Augmenting or Immersing? The Impact of Technology Choice on Cognitive Load and Enjoyment in Escape Rooms

Vivien Landgrebe
LMU Munich
Munich, Germany
landgrebe.vivien@web.de

Elizabeth Bouquet
University of Television and Film Munich
Munich, Germany
e.bouquet@hff-muc.de

Simon von der Au
University of Television and Film Munich
Munich, Germany
s.vonderau@hff-muc.de

Sven Mayer
TU Dortmund University
Dortmund, Germany
info@sven-mayer.com

Abstract

We investigate the impact of Augmented Reality (AR) and Virtual Reality (VR) on cognitive load and enjoyment in the context of games and immersive experiences. To investigate this issue, we conducted a user study with 60 participants, who played in an escape room either in AR or VR using the same hardware. We measured the mental demands of these technologies both subjectively with standardized questionnaires and objectively with eye-tracking and behavior measurements. Our analysis reveals that the cognitive load in AR is significantly higher than in VR. Moreover, presence and usability were higher in VR despite the same hardware and game used. Thus, our results suggest that AR game designers need to consider increased cognitive load to ultimately enhance user experience and performance. In summary, our findings show the different prerequisites that the two technologies have regarding their impact on cognitive load and enjoyment.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**.

Keywords

human computer interaction, game, escape room, cognitive load

ACM Reference Format:
Vivien Landgrebe, Elizabeth Bouquet, Simon von der Au, and Sven Mayer. 2025. Augmenting or Immersing? The Impact of Technology Choice on Cognitive Load and Enjoyment in Escape Rooms. In *24th International Conference on Mobile and Ubiquitous Multimedia (MUM '25), December 01–04, 2025, Enna, Italy*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3771882.3771887

1 Introduction

Extended Reality (XR), such as Augmented Reality (AR) and Virtual Reality (VR), has been shown to be effective in extending the possibilities of games and immersive experiences [17, 53]. Both bring unique options to extend the game flow. While AR extends the physical world by presenting virtual objects in combination with the real world [1, 7], VR overlays an entirely new world on top of the real world. This brings unique opportunities but may come with its own challenges for XR game designers. The user experience is a key factor in determining the success of XR. A high cognitive load (CL) can lead to frustration, inefficiency, and ultimately to a reduction in user experience [28]. The impact of CL on immersive experiences due to technological use is yet unclear.

Previous work has focused on various aspects of AR and VR, including their potential applications, benefits, and limitations. Studies have shown that both can enhance user immersion and presence in virtual environments [31, 47, 51]. However, little attention has been paid to the specific impact of AR and VR technologies on coherence [58], CL [9, 21, 58], and enjoyment. Thus, there is a need for research investigating the cognitive demands of XR environments. While AR and VR are often used for different purposes, the same game can often be played in different form factors (e.g., on a mobile device and/or a large flat screen). The impact of experiencing the same environment in AR or VR on user experience, including coherence, cognitive load, and enjoyment, remains unclear.

We designed and implemented a single-player XR escape room with three puzzles to address this knowledge gap. We implemented the same experience in AR and in VR. For this, we used the same Varjo XR-3 headset to keep the impact of the technology on the experience low. We conducted a between-subjects user study (N=60) to measure the mental demands of these technologies both subjectively with standardized questionnaires and objectively using eye-tracking and behavioral measurements.

Our analysis confirms that presence was higher in VR despite the same hardware and game being used. Combined with the MREQ questionnaire results, we show lower coherence in the AR condition. We attribute this to the split attention between real and digital elements in the game. Moreover, with respect to CL measures, pupil change and task completion time in AR are significantly higher than in VR; however, we could not identify a difference via blink rate and raw NASA-TLX. Finally, we attribute VR to a generally higher enjoyment, but this might be affected by the fact that we used a single-user escape room. Based on our results, we derive general design implications for cognitively demanding applications in XR as well as separate requirements for AR and VR.

2 Related Work

Next, we give a brief overview of existing research that deals with (XR) escape rooms, measuring CL and enjoyment in XR environments, and directly comparing CL and enjoyment in AR and VR. Based on our insights, as well as the concepts of the split-attention effect and processing fluency, we derive three hypotheses that we want to investigate in this paper.

MUM '25, Enna, Italy

ACM ISBN 979-8-4007-2015-4/25/12
https://doi.org/10.1145/3771882.3771887

2.1 (XR) Escape Rooms as Research Tool

Previous research shows that escape rooms have the potential as gamified environments for research [14, 25, 33], education [16, 23, 34, 36, 42, 56], and collaboration [5, 18, 26, 52, 57]. Kleinman and Harteveld [25] state that "escape rooms have much untapped potential as opportunities to research complex problem solving, learning, and reflection in a gamified environment." Moreover, according to Krekhov et al. [29] the amount of research in this area has increased due to the raising popularity of escape games in society. One common practice is to enhance escape room designs with digital aspects [2, 4, 16, 23–25, 36, 43, 44, 52, 56]. Beyond adding digital aspects, researchers have also incorporated XR escape elements into the game rooms [5, 15, 18, 26, 29, 33, 40, 42, 57]. Finally, only a few studies have investigated fully digital escape rooms [14, 34, 61].

In the context of education, the concept "escape room" was tested as learning environment for schools [5, 23, 36, 43], and universities [4, 16, 44, 56]. There, it was primarily used for teaching computer science-related concepts and developing analytical thinking. For example, Back et al. [2] investigated an escape room as an interactive learning experience within a museum exhibition.

Kihara et al. [24], Maragkoudaki and Kalloniatis [34] used escape rooms as a means to provoke an increase in privacy awareness, and Paraschivoiu et al. [42] developed an AR escape room to educate people on how to identify and fight fake news. Holly et al. [15]'s work showed, i.e., how an XR escape room can give insights about the skills of a potentially new co-worker during the hiring process.

As escape rooms are often designed to be solved by groups, there is also research that uses them as motivation for collaboration. While Shakeri et al. [52] used a distributed real-life escape room to gain insights about remote collaboration, Itenge et al. [18], Knoll et al. [26], Warmelink et al. [57] created escape rooms to show their potential for co-located and remote collaboration when utilizing the benefits of AR/MR. Zou et al. [61] showed how an VR escape room can be a place for collaboration that also mitigates ageism.

Furthermore, Madsen et al. [33] studied the fear-inducing effects of different NPCs in an AR horror escape room, while Hirsch et al. [14] examined the emotional connectedness between asynchronous players in an escape room by visualizing the players' heart rate as hearts in VR. Finally, Nikkarikoski et al. [40] presents the possibility of AR offering escape rooms independent of the room size, as they created a dollhouse-size escape room for a tabletop game.

2.2 Cognitive Load in AR and VR Environments

Generally, CL can be measured using a variety of metrics; for a full literature review, see Kosch et al. [28]. The NASA-TLX questionnaire [12] is commonly used to measure overall workload, which includes CL in HCI [11, 21]. Most papers about CL in XR environments use it to measure the participants' CL during their user studies [9, 58, 60]. Kockord and Bodensiek [27] employed an alternative method and utilized various subjective measurement methods, including the presence questionnaire and interviews. Moreover, research utilizes measures to assess CL behavioral metrics, such as time performance, task performance, and error rate (cf. [9, 31, 32]). Furthermore, objective measurements capture data in real-time that can correlate with subjective data [21]; for example, physiological data such as eye-tracking or electroencephalogram (EEG), e.g., Chao et al. [9], Drouot et al. [11], Jeffri and Awang Rambli [21]. Two of them utilized eye-tracking, specifically pupil size and blink rate, as additional measurement instruments [11, 21]. Three mentioned the use of EEG [21], head tracking [58], and a combination of heart rate, galvanic skin response, and heart rate variability [9].

Jeffri and Awang Rambli [21] reviewed correlations between CL and task performance in AR. They show a positive correlation between CL and task performance, meaning that a low CL can lead to better time performance. In an AR assembly task, Drouot et al. [11] compared AR manuals with digital text-based manuals, and the level of CL was equal for both conditions in the complex task variant but not in the simple tasks. Thus, Drouot et al. [11] recommend carefully weighing the advantages of AR. In a similar task, Kockord and Bodensiek [27] found that MR increases the CL in cases where users have little to no experience with MR before. In contrast, Loch et al. [31] show that the AR system has significantly improved task performance as well as error rates compared to video instructions. Chao et al. [9] investigated CL differences in VR compared to regular manuals and multi-media films. While the NASA-TLX results had no significant differences, participants' CL was slightly lower in VR. Additionally, VR showed the best results in terms of time performance and error rate. Luong et al. [32] investigating CL in more complex scenarios through scenarios in which multiple tasks must be performed. They show that CL increases more when addressing two different sensory channels. Xi et al. [58], Zhao et al. [60] investigated CL within a purchasing-related task, directly comparing AR and VR. Zhao et al. [60] showed that CL in AR and VR differed, which mainly depends on their different effects on sensory channels. In a shopping task, Xi et al. [58] found a significantly higher CL when using XR technologies compared to the non-XR condition.

