions in a business setting. Some participants emphasised the need for increased realism in the XR environment and technological advancements to improve the user experience.

- Usability testing results (iteration phase 2)

Participants did not encounter navigational difficulties when selecting models in iteration phase 2, but some needed help establishing the selection criteria. Participants remarked that the XR environment was blurry, especially towards the edges. As a result, many participants needed help understanding the characteristics of the models.

Some criticised the software's user interface for having elements that were too closely positioned and being interrupted by disruptive pop-up windows.

Participants rapidly left the "Rooftop" view due to distractions caused by the white background behind the models, prompting several users to quickly switch to the "Studio" view because of feelings of spatial anxiety.

Participants in Iteration 2 were concerned about models approaching too closely in catwalk mode, prompting them to seek software solutions to prevent this.

The low audio recording quality made understanding and analysing the models' introduction challenging: “I was annoyed by the models' constant speech, so I tried to keep it under control.” (ID6)

Participants' perceptions of being within the XR environment space varied. Some participants were dissatisfied as they could not experience detachment from reality: "I could see through the bottom of the glasses, so it was less realistic.” (ID5)

In addition, some participants in the usability testing expressed concern about being in a physical environment: "I was afraid of hitting the real wall." (ID7)

The hand-tracking feature generated mixed reactions from the participants. All of them were irritated since the software often failed to react to button pushes, and it was unclear how to click. Nonetheless, they enjoyed it due to the novel and unusual experience: "It was interesting to press into the space and interact with something." (ID9)

Several participants, however, claimed that calculating the proper distances to reach the buttons took much work. Some said that activating the buttons with one finger was complicated and annoying (both found it easier to contact the button with two fingers). Some participants recommended creating an onboarding solution that includes all manipulation operations to overcome navigation issues.

Participants maintained a positive impression of the XR software during the interviews in iteration 2, describing it as "great," "exciting," and "particularly interesting." Interviewees appreciated the realistic environment and the opportunities for controller-free interactions. Feedback was generally positive, although some struggled with pop-up windows. The importance of onboarding was highlighted. Participants predicted increased adoption of XR software for design, medical applications, office work, entertainment, and business meetings but emphasised the need for improved immersion.

- Usability testing results (iteration phase 3)

In iteration 3, the model selection process was relatively easy. However, there were visual disturbances in the XR environment, with the space shaking and blinking at the start. Several individuals have raised concerns regarding the image quality of the stationary models.

The revised panel layout implementation after iteration phase 2 was successful, leading to enhanced clarity and transparency in the user interface.

Pressing buttons continued to present difficulties due to the need for tactile feedback. Furthermore, the system's slow performance impeded the software's navigation.

The participants found the software's response time slow and were unsure why they couldn't press a button or activate a function. Participants also complained about the lack of clarity in the system's feedback: “The response of the overall environment is painfully sluggish.” (ID11)

Frequently, the pop-up panels were positioned too close to the user, causing the text to be difficult to read. Most people instinctively reacted by stepping back: "I want to move away”. (ID10)

The "Rooftop" perspective caused users to lose their sense of space, leading them to prefer the "Studio" view: “The studio perspective was more realistic.” (ID13)

Hand-tracking generated negative feedback during this iterative phase, most likely as a result of the slower system speed of the improved prototype: "The hand control is unusual and odd. I don’t like it." (ID12)

Participants in the interviews expressed positive assessments of the design of the XR space, particularly appreciating its cleanliness and intuitive structure. They recommended improvements like increasing the size of action buttons and text and adding arrows for straightforward navigation. The need for more onboarding information and irritating pop-up windows were also mentioned. Some users expressed dissatisfaction with the display quality and the lack of background music. Participants acknowledged the software's potential for academic and recreational purposes, foreseeing a promising future but recognising the necessity for additional improvements.

- System Usability Scale (SUS) results

The SUS questionnaires validate the findings of the usability studies. During iterative phase 1, participants assessed the software product's usability with a mean score of 77.5% and a standard deviation (SD) of 6.12. In iterative phase 2, the mean was 76% (SD=11.81); in iterative phase 3, it was 78.75% (SD=4.33).

The high SUS values were attributed to the participant's ability to appreciate the novelty of the software. The XR Environment and the innovative hand-tracking control overcame several minor issues identified for improvement in the usability evaluation.

4.2 Analytical Assessment Results

The results of the analytic examination are provided alongside the follow-up interviews, in which the experts primarily re-emphasized, clarified, and concluded the issues raised throughout the inquiry.

