SwitchAR: Enabling Perceptual Manipulations in Augmented Reality

SwitchAR: Wahrnehmungsmanipulationen in der erweiterten Realität
Master thesis in the department of Computer Science by Jonas Julian Wombacher
Date of submission: September 27, 2024

1. Review: Prof. Dr. Jan Gugenheimer
2. Review: Dr. rer. pol. Veronika Krauß
Darmstadt

TECHNISCHE UNIVERSITÄT DARMSTADT

Computer Science Department
HCI Lab

Erklärung zur Abschlussarbeit gemäß § 22 Abs. 7 APB TU Darmstadt

Hiermit erkläre ich, Jonas Julian Wombacher, dass ich die vorliegende Arbeit gemäß § 22 Abs. 7 APB TU Darmstadt selbstständig, ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt habe. Ich habe mit Ausnahme der zitierten Literatur und anderer in der Arbeit genannter Quellen keine fremden Hilfsmittel benutzt. Die von mir bei der Anfertigung dieser wissenschaftlichen Arbeit wörtlich oder inhaltlich benutzte Literatur und alle anderen Quellen habe ich im Text deutlich gekennzeichnet und gesondert aufgeführt. Dies gilt auch für Quellen oder Hilfsmittel aus dem Internet.

Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Mir ist bekannt, dass im Falle eines Plagiats (§38 Abs.2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird. Abschlussarbeiten dürfen nur einmal wiederholt werden.

Bei einer Thesis des Fachbereichs Architektur entspricht die eingereichte elektronische Fassung dem vorgestellten Modell und den vorgelegten Plänen.

English translation for information purposes only:

Thesis Statement pursuant to § 22 paragraph 7 of APB TU Darmstadt

I herewith formally declare that I, Jonas Julian Wombacher, have written the submitted thesis independently pursuant to § 22 paragraph 7 of APB TU Darmstadt without any outside support and using only the quoted literature and other sources. I did not use any outside support except for the quoted literature and other sources mentioned in the paper. I have clearly marked and separately listed in the text the literature used literally or in terms of content and all other sources I used for the preparation of this academic work. This also applies to sources or aids from the Internet.

This thesis has not been handed in or published before in the same or similar form.

I am aware, that in case of an attempt at deception based on plagiarism (§38 Abs. 2 APB), the thesis would be graded with 5,0 and counted as one failed examination attempt. The thesis may only be repeated once.

For a thesis of the Department of Architecture, the submitted electronic version corresponds to the presented model and the submitted architectural plans.

Datum / Date: 27.09.2024

Unterschrift/Signature: J. Wombacher

Zusammenfassung

Wahrnehmungsmanipulationen (PMs) wie beispielsweise Redirected Walking (RDW) werden in der Virtuellen Realität (VR) häufig eingesetzt, um technologische Einschränkungen zu überwinden. Diese PMs verändern die visuelle Wahrnehmung der Nutzer, etwa durch Rotationsverstärkungen, was in der Erweiterten Realität (AR) weiterhin eine Herausforderung darstellt. Wir stellen SwitchAR vor, eine PM, die für Video Pass-Through AR entwickelt wurde und die Phänomene der Veränderungsblindheit und Unaufmerksamkeitsblindheit nutzt, um unbemerkt zwischen dem Live-Kamerabild der realen Welt und einer 3D-Rekonstruktion zu wechseln. Dadurch wird die Anwendung von VR-Umleitungstechniken ermöglicht, während die Nutzer weiterhin das Gefühl haben, sich in AR zu befinden. Unsere Pipeline besteht aus vier Schritten: (1) Rekonstruktion, (2) Wechsel von AR zu VR, (3) PM und (4) Wechsel von VR zu AR, wobei jede Phase zusammen mit ihren jeweiligen Herausforderungen und unseren Lösungsansätzen erläutert wird. In einer RDW-Nutzerstudie mit 20 Teilnehmern bemerkte niemand den Wechsel, und nur ein Teilnehmer bemerkte die Umleitung. Selbst nachdem die Teilnehmer darüber informiert wurden, dass eine Manipulation angewendet wurde, waren sie in einem anschließenden Versuch nicht in der Lage, den Wechsel zu erkennen. SwitchAR dient als grundlegender Ansatz zur Ermöglichung von PMs in AR.