In summary, the literature shows that the CL in an AR scenario tended to rise when compared to other scenarios, such as purely physical scenarios [27], information displayed on a desktop [11], video-based information [31], and even compared to VR [58]. VR technology, on the other hand, had a significantly lower CL compared to text-based content and virtual 2D content [9].

2.3 Enjoyment in AR and VR Environments

Humans generally perceive enjoyment when their intrinsic needs are satisfied [46]. Martin et al. [35] argue that the level of motivation is closely linked to the degree of enjoyment. Enjoyment is also dependent on the perceived sense of presence [19]. Furthermore, Müller et al. [38] emphasize the positive effects of enjoyment on user experience, whereby the latter is closely related to the level of CL [11]. Ou Yang et al. [41] investigated the effects of AR on learning effectiveness and enjoyment of learning. They emphasize that AR can make learning applications more effective and engaging and that AR has great potential for use in the context of learning and teaching. Qin [46] examined the appearing levels of enjoyment and user experience in the mobile AR game Pokémon Go. They found a relation between the degree of enjoyment and the satisfaction of two needs: autonomy and competence. Furthermore, enjoyment showed no notable interconnection with the feeling of presence [46]. In the context of VR gaming, Jang and Park [19] investigated factors to increase willingness to play VR games. They show that extensive interactivity and perceived elaborate control in the application positively affected enjoyment. Furthermore, Jang and Park [19] emphasize that a high engagement can be evoked by rich interactivity and immersion, as well as an appealing user experience design. Mouatt et al. [37] highlighted that high-immersion VR applications achieved more positive results than low-immersion or no VR. With respect to AR and VR, Jung et al. [22], Yu et al. [59] investigated how AR and VR can enhance the enjoyment of visitors in the tourist industry. Within their research, however, AR and VR technologies were always considered as a whole and were not separated. [22] emphasize that AR and VR technologies can improve the visitor experience in tourism-related contexts. This relates, above all, to the satisfaction and enjoyment of tourists. [22] highlight that these effects are mainly caused by high-quality applications rather than low-quality implementations. Finally, Yu et al. [59] argue that AR and VR can foster enjoyment and motivation in general.

2.4 Split-Attention-Effect & Processing Fluency

The emergence of cognitive workload can be explained by the nature and functioning of human memory. According to the Cognitive Load Theory, the split-attention effect can occur during the processing of information [8, 55]. This effect occurs when several possibly interlinked sources of information are displayed separately [55]. This means that the user has to switch their attention between these two sources to process all the presented stimuli. This can raise the number of mental resources required, which ultimately leads to an increase in the extraneous CL. The split-attention effect, therefore, occurs above all when the user has difficulty recognizing the connection between two or more presented interconnected stimuli in the same environment [8]. This may be because they are not designed uniformly enough and are too far apart.

In AR environments, real and virtual objects can differ in their visual representation and can, therefore, be perceived as visual stimuli originating from two distinct sources. Moreover, according to the concept of processing fluency, a break in the consistency of multiple stimuli can lead to difficulties in information processing [45]. A high level of processing fluency can be achieved by providing consistency between the stimulus and its context. This fact shows that it can be problematic for AR applications to provide a high level of processing fluency, as there is a break in the consistency of the environment due to the virtual objects placed in the real world. The difficulty with AR lies in connecting and merging two different worlds, the real and the virtual world, in such a way that the user perceives their environment as consistent. If this is not the case, it can lead to a lower level of processing fluency and thus to an increased CL [50]. In VR, the kind of break in consistency described above cannot exist, as it consists purely of virtual objects.

2.5 Implications and Research Gap

Based on related work, we formulate three hypotheses that we want to investigate with this research. In an AR scenario, the user must categorize virtual and physical objects simultaneously and recognize their connection [58]. Thus, the user's memory is addressed by two different sources, which is also related to the split-attention-effect [55]. Furthermore, this inconsistency in visual representation of these two stimuli makes information processing harder according to the concept of processing fluency [45]. In a well-designed AR application, users should not encounter any difficulties recognizing the common context between the real and virtual environments. This is because both need to be merged seamlessly throughout the whole AR experience [50]. However, this break in consistency is one of the natural characteristics of current AR scenarios, which is not the case in VR. From this, we derive our first hypothesis (**H1**): **The perceived coherence in VR is higher than in AR.**

Furthermore, more cognitive resources are required when people must consider objects presented by two different technologies in the same environment, such as AR and VR, than when considering objects presented by just one [58]. Furthermore, AR tends to evoke higher levels of CL than VR because users need to pay attention to real-world elements and virtual objects simultaneously. Related work provides evidence that AR led to the highest amounts of CL in comparison to non-AR scenarios [11, 27, 31], and even compared to VR [58, 60]. These findings lead us to formulate the following hypothesis (**H2**): **The CL in AR is higher than in VR.**

Our literature review found that both AR and VR have great potential to be enjoyable for their participants; however, there are no direct comparisons regarding their different effects on enjoyment. Thus, there is no verifiable evidence yet as to whether one of the two technologies generally triggers a higher level of enjoyment than the other [22, 59]. Since a high sense of presence has a positive influence on the CL [39] and given our H2, the telepresence in VR is stronger than the local presence in AR. Further, the degree of presence can positively influence enjoyment [19]. In addition, AR applications require more cognitive resources from their users than VR applications [55], which may lead to higher frustration [13] and, consequently, less enjoyment compared to VR. Consequently, we propose our third hypothesis (**H3**): **Participants perceive more enjoyment in VR than in AR.**

With this contribution to our research, we want to make statements regarding CL in XR applications. We will also examine various triggers for CL, including enjoyment and a sense of presence. To the best of our knowledge, we are the first to compare CL between AR and VR using the same HMD for both technologies, and we also provide a direct comparison of their effects on enjoyment.

3 Escape Room Design

To investigate the mental demands that AR and VR each cause, we have developed an escape room that can be played in both technologies. In this section, Section 3.1 will first deal with the story behind the escape room, and in Section 3.2, we will explain in detail the puzzles that had to be solved to win the escape room.

3.1 Story

The fictional story of the escape room revolves around Dr. Noah Johnson, a former university employee who has recently retired. Unfortunately, he forgot to pass the password on to his email inbox, which receives daily emails from students seeking study advice. The player takes on the role of his successor, who must find the credentials to answer them. Therefore, they have to search for clues in the former office of Dr. Johnson that point to the password.

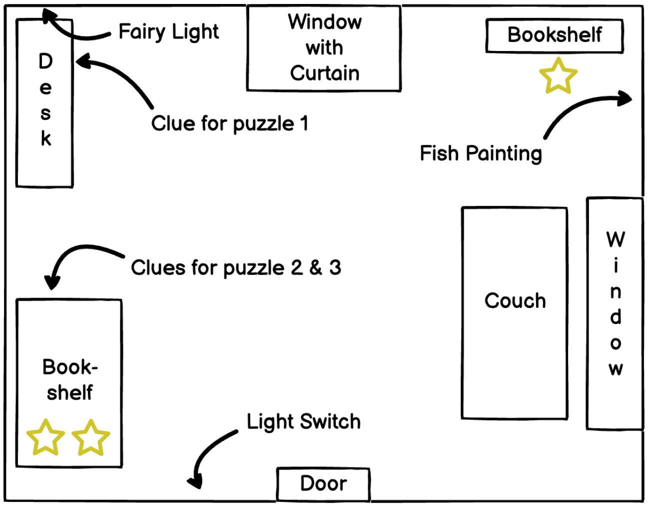


Figure 1: Basic sketch of the escape room with the positions of the props, clues, and stars.

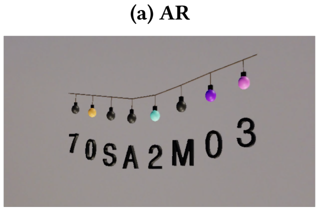
3.2 Puzzles

The player had 10 minutes to solve the escape room. The main theme was related to stars and light, while the task was: "*Follow the Stars!*" To win, the player had to solve three puzzles in the office, which was a rectangular room and consisted of two bookshelves, a desk, two windows, and a sofa, as shown in Figure 1.

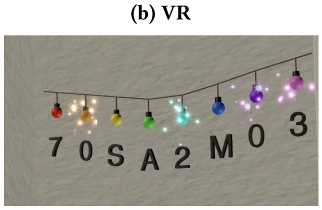
There were also smaller props that the player could gradually discover. In the beginning, the player did not yet see any stars, which meant they had to search the room for them. An abstract painting of a fish hung on the right-hand wall. The fish was a clue that the player should have kept in mind as the game progressed. There was also a fairy light hanging above the desk, which Dr. Johnson had hung there as decoration. This had no function at first. The positions of all the smaller props are also shown in Figure 1.

(a) AR

(b) VR



(c) AR



(d) VR

Figure 2: Components of the puzzle in AR and VR.

3.2.1 Puzzle 1. A paper note was hidden on the desk that said: "*STARt at the window!*" Of the two windows in the room, one of them had closed curtains, while the other was uncovered. If the player opened the curtain in front of one of the windows and looked out, a virtual constellation of stars could be seen, as illustrated in Figure 2. This constellation belonged to the zodiac sign Pisces. Once this constellation was discovered, the first puzzle was completed.