- Heuristic evaluation results

The experts' overall product assessment was optimistic based on the heuristic tests.

All experts felt confused throughout the onboarding process since photos were utilised to demonstrate non-interactive capabilities. As a result, most of them raced through this portion without thoroughly reading the directions.

During model selection, there were several instances where participants mistakenly pushed the selection (+/-) button while scrolling (swiping) between models, deleting or adding models to the selection.

The models' proximity distracted everyone in the area, leading users to step back: "I have to step back to see the legs." (ID2)

The aesthetics of the panels provoked various emotions among the participants. They were all pleased with how the content was organised, but using colour, contrast, proportions, and alignment within the panels was controversial.

Moving the panels was a common challenge. Some couldn't recall how to move them, while others decided not to make the appropriate motion. Many people saw the panels as static, preferring to adapt their stance to their surroundings.

At the end of the selection process, everyone was frustrated and unclear about the "done" feedback, which appeared interactive but produced no change or input. Moreover, it was necessary to clarify what actions to take and whether the selection had been accurately submitted.

Hand tracking received a wide range of reviews, with some considering it immature and complicated, while others found it simple and enjoyable to use.

> "What I found more challenging than my prior controller experience was the absence of tactile input when I hit the button. It made it much more difficult to understand how hard to press was. It's also unusual to see a panel and occasionally have to press behind it." (ID2)

In other situations, the issue was that they found it challenging when they attempted to break from the tool's usage patterns (2D tools). Manual navigation includes all interactions inside the software; even specialists reported a high sense of uneasiness among users about whether they were doing something wrong. When a gesture did not work the first time, many people would try a more extensive interface, palm or multiple fingers, and maybe double- or many consecutive clicks to remedy the problem.

> "I was trying to pick up patterns of behaviour that we use on 90% of devices, so it was a bit weird for me to find out ... that I was touching it and moving it, but it was cool."(ID4)

Overall, participants indicated good feelings about the XR program and constructive remarks concerning the space's architecture and increased realism.

> "The direction is nice from a UX standpoint. I believe it tries too hard to mimic real-life, realistic model choices while failing to realise the space's possibilities completely. It's a good thing there aren't any hidden panels because I believe they detract from the added real-world potential that comes from the fact that the area is completely changeable. I suppose people are seated in a typical model selection, but I can also envision them moving about, etc." (ID6)

5. Discussion

The applied research methods proved effective, and the empirical and analytical studies have revealed a relevant set of user needs that can support the user-centred development of the XR prototype.

The onboarding part of the software is likely to receive greater attention and engagement if it were interactive. To inform users what actions they can take in the XR environment, additional information should be added to the software features during the onboarding process (e.g., introduction of models, catwalk function, steps of the selection process).

After learning how to move the panels, everyone was eager to customise the interface. In addition, individual panel layouts should be saved, and it would be helpful to demonstrate this to users during onboarding.

After customising the user interface, the specified location of the pop-up windows caused issues in all testing to varied degrees. This was due to information showing behind other windows outside the perimeter. Because some of these elements were fixed, the users had to adjust their location relative to the panel to proceed.

As the users perceived the introduction to allude to a spoken speech, this feature proved to be deceptive. Therefore, it would make sense to incorporate a more direct and personal introduction to the models into the software. While it is true that the selection process is impersonal in real-time, the models might be placed in the XR environment closer to the decision-makers, for example, by providing optionally displayable extra information about their residence, hobbies, and modelling experience, among other things.

Users also felt they could not control the software because of the instantaneous, self-running video of the models. The models' inconsistent stopping points at the end of walks, which caused the movements to terminate in the middle of the sequence and gave the appearance of a software freeze, proved confounding in a few instances.

The demand for personal space is also evident in the XR environment because the models come uncomfortably close to the users. The design should ensure that the users do not feel that any part of the user interface is disturbing their personal space.

Users found the hand-tracking feature innovative and intriguing, yet they believed it required improvement.

Making the software more user-friendly would include translating components from ordinary life and reality into immersive environments. These solutions may include lifting the panel with two hands, like lifting a sheet of paper, or even a virtual board, like a corkboard, on which you could use tweezers to place pictures of particular models. (Figure 5).