Abstract

Perceptual manipulations (PMs) such as redirected walking (RDW) are commonly employed in Virtual Reality (VR) to address technological constraints. These PMs alter users' visual perceptions, for example through rotational gains, a task that remains challenging in Augmented Reality (AR). We introduce SwitchAR, a PM designed for video pass-through AR, which takes advantage of change blindness and inattentional blindness to unnoticeably transition between the live camera feed of the real world and a 3D reconstruction. This allows for the application of VR redirection techniques while users continue to perceive being in AR. Our pipeline consists of four key stages: (1) Reconstruction, (2) Switch from AR to VR, (3) PM, and (4) Switch from VR to AR, each of which is discussed along with its respective challenges and our solutions for them. In an RDW user study with 20 participants, none detected the switch, and only one noticed the redirection. Furthermore, even after being informed that a manipulation was being applied, participants were still unable to detect the switch in a subsequent trial. SwitchAR serves as a foundational approach for enabling PMs in AR.

Contents

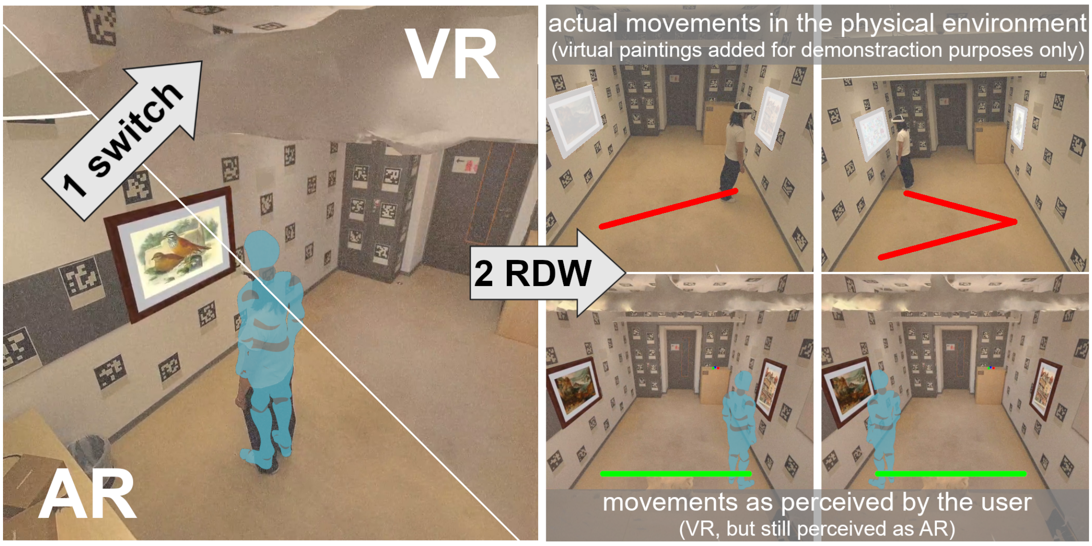

Figure 1: Overview of the key stages in SwitchAR. 1: Using an aligned virtual reconstruction of the environment, the switch from Augmented Reality (AR) to Virtual Reality (VR) is performed without the user's awareness. 2: While in VR, established perceptual manipulations such as redirected walking (RDW) can be utilized. In this case, RDW is used to create a misalignment between the user's perceived virtual position and their actual physical position within the real environment.