3.2.2 Puzzle 2. There was a card stuck to the bookcase in the room. This card read: "*STARS CAN'T SHINE WITHOUT DARKNESS.*" This implied that the player should switch off the ceiling light by using the light switch. As soon as the room was darker, a virtual starry sky appeared on the ceiling. There were 18 stars in this starry sky, five of which glowed yellow. The remaining thirteen stars were grayed out. The player now had to recreate the constellation that they discovered on the window in Puzzle 1. As this consisted of eight stars, the player had to find the remaining three in the office to complete the constellation in the sky. Two stars could be found inside books placed on the bookshelf on the left-hand side of the room. The third star was located in the top drawer of the bookshelf on the right-hand side of the room. Each of the hidden stars had a fixed position in the constellation in the virtual sky. The grayed-out stars served as placeholders under which the player had to find the correct positions. As soon as the player had properly placed all three stars, the constellation began to twinkle, and the second puzzle was solved.

3.2.3 Puzzle 3. Once Puzzle 2 was solved, the fairy light hanging above the desk was activated. Not all of the differently colored bulbs were flashing, but only four of the eight in total. Furthermore, an eight-digit code appeared under it: "*70SA2MO3.*". The player had to check which characters were under the flashing light bulbs. These were the numbers "*0203,*" which were the first part of the password. On the left-hand bookshelf, there was a picture of a wedding by Dr. Johnson, showing the wedding date. However, only the year "*1991*" was readable, which was the second part of the password. Dr. Johnson's full wedding date is, therefore, March 2, 1991. 1991 has the same astrological ruler as 1956, Dr. Johnson's birth year. On this day in March, the sun is in the constellation of Pisces. This is the connection between the Pisces painting and the star constellation from Puzzles 1 and 2. As soon as the player entered the correct password "02031991" into the computer on the desk, they gained access to the mailbox, and the escape game was won. The exact sequence of the escape room is visualized in Figure 3. Ideally, the puzzles would be solved in this order, but elements from later puzzles could be unlocked without completing earlier ones. Players could even solve Puzzle 2 through trial and error without seeing the constellation template from Puzzle 1. However, solving Puzzle 2 was required to activate the props of Puzzle 3, which ultimately revealed the password.

4 User Study

We implemented our escape room idea in AR and VR. This allows us to investigate our hypotheses. Moreover, we opted for a between-subject design, as the puzzles were identical in both versions. As the puzzles can not be solved twice. In the end, we had 60 players play our escape room game.

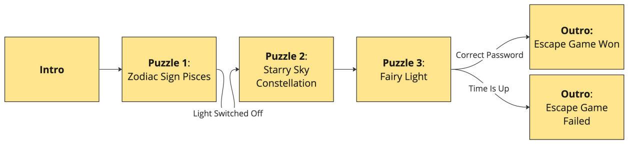


Figure 3: Ideal sequence of puzzles.

4.1 Apparatus

We developed an escape room using the Varjo XR-3 MR headset, which allowed us to use the same device for both AR and VR conditions. This made the results of our user study more comparable.

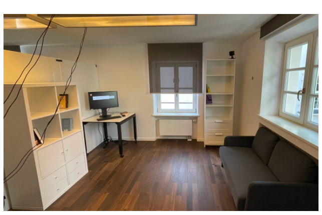
We provided HTC Vive controllers for grabbing virtual objects in both scenarios, utilizing the trigger button. A white ray was displayed from the controller to indicate tangible objects and provide distance visualization. To enhance immersion, we used over-ear headphones to deliver sound effects that responded to users' actions. This created a higher degree of immersion and eliminated distracting environmental noises. In both conditions, the timer for the escape room was shown above the room entry/exit door.

Both experiences took place in the same room. In the AR condition, users interacted with physical props in the room, see Figure 4a. In the VR condition, users interacted with virtual counterparts, which were offset from the real purpose to avoid injuries, see Figure 4b. As a result, the VR play area was smaller (4×3m vs 1.5×2m). The study took place in the same university laboratory.

4.1.1 AR Condition. In AR, participants used only one controller, with which they could grab and move certain virtual objects. The reason for this is that they had to interact with real objects with their other free hand or put the controller down.

We have used markers to integrate virtual objects that were not initially present into the AR escape room. As soon as the headset scanned these in MR, a virtual object appeared. We used a total of seven markers for our application. Some of the objects that appeared were static and drawn to the position of the marker, while others, namely the three stars, could be interacted with after they appeared. The static objects were deactivated once a new marker was scanned, but could be reactivated by scanning the respective marker.

In Puzzle 1, the players had to open a real curtain at the window. A marker was placed behind it, through which the star constellation was represented. For Puzzle 2, a real light switch had to be operated to display the virtual starry sky. The three markers used to activate the missing virtual stars in the constellation were hidden in three different physical books. The password entry in Puzzle 3 was done via a physical desktop PC with a real keyboard and mouse.



(a) AR

(b) VR

Figure 4: The room layout in both conditions.

4.1.2 VR Condition. In VR, two controllers were tracked and displayed to the user as virtual hands. The participants physically walked while using the VR escape room. We implemented the password entry in Puzzle 3 with an input field and a virtual keyboard as UI elements. The virtual objects that appeared in AR by scanning markers were triggered in VR by user interactions. All physical objects from the AR condition were recreated as 3D models in VR to design the two escape rooms as identically as possible.

4.2 Procedure

Upon arrival, we explained the study to the participants, obtained their consent, and they completed their demographic data (see Appendix A). We assigned them randomly to the conditions (AR or VR). To familiarize themselves with the controls of the respective technology before the actual escape game, the test subjects were asked to play a short tutorial. In the tutorial application, they found themselves in the same environment as later in the escape room, but without the puzzle elements. The participants then had to take turns grabbing the two objects and placing them on the table, allowing them to understand the controls. After completing the tutorial, the participants played the escape game in the technology assigned to them. In line with real escape rooms, we set a time limit to complete the room of 10 minutes. After the time was up, participants completed the final questionnaires, and we gave them the option to remove final comments on their general experience of the study in the form of qualitative data.

4.3 Measurements

To investigate our hypotheses, we collected a set of measurements based on related work (see Section 2). CL can be measured using a variety of metrics, including subjective measurement methods such as questionnaires, as well as objective metrics such as user behavior and physical reactions [28].

We asked participants a set of demographic questions (see Appendix A). Moreover, we asked participants to fill in a raw NASA-TLX to measure task load [12], IPQ [6], User Experience Questionnaire (UEQ) [49], and Mixed Reality Experience Questionnaire (MREQ) [48]. Moreover, we recorded the user interactions to extract task performance and task completion time (TCT). In addition, we collected eye-tracking data to understand the users' CL. While pupil dilation can be interpreted as a sign of changes in the CL, a high frequency of eye blinks can reveal an increased CL as well as frustration during task-solving [11, 21].

4.4 Participants

We recruited 60 participants via snowball sampling. Those had an average age of 28.5 years with a standard deviation of 7.6, while 35 identified as female and 25 as male. On average they rated their AR experience ($Q_{Demo}4$) as a 2.8 (SD = 1.6, min = 1, max = 7) VR experience ($Q_{Demo}8$) as a 3.4 (SD = 1.6, min = 1, max = 7) on ca scale from 1 to 7 (*No previous experience* to *Expert*, respectively. With respect to AR and VR usage frequency, they use AR and VR similarly frequently ($Q_{Demo}5$ and $Q_{Demo}9$). For both, most (23) stated that they use it "Once a year or less" and "Never", 19 for AR and 13 for VR. Followed by "Once every few months", 10 for AR and 14 for VR. Finally, 4 used AR more than once per month and 10 used VR more than once per month. Eight participants owned a VR device ($Q_{Demo}10$), comprising five Meta Quest 2 and one Meta Quest 3. Furthermore, the Lenovo ThinkReality VRX glasses were mentioned once, and a not further defined VR headset. 16 people stated that they owned an AR device ($Q_{Demo}6$). Among these, the cell phone (15 times) or tablet (1 time) was mentioned above. Meta Quest 3 was also mentioned once. We compensated participants with 10€ per hour.

5 Results

We employed several measurements to thoroughly and accurately examine the hypotheses.

5.1 Task Performance

For Puzzle 1, AR had a slightly lower completion rate (93.33%) than *VR* with a 100% completion rate. Both conditions performed well on Puzzle 2, with *VR* slightly outperforming *AR* when comparing their success rates of 90% vs. 80%. Since the entire escape room was considered won once Puzzle 3 had been successfully solved, both had the same success rate. In both, the *AR* condition had a lower completion rate (33.33%) than the *VR* condition with a 60% completion rate. Overall, *VR* condition performed better across all puzzles and the whole escape game compared to *AR* condition. Since the puzzles of the escape room were identical in both conditions, it could be assumed that the nature of the technologies used (AR and VR) caused the differences in the performances of the two conditions to be presented here. However, Fisher's Exact Test showed no significant difference ($p = .069$).