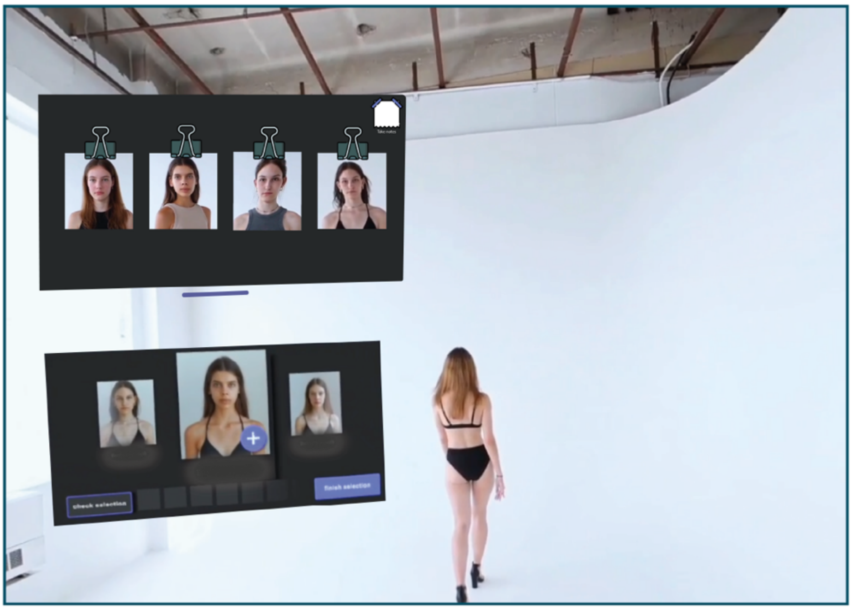

Figure 5. An example of using real-life boards and tweezers to increase the realism of the workflow in the XR environment.

The "done" caption at the end of the selection process was troublesome and caused doubt and annoyance, significantly reducing the UX. The system did not indicate that clicking on it would prompt the user to take action.

These insights will significantly help customise the software and XR environment to meet the user's needs. In the future, educating potential users through well-designed onboarding will be essential for users of XR software to be informed about, utilise, and benefit from the opportunities provided by the immersive environment.

6. Conclusion

In terms of the applied methodologies, participants in the empirical research found a notably more significant number of design flaws (such as distracting background colour and smaller button sizes) that were easily fixed but had no impact on UX. Of course, similar observations were also made during the analytical assessment. Still, the experts were concerned about more profound factors that could impact UX, such as overcrowding, consistency problems, and recurring errors in design (e.g. inadequate location of pop-up windows, lack of feedback).

The empirical research highlighted the lack of guidance and training on correct gestures, particularly about software usage. Furthermore, the analytical study discovered other significant results. These insights involved integrating components that were already present in various applications and real-world scenarios. The analysis confirms that XR software design should follow traditional UX design standards (e.g. Shneiderman's guidelines). The heuristic assessment found instances where it was suggested that the XR environment should allow the user to take control of the conversation, such as following traditional design principles. Users should be able to start and stop the video manually, and a progress indicator should display their position on the timeline.

The models selected in several cases were chosen randomly, and the analytical assessment offered a sophisticated improvement proposal to address this. The experts emphasised that errors could be reduced by placing the action button that triggers the selection in a separate bar from where users intended to scroll. Moreover, using the familiar check box instead of the selection option can help streamline the process. The participants in the empirical study made many general suggestions about expanding or limiting the selection options but did not offer any specific advice.

Aside from optimising load times and durations in different use cases, focusing more on specific aspects like hand gestures, visual prompts, movement, and sound design in XR environments is crucial.

Experts assert that allowing the user to have control over the design of their XR environment and the ability to add panels to customise it is crucial.

Creating software versions for different age groups is necessary due to the product's compatibility with diverse users and their significant usage habits. One way to achieve this is by incorporating various features to help users unfamiliar with technology, such as providing thorough onboarding, implementing hand tracking, and including traditional features like a drop-down menu. These characteristics should be optional or deactivated once abilities and routines are established.

Moreover, the analytical assessment also showed an improvement in ecological validity. Examples include integrating functionalities from different software, such as note-taking, video scrolling, and profile settings, and implementing real-life gesture recognition, like lifting a panel with two hands like a sheet of paper. Contact with other individuals in the XR area and realistic mapping of real-life movements, like sitting on a couch, changing perspectives, and exploring models, have been proposed as essential.

It can be said that both empirical and analytical studies have provided useful feedback. The fundamental difference is that the analytical insights were more comprehensive and thus can be used to design any XR circular design so that their consideration highly enhances the UX of the software. In contrast, the empirical results gave concrete development ideas for concrete software. The joint observation involved enhancing non-design parts (such as controls and audio-based sub-elements) to enhance the overall realism of the environment (Table 3).