1 Introduction

Virtual Reality (VR) applications frequently utilize a range of perceptual manipulation (PM) techniques. A specific category, Virtual-Physical Perceptual Manipulation (VPPM)[34], allows developers to subtly alter users' physical movements. One common use case for VPPMs is redirected walking (RDW)[25], where users are unconsciously guided to alter their walking paths, making it possible to explore virtual environments larger than the available physical space. Another example is the manipulation of hand movements[1, 18, 21], which increases the availability of passive haptic feedback without requiring additional physical props. This is possible because VR head-mounted displays (HMDs) fully obscure the real world, granting complete control over the user's visual field of view (FOV) and allowing for the subtle manipulation of the virtual environment to exploit perceptual thresholds.
In contrast, current Augmented Reality (AR) HMDs primarily add virtual elements to the real world without fully occluding it, limiting the degree of control over the environment. This visibility of the real world restricts the ability to implement traditional PMs in AR. Recent advances in camera technology and expanded FOV have led to the rise in popularity of pass-through AR HMDs, such as the Meta Quest 3 and Apple Vision Pro, making pass-through AR a widespread form of AR. These devices are capable not only of running AR, but also of fully occluding the real world to support VR experiences.
We introduce SwitchAR, a PM technique for pass-through AR that leverages change blindness and inattentional blindness to discreetly switch between the live camera feed of the real world and a virtual 3D reconstruction. By utilizing tricks like distraction tasks, persistent visual noise, and minimizing visible changes, the switch remains undetected despite minor reconstruction inaccuracies. This enables users to continue believing they are interacting with the physical environment in AR, while developers gain full control over the reconstructed environment, allowing the application of VR redirection techniques.
The SwitchAR pipeline consists of four stages: (1) generating a 3D reconstruction of the environment, (2) covertly switching from the real-world pass-through feed to the

virtual reconstruction, (3) applying a PM, such as RDW in our case, and (4) optionally switching back to the pass-through feed. Throughout this process, it is critical that users, while wearing the HMD, consistently believe they are experiencing AR.

To validate our approach, we implemented SwitchAR together with an RDW technique. Our RDW method applied rotational gains when users turned at waypoints, similar to the technique described by Razzaque[25]. We evaluated this implementation in a user study with 20 participants, conducting three rounds per participant. During the study, participants engaged in a distraction task, walking between virtual paintings while believing they were participating in a memory experiment. The second round ended with a reveal of the redirection.

The study aimed to measure: (1) whether unsuspecting participants noticed the switch, (2) if they noticed the RDW, and (3) whether the PMs were still effective when participants were aware of the manipulations. The noticeability of the switch was measured using a set of increasingly direct questions, similar to the approach of Simons et al. in their inattentional blindness study[28]. Participants were asked, for example, if they noticed any changes in the environment and if they had been using AR for the entire study.

For the open-ended questions, responses were coded, and any mention of visual changes in the environment was considered as detecting the switch (the raw responses are provided in appendix A and B). We found that none of the participants detected the switch from AR to VR across any of the three rounds, and only one participant consistently noticed the redirection. Even after the manipulation was revealed in round two, 17 participants were unable to correctly explain how RDW was achieved within AR. In the third round, when participants were already aware of the RDW manipulation, none detected the switch, and only one additional participant noticed the redirection. These results indicate that SwitchAR can serve as a robust foundation for applying VR PMs within video pass-through AR.

Our main contributions are: (1) the concept and implementation of SwitchAR, (2) the detailed description of the pipeline and its challenges, along with proposed solutions, and (3) a demonstration of SwitchAR's feasibility, even after participants were made aware of the manipulation, through a user study (n=20).

This thesis is structured in nine chapters. Chapter two reviews related work. Chapter three introduces the concept of SwitchAR, followed by a description of the tricks we applied in our implementation, in chapter four, and details on the implementation itself in chapter five. Chapter six presents the user study conducted to evaluate the system, including the study design and results. Chapter seven then discusses the findings, while chapter eight addresses the system's limitations and potential future work. Finally, Chapter nine concludes the thesis with a summary.

2 Related Work

Our research intersects with four key areas: it draws on concepts from Change and Inattentional Blindness and Cross Realities, and differentiates itself from prior work in Perceptual Manipulation within both Virtual Reality and Augmented Reality.

2.1 Change Blindness and Inattentional Blindness

Change blindness and inattentional blindness are two closely related phenomena that exploit limitations in human perception, allowing for changes in users' environments to occur without their awareness.

2.1.1 Change Blindness

As outlined by Simons et al., change blindness refers to "*the inability to detect changes to an object or scene*"[30]. This phenomenon can occur both after extended periods, such as when a user looks away from an object and returns to it later[32], and after brief interruptions like eye movements or a short occlusion of the entire FOV[30]. Change blindness has been applied in Virtual Reality (VR) to achieve various objectives, including optimizing rendering times by adjusting rendering fidelity[4], automatically turning pages without the user's awareness[35], and redirecting users to navigate space constraints during natural walking[32].