5.2 Task Completion Time

For the TCT, we found a significant difference between the two presentations, see Table 1. Here, participants completed the *VR* condition faster than the *AR* condition, see Figure 5. When looking at the individual puzzles, we found that participants finished *Puzzle 1* statistically significantly faster in *VR* than in *AR* (Shapiro–Wilk test: $W = .914$, $p < .001$; Mann–Whitney U test: $W = 668$, $p < .001$). However, there was no difference for *Puzzle 2* (Shapiro–Wilk test: $W = .914$, $p = 0.682$; t-test: $t = -0.992$, $p = 0.326$) and *Puzzle 3* (Shapiro–Wilk test: $W = 0.859$, $p < .001$; Mann–Whitney U test : $W = 97$, $p = 737$).

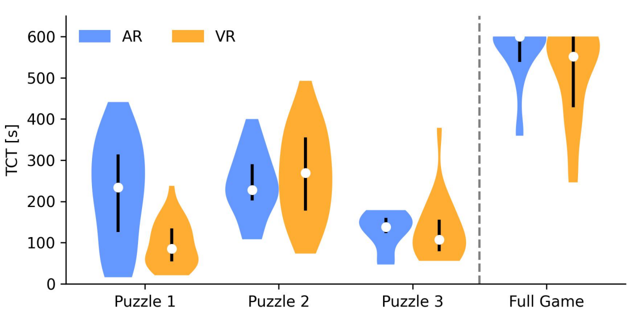
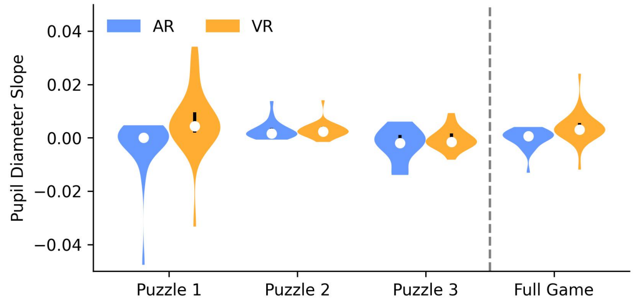
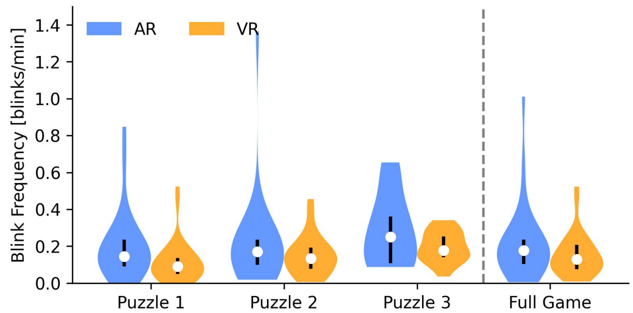


Figure 5: The TCT for the three puzzles.



(a) Pupil Diameters



(b) Blink Rates

Figure 6: Violin plot comparing the means of the a) pupil diameter and b) blink rates of the AR and VR conditions.

5.3 Eye-Tracking Measures

First, we investigated the change in **pupil diameter** as a workload measure. We calculated the linear slope for each completed Puzzle. We ignored incomplete puzzles as they might have given up, and thus, the workload might have dropped as a result. For the entire game experience, a statistically significant difference was observed between the two conditions, as shown in Table 1. The pupil diameter shows more workload in *VR* than in *AR*, see Figure 6a. When looking at the individual puzzles, we found that in *Puzzle 1*, participants had a statistically significant increase in pupil diameter in *AR* than in *VR* (Shapiro–Wilk test: $W = 0.786$, $p < .001$; Mann–Whitney U test: $W = 158$, $p < .001$). However, there was no difference for *Puzzle 2* and *Puzzle 3* (Shapiro–Wilk test: $W = .805$, $p < .001$; Mann–Whitney U test: $W = 257$, $p = .421$ and Shapiro–Wilk test: $W = .805$, $p = .052$; t-test: $t = -0.974$, $p = 0.348$; respectively).

Next, we investigated the *blink frequency* as a workload measure. Here, we did not find a significant difference in the average blink rate across the puzzles, see Table 1 and Figure 6b. However, when investigating the individual puzzles, we found that participants had a significantly higher blink frequency in *AR* in *Puzzle 1* compared to *VR* (Shapiro–Wilk test: $W = 0.760$, $p < .001$; Mann–Whitney U test: $W = 527.5$, $p = .024$), indicating a higher CL in AR than VR. Nevertheless, there was no difference for *Puzzle 2* and *Puzzle 3* (Shapiro–Wilk test: $W = .587$, $p < .001$; Mann–Whitney U test: $W = 348$, $p = .305$ and Shapiro–Wilk test: $W = .879$, $p = .004$; Mann–Whitney U test: $W = 105$, $p = 0.472$; respectively).

Table 1: The statistical results for the comparison *VR* vs *AR*. † If the Shapiro–Wilk test shows data normality, we performed t-tests. ∗ If the Shapiro–Wilk test shows data normality, we used the Vargha and Delaney A effect size measure, else Cohen's d.

	AR			VR			Shapiro–Wilk test		Mann–Whitney U/t test†		
		CI95			CI95						
Measurement	**Median**	**low**	**up**	**Median**	**low**	**up**	**W**	**p**	**Z/t**	**p**	**Effect Size***
TCT	600.	578.26	600.	551.38	513.45	600.	.723	<.001	577.	.042	.641
Pupil Change	.	−.001	.002	.	.002	.005	.861	<.001	176.	<.001	.245
Blink Rate	.18	.12	.23	.13	.09	.17	.742	<.001	488.	.185	.617
raw NASA TLX [12]											
raw NASA TLX	8.67	7.33	9.83	7.42	6.25	8.75	.742	.552	1.956	.055	.505
Mental Demand	11.	6.5	12.	8.5	7.	12.	.925	.001	507.5	.392	.564
Physical Demand	5.	4.	6.5	3.	2.	5.	.897	<.001	545.	.157	.606
Temporal Demand	7.	6.	10.	7.	4.5	10.	.951	.017	468.	.789	.520
Performance	13.	11.	14.	8.5	5.	16.	.935	.003	525.5	.263	.584
Effort	9.	7.5	13.	6.5	4.	10.	.949	.014	579.5	.055	.644
Frustration	4.	1.5	10.	2.	1.	3.5	.834	<.001	560.	.101	.622
UEQ [30]											
Attractiveness	1.75	.50	2.25	2.25	2.	2.67	.858	<.001	272.	.008	.302
Perspicuity	1.5	1.25	2.	2.	1.75	2.25	.898	<.001	335.5	.089	.373
Efficiency	1.	.50	1.62	1.88	1.5	2.12	.917	.001	270.5	.008	.301
Dependability	1.	.62	1.5	1.12	.88	1.38	.960	.048	378.5	.288	.421
Stimulation	2.	1.25	2.25	2.25	1.75	2.62	.900	<.001	339.	.097	.377
Novelty	2.	1.62	2.25	2.25	1.5	2.62	.928	.002	403.5	.489	.448
Pragmatic Quality	1.33	.62	1.62	1.5	1.29	1.92	.942	.007	321.5	.057	.357
Hedonic Quality	1.94	1.44	2.19	2.12	1.75	2.56	.932	.002	366.	.213	.407
IPQ [6]											
Presence	3.35	2.82	3.74	4.22	3.88	4.68	.952	.020	178.5	<.001	.198
Spatial Presence	4.2	3.6	4.8	4.8	4.6	5.2	.897	<.001	279.	.011	.310
Involvement	1.5	1.25	2.	4.	3.38	4.5	.960	.046	92.5	<.001	.103
Experienced Realism	2.75	2.25	3.	2.88	2.38	3.38	.960	.569	−1.912	.061	.494
MREQ [48]											
P(RE)	2.	1.	3.	−1.	−2.5	.50	.864	<.001	749.	<.001	.832
P(VE)	2.	1.5	2.5	3.	2.	3.	.788	<.001	306.	.022	.340
P(RO)	3.	2.	3.	−2.	−3.	−1.	.792	<.001	796.5	<.001	.885
P(VO)	3.	2.	3.	3.	3.	3.	.553	<.001	394.	.302	.438
Usr-RE	2.	1.5	2.5	−1.	−2.	.	.892	<.001	840.5	<.001	.934
Usr-VE	1.	1.	2.	2.	2.	3.	.888	<.001	254.5	.003	.283
Usr-VO	1.	1.	1.5	2.	.	2.	.881	<.001	409.5	.541	.455
VO-VE	3.	2.	3.	3.	2.	3.	.707	<.001	419.	.606	.466
VO-VO	2.	1.	2.	2.	1.	3.	.849	<.001	365.5	.196	.406

5.4 Perceived Workload

For the raw NASA-TLX [12], we found no significant differences, see Table 1. All subscales of the NASA TLX exhibit the same trend, where AR is rated higher than VR; see Figure 7. Thus, the descriptive statistics suggest a slightly higher workload for *AR*, see Figure 7.