Table 3. Summarising the results of empirical and analytical usability studies

Empirical research	Analytical assessment
Design flaws (e.g. distracting background colour, smaller button sizes)	Overall design insights (e.g. overcrowding, consistency, recurring design errors)
Inappropriate visualisation, location and grouping of functions	Ability to customise panels and settings, furthermore creating software versions
Absence of onboarding	Developing product learnability (beyond onboarding)
Greater emphasis on certain aspects of the whole environment (e.g. hand gestures, visual prompts, movement, sound design)	

6.1 Limitations

Consequently, when creating an XR software product, an analytical assessment is recommended after developing a high-fidelity prototype in the initial iteration. After considering the suggestions of analytical assessment and constructing the product, empirical usability testing is a valuable tool.
The empirical research should have a 2-week development duration to implement the development ideas effectively. One limitation of our study was that the participants fell within the age range of 20 to 34. Thus, we could not collect the viewpoints and experiences of other potential users.

We could not hire fashion industry experts because of the significant financial implications. Upon product completion, it would also be advantageous to gather their feedback.

The experts who participated in the analytical assessment included some with less experience in UX (e.g. ID2) or no experience in VR/AR development (e.g. ID4). This may have slightly affected the results.

References

Aghaei, S., Nematbakhsh, M. A., & Farsani, H. K. (2012). Evolution of the World Wide Web: From Web 1.0 to Web 4.0. *International Journal of Web & Semantic Technology (IJWesT), 3*(1), 1–10. https://doi.org/10.5121/ijwest.2012.3101

Ahuja, A. S., Polascik, B. W., Doddapaneni, D., Byrnes, E. S., & Sridhar, J. (2023). The digital metaverse: Applications in artificial intelligence, medical education, and integrative health. *Integrative Medicine Research, 12*(1), 100917. https://doi.org/10.1016/j.imr.2022.100917

Almeida, F. L. (2017). Concept and dimensions of Web 4.0. *International Journal of Computers & Technology, 16*(7), 7040–7046. https://doi.org/10.24297/ijct.v16i7.6446

Alshammari, T., Alhadreti, O., & Mayhew, P. (2015). When to ask participants to think aloud: A comparative study of concurrent and retrospective think-aloud methods. *International Journal of Human Computer Interaction, 6*(3), 48–64.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Becker, A., & Freitas, C. M. D. S. (2023). Evaluation of XR applications: A tertiary review. *ACM Computing Surveys, 56*(5), 1–35. https://doi.org/10.1145/3626517

Benford, S., Giannachi, G., Koleva, B., & Rodden, T. (2009). From interaction to trajectories: Designing coherent journeys through user experiences. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 709–718. https://doi.org/10.1145/1518701.1518812

Berryman, D. R. (2012). Augmented reality: A review. *Medical Reference Services Quarterly, 31*(2), 212–218. https://doi.org/10.1080/02763869.2012.670604

Bharadiya, J. P. (2023). Artificial intelligence and the future of web 3.0: Opportunities and challenges ahead. *American Journal of Computer Science and Technology, 6*(2), 91–96. https://doi.org/10.11648/j.ajcst.20230602.14

Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. *Foundations and Trends® in Human–Computer Interaction, 8*(2–3), 73–272. https://doi.org/10.1561/1100000049
Boas, Y. (2013). Overview of virtual reality technologies. *Interactive Multimedia Conference*, 1–6.

Choudhury, N. (2014). World Wide Web and its journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Information Technologies (IJCSIT), 5*(6), 8096–8100.

Delice, E. K., & Güngör, Z. (2009). The usability analysis with heuristic evaluation and analytic hierarchy process. *International Journal of Industrial Ergonomics, 39*(6), 934–939. https://doi.org/10.1016/j.ergon.2009.08.005

Eftekhari, M. H., Barzegar, Z., & Isaai, M. T. (2011). Web 1.0 to web 3.0 evolution: Reviewing the impacts on tourism development and opportunities. In F. V. Cipolla Ficarra, C. de Castro Lozano, E. Nicol, A. Kratky, & M. Cipolla-Ficarra (Eds.), *Human-Computer Interaction, Tourism and Cultural Heritage* (Vol. 6529, pp. 184–193). Springer. https://doi.org/10.1007/978-3-642-18348-5_17

Fjeld, M. (2004). Usability and collaborative aspects of augmented reality. *Interactions, 11*(6), 11–15. https://doi.org/10.1145/1029036.102904

Fleming, J., & Koman, R. (1998). *Web navigation: Designing the user experience*. O'Reilly Media.