2.1.2 Inattentional Blindness

Inattentional blindness occurs when parts of the FOV that are outside the center of attention are not consciously perceived[28]. This effect can also cause users to miss changes, such as the appearance of unexpected objects[16].

In their study, Simons et al.[28] demonstrated that 46% of participants failed to notice a person carrying an umbrella or dressed in a gorilla costume walking through a scene while they were focused on counting basketball passes in a video. The addition of a distraction task, counting the passes, permanently captured participants' attention, inducing inattentional blindness. The study further found that increasing the difficulty of the distraction task reduced the rate of the unexpected person being detected.

In VR, Marwecki et al.[17] utilized inattentional blindness to develop a system capable of making unnoticed changes to the virtual environment within the FOV. Application scenarios included subtle difficulty adjustments, tailoring experiences to user preferences, and reducing motion sickness.

Both change blindness and inattentional blindness are fundamental to SwitchAR, providing the theoretical basis for enabling an imperceptible switch from pass-through AR to a virtual reconstruction.

2.2 Cross Realities

Recent advances have improved the quality and accessibility of HMDs that implement Augmented Reality (AR) using pass-through camera feeds. Standalone devices like the Meta Quest 3 and Apple Vision Pro are fully enclosed, displaying the pass-through feed on their screens to enable AR. This design allows them to seamlessly transition across various points on the reality-virtuality continuum[20, 19], including the ability to completely obscure the real world and operate in VR.

2.2.1 Application Scenarios

Sun et al.[33] introduced a system that monitors users' emotions and task efficiency. If the system determines that the user could be more productive or in a better mood in a different environment, it suggests switching between AR and VR to continue the task in the new setting.

Gottsacker et al.[10] explored various methods for visualizing interruptions by outsiders in VR. While most techniques involved representing the outsider virtually within the VR environment, one approach involved transitioning between VR and AR. In this case, the system would prompt users to activate the HMD's pass-through cameras, enabling interaction with the outsider in AR.

2.2.2 Switching Techniques

Beyond simply replacing the entire view from one environment with one from another in a single frame, there are various methods to implement transitions between different levels of the reality-virtuality continuum.
Pointecker et al.[23] developed and tested four different switching techniques in a user study. Participants' preferences varied depending on the scenario. For instance, a fading transition was favoured for situations requiring frequent environment switches, while walking through a portal was the preferred technique for transitioning between dissimilar environments.
In another project, Pointecker et al.[24] employed an intermediate replica environment, transitioning from AR to this replica before switching to VR. They compared various switching techniques in a user study, focusing on user experience. The results showed that incorporating the intermediate environment made the transition less confusing compared to a fade transition without the replica.

2.2.3 Imperceptible Transition

The concept of transitioning between different points on the reality-virtuality continuum is a core aspect of SwitchAR. Unlike the previous research on transitions, where users are aware of the transition, SwitchAR focuses on an imperceptible switch to enable the application of PMs like RDW while users continue to believe they are still in AR.
Achieving such an imperceptible switch requires the target environment to be a virtual reconstruction of the real-world environment visible in AR. Lindlbauer et al.[15] created a virtual reconstruction using depth cameras, allowing access to arbitrary virtual camera positions. This system enabled users to perform actions like removing objects from their environment or teleporting within it.
Kari et al.[12] used an iPhone and the Polycam app to create a virtual reconstruction of the environment. By aligning this reconstruction with the physical world, they could replace parts of the real environment with virtual counterparts. This setup allowed

users to experience virtual actions that appeared to affect the physical space, such as a virtual character lifting and moving a physical chair.
For SwitchAR, the reconstruction needs to closely resemble the real environment to enable a seamless switch from AR to the virtual environment without detection, replacing the entire environment rather than just parts. Thus, our work focuses on achieving a more realistic reconstruction compared to previous approaches.

2.3 Perceptual Manipulations in Virtual Reality

The concepts of perceptual manipulations (PMs), and more specifically virtual-physical perceptual manipulations (VPPMs), were first defined by Tseng et al.[34] and Bonnail et al.[2]. PMs are described as "*mechanisms grounded in limitations of users' cognition and perception with a clear intention to influence users towards a specific outcome*."[2]. VPPMs, a subset of PMs, specifically "*are perceptual manipulations that are grounded in visual-haptic limitations with the intention to nudge the user's physical movements*"[2].