5.5 Perceived User Experience

We asked the participants to rate the perceived user experience measured by the UEQ [30]. We found that *VR* generally scores higher on the six sub-scales of the UEQ, see Figure 8. However, we only found statistically significant differences for two measures (Attractiveness, and Efficiency), see Table 1.

5.6 Users' Sense of Presence

We found that the general presence measured by the IPQ scale [6] is significantly different between the two conditions, see Table 1. Moreover, we found that *Spatial Presence* and *Involvement* are significantly higher in the VR condition. However, *Experienced Realism* was not different, see Figure 9.

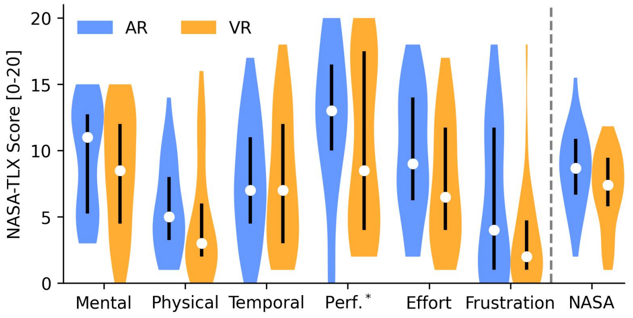


Figure 7: Comparison of the raw NASA-TLX [12] results between the AR and VR conditions. * lower indicates better perceived performance.

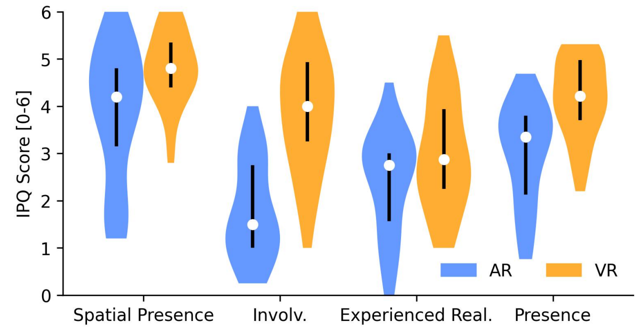


Figure 9: Comparison of the IPQ scale [6] means between the AR and VR conditions.

5.7 Users' Perception of the AR and VR Environments

We found that the perception questions P(RE), P(VE), and P(RO) are all statistically different in the two conditions; however, not P(VO), see Figure 10. As such, in both conditions, virtual objects were perceived as not different. Similarly, the user-perceived actuality questions Usr-RE and Usr-VE are different, but not the interaction with the virtual objects (Usr-VO). Thus, the virtual objects in both conditions had similar perceived distances.

5.8 Users' General Impression

In AR (n=30), 23.3% mentioned positive aspects, for example, "This was really fun!" (P54). With respect to coherence, participants commented that the real and virtual elements were easily distinguishable (P42), and a slight disconnect from the real world was perceived (P36). One participant commented: "I was especially disgusted by the moment we needed to look up. The stars on the ceiling were very hard to interact with because the AR device's weight made the work quite annoying and tedious." (P34). Another test subject felt disturbed by the cables and commented, "It was difficult to focus on the virtual world while trying not to cause any harm to the headset, cables, etc. Hence, naturally, the experience was not fully virtual." (P59). One participant mentioned that the tracking and the motions did not work correctly, which negatively impacted their overall experience (P37). Further, 13,33% submitted suggestions for improvement, like a broader time limit (P39 & P41).

In VR (n=30), 30% of the participants did not provide relevant answers. 40% gave positive feedback about the escape room or the overall experience, like "Interesting and funny experience" (P1). 6.67% mentioned that the headset's cables and its weight disturbed them during the experience. One participant stated that the sound effects via the headphones contributed to the immersion (P25). Another participant reported a high level of immersion, as they wanted to place the physical controllers on the virtual table (P3). This participant further mentioned that the tracking of the controllers did not work reliably and, thus, negatively influenced the experience. In this condition, 10% noted that the escape game could have been longer and more diversified. Another test subject remarked that even though the graphics were convincing, there was still a huge difference when compared to the real world (P9).

Finally, 13.33% of participants had a negative opinion of the Varjo XR-3, especially with regard to wearing comfort.

6 Discussion

We discuss our research findings based on our three hypotheses. Afterward, we provide design implications for future cognitively demanding AR/VR escape rooms.

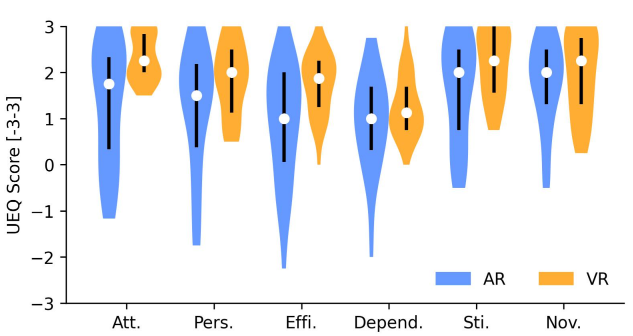


Figure 8: Violin plots of the UEQ sub-scales [30] between the AR and VR conditions.

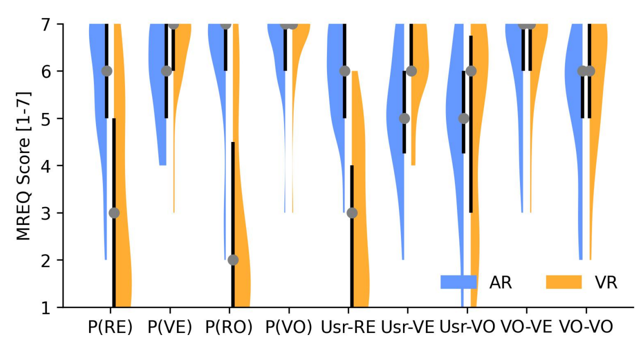


Figure 10: Violin plots comparing the MREQ items between the AR and VR conditions.

6.1 H1: Differences in Perceived Coherence

Based on the split-attention-effect [55] and the concept of processing fluency [45] we derived our **H1** as current AR applications have a natural break in consistency in the visual representation of physical and virtual objects in contrast to VR applications, which can challenge users while recognizing the connection between the differently visualized objects, see Section 2. With our **H1**, we investigated whether perceived coherence is higher in VR than in AR. There were no statistically significant differences in *Experienced Realism* between AR and VR, see Section 5.6. The MREQ item VO-VE (see Section 5.7) revealed high coherence in VR since virtual objects were strongly assigned as belonging to the virtual environment. In contrast, AR participants were more convinced that the virtual objects belonged to the virtual environment than to the real environment. This indicates a low perceived coherence between the real and virtual environments in AR. The qualitative data further promotes low coherence in AR. The collected data suggest that perceived coherence was higher in VR than in AR, which aligns with related work [58]. However, **H1** can not be confirmed with statistically significant differences.

6.2 H2: Differences in Cognitive Load

Related work provides evidence that AR induced higher CL compared to non-AR scenarios [11, 27, 31] and in direct comparison to VR [58, 60]. Based on these previous findings, our **H2** states that "the CL in AR is higher than in VR." This is supported by the fact that task performance was significantly greater and the task completion time significantly lower in VR (see Section 5.1 and 5.2). While VR further evoked a lower blink rate than AR, the pupil diameter slope was greater in VR than in AR (see Section 5.3). Moreover, the significantly higher feelings of *Involvement* and *Spatial Presence* in VR (see Section 5.6) indicate lower CL in VR [39]. In contrast to VR, AR participants tended not to feel as part of the virtual environment (see Section 5.7). Given these five metrics (task performance, task completion time, blink rate, IPQ, and MREQ), VR evokes a lower level of CL than AR. Thus, we can verify our **H2**, which resembles our findings from our literature research.

6.3 H3: Difference in Game Enjoyment

Regarding our literature review, both AR and VR can have a positive effect on enjoyment, while there has not been a direct comparison between the two technologies in the context of enjoyment. A high sense of presence positively influences CL [39] as well as enjoyment [19]. And given our **H2**, telepresence in VR is stronger than the local presence in AR. Additionally, AR applications require more cognitive resources than VR applications [54], which might lead to higher frustration and, therefore, less enjoyment [13]. Our **H3** states that "in VR, the participants perceive more enjoyment than in AR." Participants completed the UEQ to investigate enjoyment. In our study, VR achieved better results than AR on all scales (see Section 5.5). Therefore, VR tends to provide a better user experience than AR. Nevertheless, hedonistic qualities and the NASA-TLX subscale *Frustration*, which are particularly important for enjoyment, did not show statistically significant differences between VR and AR. However, since we frequently argued that presence is closely linked to enjoyment [19], the first of which was significantly higher in VR (see Section 5.6), we assume that VR produced higher levels of enjoyment than AR. Hence, our results confirm **H3**.

6.4 Design Implications

First, we derive general design implications for cognitively demanding XR applications and then address the specific requirements of AR and VR separately. To keep the CL low, we recommend making interaction with virtual objects in XR environments as natural as possible, for example, through hand tracking. This allows users to move within an application as they would in a purely physical world. Thus, they do not have to learn new interaction concepts, which would require additional CL.