Grier, R. A., Bangor, A., Kortum, P., & Peres, S. C. (2013). The system usability scale: Beyond standard usability testing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 187–191. https://doi.org/10.1177/154193121357104

He, W., Zhang, Z. J., & Li, W. (2021). Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic. *International Journal of Information Management, 57*, 102287. https://doi.org/10.1016/j.ijinfomgt.2020.102287

Hendler, J. (2009). Web 3.0 emerging. *Computer, 42*(1), 111–113. https://doi.org/10.1109/MC.2009.30

Hertzum, M., & Jacobsen, N. E. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction, 13*(4), 421–443. https://doi.org/10.1207/S15327590IJHC1304_05

Hillmann, C. (2021). *UX for XR: User experience design and strategies for immersive technologies*. Springer.

Hiremath, B. K., & Kenchakkanavar, A. Y. (2016). An alteration of the web 1.0, web 2.0 and web 3.0: A comparative study. *Imperial Journal of Interdisciplinary Research, 2*(4), 705–710.

Hung, W. T., Tsang, S. S., & Liu, H. Y. (2010). Website characteristics and the impact of user perceived value on user behavior in Web 2.0. *Journal of Management Information and Decision Sciences, 13*(1), 1.

Issa, T., & Isaias, P. (2022). Usability and human–computer interaction (HCI). In *Sustainable design: HCI, usability and environmental concerns* (pp. 23–40). Springer.

Jääskeläinen, R. (2010). Think-aloud protocol. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (pp. 371–372). John Benjamins Publishing Company. https://doi.org/10.1075/hts.1.thi1

Javornik, A. (2016). Augmented reality: Research agenda for studying the impact of its media characteristics on consumer behaviour. *Journal of Retailing and Consumer Services, 30*, 252–261. https://doi.org/10.1016/j.jretconser.2016.02.004

Jeon, C. (2015). The virtual flier: The link trainer, flight simulation, and pilot identity. *Technology and Culture*, 28–53. https://doi.org/10.1353/tech.2015.0017

Khaleel Ibrahim, A. (2021). Evolution of the Web: From Web 1.0 to 4.0. *Qubahan Academic Journal, 1*(3), 20–28. https://doi.org/10.48161/qaj.v1n3a75

Kuhn, W., & Ballatore, A. (2015). Designing a language for spatial computing. Springer. https://doi.org/10.1007/978-3-319-16787-9_18

Law, E. L.-C., & Van Schaik, P. (2010). Modelling user experience–An agenda for research and practice. *Interacting with Computers, 22*(5), 313–322. https://doi.org/10.1016/j.intcom.2010.04.006

Lee, Y., Moon, C., Ko, H., Lee, S.-H., & Yoo, B. (2020). Unified representation for XR content and its rendering method. *The 25th International Conference on 3D Web Technology*, 1–10. https://doi.org/10.1145/3424616.3424695

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human–Computer Interaction, 34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Machado, L. M. O., Souza, R. R., & da Graça Simões, M. (2019). Semantic web or web of data? A diachronic study (1999 to 2017) of the publications of Tim Berners-Lee and the World Wide Web Consortium. *Journal of the Association for Information Science and Technology, 70*(7), 701–714. https://doi.org/10.1002/asi.24111

Medlock, M. C., Wixon, D., McGee, M., & Welsh, D. (2005). The rapid iterative test and evaluation method: Better products in less time. In *Cost-justifying usability* (pp. 489–517). Elsevier. https://doi.org/10.1016/B978-012095811-5/50017-1

Megyeri, M., & Szabó, B. (2024). Investigating the effectiveness of user onboarding solutions with eye tracking: A case study on Paint 3D. *Ergonomics in Design, 32*(1), 1–8. https://doi.org/10.1177/1064804621102602

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems, 77*(12), 1321–1329.

Nielsen, J. (1994). Heuristic evaluation. In *Usability inspection methods* (pp. 25–62).

Pangilinan, E., Lukas, S., & Mohan, V. (2019). *Creating augmented and virtual realities: Theory and practice for next-generation spatial computing*. O'Reilly Media.

Pfeiffer, T., Hellmers, J., Schön, E.-M., & Thomaschewski, J. (2016). Empowering user interfaces for Industrie 4.0. *Proceedings of the IEEE, 104*(5), 986–996. https://doi.org/10.1109/JPROC.2015.2508640

Pranay, K., & Chhetri, B. (2015). Evolution of World Wide Web: Journey from Web 1.0 to Web 4.0. *International Journal of Computer Science and Technology, 6*(1), 134–138.

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., & Alt, F. (2022). What is XR? Towards a framework for augmented and virtual reality. *Computers in Human Behavior, 133*(March), 107289. https://doi.org/10.1016/j.chb.2022.107289

Ribeiro, R., Ramos, J., Safadinho, D., Reis, A., Rabadão, C., Barroso, J., & Pereira, A. (2021). Web AR solution for UAV pilot training and usability testing. *Sensors, 21*(4), 1456. https://doi.org/10.3390/s21041456

Rokhsaritalemi, S., Sadeghi-Niaraki, A., & Choi, S.-M. (2020). A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences, 10*(2), 636. https://doi.org/10.3390/app10020636

Rubin, J., Chisnell, D., & Spool, J. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Wiley.

Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is mixed reality? *Conference on Human Factors in Computing Systems - Proceedings*, 1–15. https://doi.org/10.1145/3290605.3300767

Svedberg, J. J. (2020). Reality check: A review of design principles within emergent XR artefacts.

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of Neuroengineering and Rehabilitation, 1*, 1–8. https://doi.org/10.1186/1743-0003-1-10

Szabó, B., & Hercegfi, K. (2023). User-centered approaches in software development processes: Qualitative research into the practice of Hungarian companies. *Journal of Software: Evolution and Process, 35*(2), 1–28. https://doi.org/10.1002/smr.2501

Tarnoff, B. (2022). *Internet for the people: The fight for our digital future*. Verso Books.

Usability.GOV. (2013). Running a usability test. https://www.usability.gov/how-to-and-tools/methods/running-usability-tests.html

Ziker, C., Truman, B., & Dodds, H. (2021). Cross reality (XR): Challenges and opportunities across the spectrum. In J. Ryoo & K. Winkelmann (Eds.), *Innovative learning environments in STEM higher education* (pp. 55–77). Springer. https://doi.org/10.1007/978-3-030-58948-6_4

About the Authors:

Bálint Szabó, assistant professor at Budapest University of Technology and Economics, delves into user-centered software development in his doctoral dissertation. Specializing in Human-Computer Interaction (HCI) research, he focuses on eye-tracking methodologies.

Rita Ladjánszky, an industrial design engineering student at Budapest University of Technology and Economics. Researching XR products and software usability testing methodologies, she earned top honors at the 2023 Scientific Student Conference and a Pro Progressio fellowship for her work.

Pataki-Bittó Fruzsina, assistant professor at Budapest University of Technology and Economics, focuses on employee-centered development of office work environments. She leads the Future of Work community within EELISA, collaborating with Hungarian and international students on future workplace projects.

Sarolta Tóvölgyi, assistant professor and Erasmus coordinator at Budapest University of Technology. Her doctoral focus is on ergonomics-driven product management for blood group serology equipment. She excels in user-centric product development across various physical products.

Kata Kapusy, assistant professor at Budapest University of Technology and Economics, explored consumer purchasing behavior in second-hand clothing shopping for her doctoral dissertation. With a focus on human factors, she delves into in-store and user-centered XR software research.

Acknowledgements: We would like to thank the students involved in the product-user interaction for their assistance in carrying out the empirical investigation. We would also like to express our gratitude to the UX Budapest group members for their enthusiastic participation in the analytical investigation.

P/REFERENCES OF DESIGN

This contribution was presented at Cumulus Budapest 2024: P/References of Design conference, hosted by the Moholy-Nagy University of Art and Design Budapest, Hungary between May 15-17, 2024.

Conference Website
cumulusbudapest2024.mome.hu

Conference Tracks
Centres and Peripheries
Converging Bodies of Knowledge
Redefining Data Boundaries
Bridging Design and Economics
Speculative Perspectives
The Power of Immersion
The Future of Well-being
Taming Entropy: Systems Design for Climate and Change
Ways of Living Together
Cumulus PhD Network

Full Conference Proceedings
https://cumulusbudapest2024.mome.hu/proceedings

ISBN Volume 1: 978-952-7549-02-5 (PDF)
ISBN Volume 2: 978-952-7549-03-2 (PDF)

DOI Volume 1: https://doi.org/10.63442/IZUP8898
DOI Volume 2: https://doi.org/10.63442/TADX4016

Conference Organisers
Moholy-Nagy University of Art and Design Budapest (MOME)
mome.hu
Cumulus Association
cumulusassociation.org