6.4.1 AR. We identified potential opportunities for enhancing the design of physical environments that utilize AR applications. Varjo provided the markers used in our AR setup, which resemble QR codes. Due to their abstract look, participants quickly realized that there was something to discover and, thus, assigned a high level of importance to the virtual objects that appeared after scanning. This sometimes led them down the wrong path and made them neglect other relevant objects. Therefore, we suggest using image targets instead of abstract markers, since they can be individually designed and, thus, better integrated into the environmental context.

The physical environment should only contain objects relevant to the AR application, thereby minimizing distractions for participants. We also observed that some participants were inhibited when using physical objects in an unfamiliar room. This could be because our user study took place in a university laboratory or because the study organizer was in the same room during the escape game. Therefore, we further suggest that the study organizers are either not in the same room or act as part of the application's story. In both cases, this would promote the immersion of the participants and, thus, prevent an increase in their CL.

6.4.2 VR. For cognitively demanding VR applications, we formulate design proposals primarily concerning interactions. In our escape game, we have provided an offset grab that works via a ray, which indicates whether a virtual object is capable of interaction. The participants sometimes had difficulty understanding this unnatural visualization. Some accidentally pressed a controller button while the ray was active and, thus, unintentionally interacted with objects. To prevent this, interaction methods should be understandable and ideally based on concepts that are familiar to the real world to avoid frustration. For VR applications that resemble the real world, we recommend designing interactions in such a way that they require the same effort as in the physical environment. Another drawback of ray visualization is that it makes it possible to recognize more quickly in VR whether a virtual object can interact with it, in contrast to the physical counterparts in AR.

6.4.3 The Digitally Enhanced Escape Room. Our escape room is based on a scenario that is not possible in the physical world (starry sky in a room, tangible stars). This is possible in both VR and AR. Nevertheless, VR offers more possibilities than AR when it comes to depicting such physically impossible scenarios, as the real world is completely replaced by the former. Our collected data shows that a higher coherence, as well as a tendency towards a lower CL and a higher level of enjoyment, were perceived in VR compared to

AR. This leads us to conclude that unnatural scenarios are more suitable for VR applications than for AR. We assume that virtual objects that cannot exist in the real world represent a greater break in the consistency of the displayed environment than those known from the physical world. This further impairs processing fluency, which ultimately harms perceived coherence and CL. Our escape room scenario is, therefore, more suitable for VR, while AR should be utilized for use cases in which concepts from the real world are being enhanced by virtual elements. Therefore, the purpose of AR is not to disguise the physical room with virtual elements but to combine the virtual and the physical environment into one harmonized whole. The aim of VR, on the other hand, is to realize scenarios in which concepts of the physical world are not relevant and cannot be implemented because of high costs and/or impossible conditions (e.g., an escape room on Mars).

Single- vs. Multi-User. While we focused on the impact of AR and VR on CL in a single-player setting, it is crucial to consider how these findings translate to multiplayer environments. As AR and VR technologies continue to advance, multiplayer experiences are becoming increasingly popular, raising questions about how social interactions and collaborative experiences might influence CL. Future studies should aim to replicate our experiment in a multiplayer context to explore whether the differences in CL between AR and VR persist. Moreover, researchers could investigate how different multiplayer configurations, such as competitive versus cooperative play, affect CL in both environments. By extending our findings to multiplayer scenarios, we can gain an understanding of the interplay between technology, social interaction, and cognition.

Co-located vs. Remote. The implications of co-located vs. remote scenarios in AR and VR escape rooms warrant consideration. In VR escape rooms, a remote scenario is more suitable as it eliminates the need for physical props and allows all players to experience the same virtual environment. This is particularly beneficial as physically recreating the same environment for multiple players would be impractical. Conversely, co-located scenarios in VR are less ideal due to the limitations of being unable to see fellow players in real life and the increased risk of accidental collisions due to a diminished sense of presence in the physical environment. In contrast, AR escape rooms are more suitable for co-located scenarios as they enable players to share the same physical space and interact in real time. However, remote AR scenarios present challenges, requiring each player to have an identical physical environment, which is often unfeasible, limiting interactions to virtual elements only.

6.5 Limitations

With our research, we aimed to directly compare CL and enjoyment in AR and VR using the same hardware and room setup. Due to restricted hardware resources and laboratory space, our contribution has some limitations. Here, as with most HMD studies, participants criticized the low comfort of wearing the HMD as well as the unstable tracking. Moreover, as see-through technology is not perfect yet, natural light may not be presented flawlessly in the headset.

Although our user study took place in the same room, the play area between the two conditions differed, as the physical props would pose a hazard to the VR participants. On the other hand, only the spaces between the puzzles are impacted by this; the puzzles themselves were not affected by the play area. Moreover, due to the nature of AR, we utilized real-world markers to trigger actions, whereas in VR, user interactions trigger actions. This difference could have affected the outcome.

The eye-tracking functionality of the HMD did not work reliably for two participants in the AR condition. As a result, we were unable to collect eye-tracking data from two test subjects and were unable to conclude their CL using this metric. Our eye-tracking data, therefore, only includes 58 participants.

In our study, we ensured that all conditions were comparable, particularly regarding head presentation quality, visibility, and overall usability. To achieve this, we used the same device for both AR and VR conditions. However, this approach limited us to using video see-through AR. Recent findings by Javaheri et al. [20] indicate differences between video and optical see-through AR, and the impact of such hardware differences on our results remains unclear. Similarly, Ballestin et al. [3] reported tradeoffs between the two approaches. Therefore, future work should investigate how our findings generalize to true optical see-through systems.

7 Conclusion

This paper investigates the impact of AR and VR on cognitive load and immersive experiences in escape rooms. Our user study with 60 participants revealed that AR induces a significantly higher cognitive load than VR despite using the same hardware and environment. Moreover, presence and usability were higher in VR. The results suggest that AR environment designers need to consider the increased cognitive load to enhance user experience and performance. We highlighted future implications for cognitively demanding AR/VR environments, such as escape rooms, including making interactions natural and promoting processing fluency as well as immersion to prevent increased cognitive load. Our results are the first to be based on a direct comparison regarding enjoyment in AR/VR, operated under identical technical conditions for both technologies.

However, further research is required to ensure the generalizability of our results. Ordinarily, escape rooms are joined in a team of about four to eight [10, 34]. To fulfill the typical application context of escape rooms, our single-player escape room concept could be expanded in the future so that it is suitable for multiple users. A user study can then be conducted again with several HMDs to investigate how the CL in a multi-user context differs from our results. Furthermore, escape room experiences usually last longer than ten minutes. As a rule, the team has a time limit of usually one hour [34]. In addition, several test subjects wished for more time to play the escape room. With a longer time limit, future work could include additional puzzles as well as more elements, and therefore investigate CL, perceived coherence, and enjoyment in more extensive escape rooms.

Open Science

We encourage readers to review, reproduce, and extend our findings. To achieve this goal, we open-sourced Unity code at GitHub https://github.com/mimuc/AR-VR-escape-room and analysis script available at Zenodo http://doi.org/10.5281/zenodo.17398134.

Author Contributions

Vivien Landgrebe: Conceptualization, Formal Analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing; **Elizabeth Bouquet**: Conceptualization, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; **Simon von der Au**: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing; **Sven Mayer**: Data curation, Formal Analysis, Visualization, Supervision, Writing – original draft, Writing – review & editing

Acknowledgments

We would like to thank the Innovative Hochschule, funded by the German Federal Ministry of Education and Research (BMBF), for supporting this research. This study has been conducted as part of the CreatiF Center research mission. This work has been partly supported by the Research Center Trustworthy Data Science and Security (https://rc-trust.ai), one of the Research Alliance centers within the UA Ruhr (https://uaruhr.de).

References

[1] Ronald T. Azuma. 1997. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. doi:10.1162/PRES.1997.6.4.355

[2] Jon Back, Svante Back, Emma Bexell, Stefan Stanisic, and Daniel Rosqvist. 2019. the Quest: An Escape Room Inspired Interactive Museum Exhibition. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Barcelona, Spain) *(CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 81–86. doi:10.1145/3341215.3356987

[3] Giorgio Ballestin, Manuela Chessa, and Fabio Solari. 2021. A Registration Framework for the Comparison of Video and Optical See-Through Devices in Interactive Augmented Reality. *IEEE Access* 9 (2021), 64828–64843. doi:10.1109/ACCESS.2021.3075780

[4] Austin Cory Bart. 2022. Let's Learn Algorithms with AlgoTutorBot! An Entire Course as an Educational Escape Room. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 2* (Providence, RI, USA) *(SIGCSE 2022)*. Association for Computing Machinery, New York, NY, USA, 1178. doi:10.1145/3478432.3499262

[5] Katja Becker, Helvi Itenge, Selma Auala, Joseph-David Isaacs, and Heike Winschiers-Theophilus. 2024. Co-design of an Extended Reality Escape Room with Primary School Students. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference* (Delft, Netherlands) *(IDC '24)*. Association for Computing Machinery, New York, NY, USA, 744–748. doi:10.1145/3628516.3659387

[6] Sarah Brown, Ilda Ladeira, Cara Winterbottom, and Edwin Blake. 2003. The Effects of Mediation in a Storytelling Virtual Environment. In *Virtual Storytelling. Using Virtual RealityTechnologies for Storytelling*, Olivier Balet, Gérard Subsol, and Patrice Torguet (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 102–111. doi:10.1007/978-3-540-40014-1_13

[7] Julie Carmigniani and Borko Furht. 2011. Augmented Reality: An Overview. In *Handbook of Augmented Reality*, Borko Furht (Ed.). Springer New York, New York, NY, Chapter 1, 3–46. doi:10.1007/978-1-4614-0064-6_1

[8] Paul Chandler and John Sweller. 1992. The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology* 62, 2 (6 1992), 233–246. doi:10.1111/j.2044-8279.1992.tb01017.x

[9] Chin Jung Chao, Sheng Yu Wu, Yi Jan Yau, Weng Yan Feng, and Feng Yi Tseng. 2017. Effects of three-dimensional virtual reality and traditional training methods on mental workload and training performance. *Human Factors and Ergonomics In Manufacturing* 27, 4 (7 2017), 187–196. doi:10.1002/hfm.20702

[10] Samantha Jane Clarke, Daryl J. Peel, Sylvester Arnab, Luca Morini, Helen Keegan, and Oliver Wood. 2017. EscapED: A Framework for Creating Educational Escape Rooms and Interactive Games to For Higher/Further Education. *International Journal of Serious Games* 4, 3 (9 2017), 14. doi:10.17083/ijsg.v4i3.180

[11] Mathilde Drouot, Nathalie Le Bigot, Emmanuel Bricard, Jean Louis de Bougrenet, and Vincent Nourrit. 2022. Augmented reality on industrial assembly line: Impact on effectiveness and mental workload. *Applied Ergonomics* 103 (9 2022), 103793. doi:10.1016/j.apergo.2022.103793

[12] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. doi:10.1177/154193120605000909

[13] Heng Cui and Ernst Biersack. 2013. Troubleshooting slow webpage downloads. In *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, New York, NY, USA, 405–410. doi:10.1109/INFCOMW.2013.6562884

[14] Linda Hirsch, Florian Müller, Francesco Chiossi, Theodor Benga, and Andreas Martin Butz. 2023. My Heart Will Go On: Implicitly Increasing Social Connectedness by Visualizing Asynchronous Players' Heartbeats in VR Games. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 411 (Oct. 2023), 26 pages. doi:10.1145/3611057

[15] Fabian Holly, Tanja Zigart, Martina Maurer, Josef Wolfartsberger, Magdalena Brunnhofer, Sabrina Romina Sorko, Thomas Moser, and Alexander Schlager. 2022. Gaining Impact with Mixed Reality in Industry – A Sustainable Approach. In *Proceedings of the 2022 8th International Conference on Computer Technology Applications* (Vienna, Austria) *(ICCTA '22)*. Association for Computing Machinery, New York, NY, USA, 128–134. doi:10.1145/3543712.3543729

[16] Beat Horat and Mareen Przybylla. 2021. CS Learning in Educational Escape Room. In *Proceedings of the 16th Workshop in Primary and Secondary Computing Education* (Virtual Event, Germany) *(WiPSCE '21)*. Association for Computing Machinery, New York, NY, USA, Article 9, 2 pages. doi:10.1145/3481312.3481327

[17] Shafaq Irshad and Dayang Rohaya Bt Awang Rambli. 2015. User experience of mobile augmented reality: A review of studies. In *Proceedings - 2014 3rd International Conference on User Science and Engineering: Experience. Engineer. Engage, i-USEr 2014*. IEEE, New York, NY, USA, 125–130. doi:10.1109/IUSER.2014.7002689

[18] Helvi Itenge, Carolina Islas Sedano, and Heike Winschiers-Theophilus. 2021. Virtually Escaping Lock Down - co-designing a mixed reality escape room narrative with Namibian learners-. In *Proceedings of the 3rd African Human-Computer Interaction Conference: Inclusiveness and Empowerment* (Maputo, Mozambique) *(AfriCHI '21)*. Association for Computing Machinery, New York, NY, USA, 103–112. doi:10.1145/3448696.3448700

[19] Yeonju Jang and Eunil Park. 2019. An adoption model for virtual reality games: The roles of presence and enjoyment. *Telematics and Informatics* 42 (9 2019), 101239. doi:10.1016/j.tele.2019.101239

[20] Hamraz Javaheri, Vitor Fortes Rey, Paul Lukowicz, Gregor A Stavrou, Jakob Karolus, and Omid Ghamarnejad. 2025. Assessing the Feasibility of Using Apple Vision Pro While Performing Medical Precision Tasks: Controlled User Study. *JMIR XR Spatial Comput* 2 (15 Aug 2025), e73574. doi:10.2196/73574

[21] Nor Farzana Syaza Jeffri and Dayang Rohaya Awang Rambli. 2021. A review of augmented reality systems and their effects on mental workload and task performance. *Heliyon* 7, 3 (3 2021), e06277. doi:10.1016/j.heliyon.2021.e06277

[22] Timothy Jung, M. Claudia tom Dieck, Hyunae Lee, and Namho Chung. 2016. Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum. In *Information and Communication Technologies in Tourism 2016*. Springer International Publishing, Cham, Switzerland, 621–635. doi:10.1007/978-3-319-28231-2_45

[23] Juho Kahila, Tuomo Parkki, Anssi Gröhn, Atte Karvinen, Elmeri Telimaa, Pekka Riikonen, Riku Tiitta, Pasi Haantio, Anssi Keinänen, Tero Kerkkänen, Ilkka Jormanainen, Silja Penttinen, and Matti Tedre. 2020. Escape Room Game for CT Learning Activities in the Primary School. In *Proceedings of the 20th Koli Calling International Conference on Computing Education Research* (Koli, Finland) *(Koli Calling '20)*. Association for Computing Machinery, New York, NY, USA, Article 9, 5 pages. doi:10.1145/3428029.3428063

[24] Tomo Kihara, Roy Bendor, and Derek Lomas. 2019. Designing an Escape Room in the City for Public Engagement with AI-enhanced Surveillance. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3290607.3313003

[25] Erica Kleinman and Casper Harteveld. 2024. The Untapped Potential of Escape Rooms as Gamified Research Environments. In *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play* (Tampere, Finland) *(CHI PLAY Companion '24)*. Association for Computing Machinery, New York, NY, USA, 276–278. doi:10.1145/3665463.3678865

[26] Theodore Knoll, Amna Liaqat, and Andrés Monroy-Hernández. 2023. ARctic Escape: Promoting Social Connection, Teamwork, and Collaboration Using a Co-Located Augmented Reality Escape Room. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 38, 6 pages. doi:10.1145/3544549.3585841

[27] René Kockord and Oliver Bodensiek. 2021. Cognitive Load During First Contact With Mixed Reality Learning Environments. In *Proceedings of Mensch Und Computer 2021* (Ingolstadt, Germany) *(MuC '21)*. Association for Computing Machinery, New York, NY, USA, 260–264. doi:10.1145/3473856.3474003

[28] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Wozniak. 2023. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *Comput. Surveys* 55, 13s (7 2023), 1–39. doi:10.1145/3582272

[29] Andrey Krekhov, Katharina Emmerich, Ronja Rotthaler, and Jens Krueger. 2021. Puzzles Unpuzzled: Towards a Unified Taxonomy for Analog and Digital Escape

Room Games. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 269 (Oct. 2021), 24 pages. doi:10.1145/3474696

[30] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 63–76. doi:10.1007/978-3-540-89350-9_6

[31] Frieder Loch, Fabian Quint, and Iuliia Brishtel. 2016. Comparing video and augmented reality assistance in manual assembly. In *Proceedings - 12th International Conference on Intelligent Environments (IE 2016)*. IEEE, New York, NY, USA, 147–150. doi:10.1109/IE.2016.31

[32] Tiffany Luong, Ferran Argelaguet, Nicolas Martin, and Anatole Lecuyer. 2020. Introducing Mental Workload Assessment for the Design of Virtual Reality Training Scenarios. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York, NY, USA, 662–671. doi:10.1109/VR46266.2020.00089

[33] Peter Madsen, Henning Pohl, and Timothy Merritt. 2022. Fear Inducing Play in an AR Escape Room with Human and Robotic NPCs. In *Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play* (Bremen, Germany) *(CHI PLAY '22)*. Association for Computing Machinery, New York, NY, USA, 38–43. doi:10.1145/3505270.3558317

[34] Sofia Niovi Maragkoudaki and Christos Kalloniatis. 2023. Virtual Reality as a mean for increasing privacy awareness: The escape room example. In *Proceedings of the 26th Pan-Hellenic Conference on Informatics* (Athens, Greece) *(PCI '22)*. Association for Computing Machinery, New York, NY, USA, 261–266. doi:10.1145/3575879.3576002

[35] Corey Thomas Martin, James Baumeister, Andrew Cunningham, and G Stewart Von Itzstein. 2021. The Spheres of Player Motivation: Towards a New Typology for Player Motivation in Digital Games. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Austria) *(CHI PLAY '21)*. Association for Computing Machinery, New York, NY, USA, 140–145. doi:10.1145/3450337.3483491

[36] Tilman Michaeli and Ralf Romeike. 2020. Investigating Students' Preexisting Debugging Traits: A Real World Escape Room Study. In *Proceedings of the 20th Koli Calling International Conference on Computing Education Research* (Koli, Finland) *(Koli Calling '20)*. Association for Computing Machinery, New York, NY, USA, Article 15, 10 pages. doi:10.1145/3428029.3428044

[37] Brendan Mouatt, Ashleigh E. Smith, Maddison L. Mellow, Gaynor Parfitt, Ross T. Smith, and Tasha R. Stanton. 2020. The Use of Virtual Reality to Influence Motivation, Affect, Enjoyment, and Engagement During Exercise: A Scoping Review. doi:10.3389/frvir.2020.564664

[38] Livia J. Müller, Elisa D. Mekler, and Klaus Opwis. 2016. Hedonic Enjoyment and Personal Expressiveness in Positive User Experiences. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 3166–3172. doi:10.1145/2851581.2892494

[39] Federica Nenna, Davide Zanardi, and Luciano Gamberini. 2023. Effects of Presence on Human Performance and Workload in Simulated VR-based Telerobotics. In *Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Greece) *(PETRA '23)*. Association for Computing Machinery, New York, NY, USA, 47–52. doi:10.1145/3594806.3594856

[40] Iina Nikkarikoski, Johanna Pehkonen, Iiris Kivelä, Paula Alavesa, and Mikko Korkiakoski. 2024. Galaxy Escape: A Game as a Tabletop Holographic Digital Twin Interface. In *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play* (Tampere, Finland) *(CHI PLAY Companion '24)*. Association for Computing Machinery, New York, NY, USA, 215–220. doi:10.1145/3665463.3678810

[41] Fang Chuan Ou Yang, Hui Min Lai, and Yen Wen Wang. 2023. Effect of augmented reality-based virtual educational robotics on programming students' enjoyment of learning, computational thinking skills, and academic achievement. *Computers and Education* 195 (4 2023), 104721. doi:10.1016/j.compedu.2022.104721

[42] Irina Paraschivoiu, Josef Buchner, Robert Praxmarer, and Thomas Layer-Wagner. 2021. Escape the Fake: Development and Evaluation of an Augmented Reality Escape Room Game for Fighting Fake News. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Austria) *(CHI PLAY '21)*. Association for Computing Machinery, New York, NY, USA, 320–325. doi:10.1145/3450337.3483454

[43] Wolfgang Pfeffer, Tobias Fuchs, Luisa Greifenstein, and Ute Heuer. 2024. Codes and Ciphers: A Digital Escape Room for Secondary Schools (Practical Report). In *Proceedings of the 19th WiPSCE Conference on Primary and Secondary Computing Education Research* (Munich, Germany) *(WiPSCE '24)*. Association for Computing Machinery, New York, NY, USA, Article 15, 6 pages. doi:10.1145/3677619.3678112

[44] Marc Pfeifer, Benjamin Völker, Sebastian Böttcher, Sven Köhler, and Philipp M. Scholl. 2021. Teaching Embedded Systems by Constructing an Escape Room. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (Virtual Event, USA) *(SIGCSE '21)*. Association for Computing Machinery, New York, NY, USA, 1103–1109. doi:10.1145/3408877.3432485

[45] Jan Preßler, Lukas Schmid, and Jörn Hurtienne. 2023. Statistically Controlling for Processing Fluency Reduces the Aesthetic-Usability Effect. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 261, 7 pages. doi:10.1145/3544549.3585739

[46] Yan Qin. 2021. Attractiveness of game elements, presence, and enjoyment of mobile augmented reality games: The case of Pokémon Go. *Telematics and Informatics* 62 (9 2021), 101620. doi:10.1016/j.tele.2021.101620

[47] Philipp A. Rauschnabel, Reto Felix, Chris Hinsch, Hamza Shahab, and Florian Alt. 2022. What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior* 133 (8 2022), 107289. doi:10.1016/j.chb.2022.107289

[48] Holger Regenbrecht, Cristina Botella, Rosa Baños, and Thomas Schubert. 2017. *Mixed Reality Experience Questionnaire (MREQ)-Reference The Information Science Discussion Paper Series.* Technical Report. Department of Information Science, University of Otago, Dunedin. https://hdl.handle.net/10523/7151

[49] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2014. Applying the user experience questionnaire (UEQ) in different evaluation scenarios. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 8517 LNCS. Springer Verlag, Berlin, Germany, 383–392. doi:10.1007/978-3-319-07668-3_37

[50] Norbert Schwarz. 2004. Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology* 14, 4 (1 2004), 332–348. doi:10.1207/s15327663jcp1404_2

[51] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. 2019. Using Presence Questionnaires in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3290605.3300590

[52] Hanieh Shakeri, Samarth Singhal, Rui Pan, Carman Neustaedter, and Anthony Tang. 2017. Escaping Together: The Design and Evaluation of a Distributed Real-Life Escape Room. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 115–128. doi:10.1145/3116595.3116601

[53] Andrei Sherstyuk, Dale Vincent, and Anton Treskunov. 2009. Towards Virtual Reality Games. In *Proceedings of the 8th International Conference on Virtual Reality Continuum and Its Applications in Industry*. Association for Computing Machinery, New York, NY, USA, 315–316. doi:10.1145/1670252.1670322

[54] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. *Cognitive Load Theory*. Springer New York, New York, NY. doi:10.1007/978-1-4419-8126-4

[55] John Sweller, Paul Ayres, and Slava Kalyuga. 2011. The Split-Attention Effect. In *Cognitive Load Theory. Explorations in the Learning Sciences, Instructional Systems and Performance Technologies.* Vol. 1. Springer New York, New York, NY, Chapter 9, 111–128. doi:10.1007/978-1-4419-8126-4_9

[56] Lydia Velázquez-García, Antonio Cedillo-Hernández, and Eduardo Antonio Cendejas-Castro. 2024. Gamification in higher education: learning computer science with an escape room. In *Proceedings of the 15th International Conference on Education Technology and Computers* (Barcelona, Spain) *(ICETC '23)*. Association for Computing Machinery, New York, NY, USA, 72–77. doi:10.1145/3629296.3629308

[57] Harald Warmelink, Igor Mayer, Jessika Weber, Bram Heijligers, Mata Haggis, Erwin Peters, and Max Louwerse. 2017. AMELIO: Evaluating the Team-building Potential of a Mixed Reality Escape Room Game. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 111–123. doi:10.1145/3130859.3131436

[58] Nannan Xi, Juan Chen, Filipe Gama, Marc Riar, and Juho Hamari. 2023. The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. *Information Systems Frontiers* 25, 2 (4 2023), 659–680. doi:10.1007/s10796-022-10244-x

[59] Jongsik Yu, Seongseop Kim, Tadesse Bekele Hailu, Junghyun Park, and Heesup Han. 2024. The effects of virtual reality (VR) and augmented reality (AR) on senior tourists' experiential quality, perceived advantages, perceived enjoyment, and reuse intention. *Current Issues in Tourism* 27, 3 (2024), 464–478. doi:10.1080/13683500.2023.2165483

[60] Xiaojun Zhao, Changxiu Shi, Xuqun You, and Chenming Zong. 2017. Analysis of mental workload in online shopping: Are augmented and virtual reality consistent? *Frontiers in Psychology* 8, JAN (1 2017), 8. doi:10.3389/fpsyg.2017.00071

[61] Ruotong Zou, Shuyu Yin, Tianqi Song, Peinuan Qin, and Yi-Chieh Lee. 2024. Mitigating Ageism through Virtual Reality: Intergenerational Collaborative Escape Room Design. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 241, 7 pages. doi:10.1145/3613905.3650914

A Appendix

Demographic questions:

Q_{Demo}1 How old are you? [Number]

Q_{Demo}2 Which gender do you identify with the most?

Q_{Demo}3 What is your current occupation?

Q_{Demo}4 How much previous experience do you have with augmented reality? [7-point Likert item, 1 = No previous experience, 7 = Expert]

Q_{Demo}5 How frequently do you use augmented reality? [7-point Likert item, Daily, Once a week, Multiple times a week, Once a month, Once every few months, Once a year or less, Never]

Q_{Demo}6 Do you own an augmented reality device? [yes/no]

Q_{Demo}7 If yes, which device? [Free text]

Q_{Demo}8 How much previous experience do you have with virtual reality?[7-point Likert item, 1 = No previous experience, 7 = Expert]

Q_{Demo}9 How frequently do you use virtual reality? [7-point Likert item, Daily, Once a week, Multiple times a week, Once a month, Once every few months, Once a year or less, Never]

Q_{Demo}10 Do you own a virtual reality device? [yes/no]

Q_{Demo}11 If yes, which device? [Free text]