2.3.1 Beneficial Applications

Several research efforts have explored VPPMs as tools to enhance the user experience. Razzaque[25] introduced a method for imperceptibly applying subtle rotational gains to users, allowing virtual paths to differ from real-world movements. This redirected walking (RDW) technique enables users to navigate virtual environments larger than the actual physical space while still relying on the real walking locomotion.
Suma et al.[32] later expanded on RDW by incorporating change blindness rather than relying solely on visual perceptual thresholds. In their approach, instead of applying rotational gains, doorways were repositioned behind users' backs while they were distracted by tasks like looking at monitors. This allowed users to explore multi-room environments significantly larger than their physical tracking space.
Another technique, haptic retargeting, was introduced by Azmandian et al.[1], who used visual perceptual thresholds to allow one physical object to provide haptic feedback for multiple virtual objects. By subtly adjusting the virtual representations of the user's hand, the virtual environment, or both, users would repeatedly reach for the same physical cube, while perceiving it as different virtual cubes.
Hand redirection has also been used to enhance ergonomics. Medeiros et al.[18] developed a technique to redirect hand movements in constrained environments, such

as airplane cabins. This allowed passengers to use physical surfaces for haptic feedback while maintaining a comfortable posture, regardless of the surfaces' orientations. Similarly, Murillo et al.[21] applied hand redirection to make mid-air interactions with virtual objects less fatiguing, thereby reducing fatigue and improving ergonomic comfort.
Combining RDW with haptic retargeting is another approach demonstrated by Clarence et al.[6]. This stacked retargeting increases the effective space for haptic feedback, reducing the need for multiple physical props and enhancing spatial interactions in virtual environments.

2.3.2 Malicious Applications

While many PM applications are beneficial, they also have the potential for malicious use. Tseng et al.[34] identified potential harmful applications through speculative design workshops, one of which was the "*puppetry attack*"[34]. In such attacks, users' body movements are manipulated in harmful ways. For instance, RDW could be used to guide a user toward a staircase, causing them to fall down. Haptic retargeting could make users unknowingly reach for hazardous objects, like knives present in their physical environment. Another malicious scenario is the "*mismatching attack*"[34], which assumes a virtual environment that closely mirrors the real world. In one example, a virtual-only chair is inserted into the scene, causing the user to fall when they attempt to sit down.
Casey et al.[3] explored potential vulnerabilities in HTC Vive and Oculus Rift VR systems, focusing on redirection attacks. By exploiting software vulnerabilities, they were able to redirect users without their awareness during VR gameplay in a user study.

While numerous PMs have been explored in VR, whether for beneficial or malicious purposes, these approaches remain confined to VR experiences. SwitchAR, on the other hand, extends these PMs, such as RDW, to video pass-through AR. The following section is going to cover research on PMs within AR contexts.

2.4 Perceptual Manipulations in Augmented Reality

While some research exists on applying PMs in AR, the work on VPPMs remains scarce. The key challenge lies in the fact that users can still see the real world, limiting the

ability to create sensory conflicts.
Cheng et al.[5] introduced PMs in AR by overlaying red and green virtual boxes on top of a 2D monitor, which also displayed red and green boxes. Their study aimed to influence users' reaction times to changes in the colours of the monitor's boxes. While this is an example of PMs in AR, it does not involve VPPMs.
Ishii et al.[11] implemented a technique similar to redirected walking (RDW) in AR, but with limitations in terms of redirection capabilities. They achieved this by cropping a portion of the pass-through video feed and shifting the cropped area to the side. This caused participants to drift left or right instead of walking straight. However, this technique was constrained by the limited space available in the camera feed to crop out of, preventing continuous redirection in the same direction. Additionally, participants had to walk in straight lines without turning their heads to the side, and the system couldn't handle turns around 90-degree corners. The redirection was only possible due to this custom cropping method, rather than reusing established VR techniques.

Just as PMs in VR enable malicious applications, this is also the case for PMs in AR. Roesner et al.[27, 26] and Lebeck et al.[13, 14] highlighted potential risks associated with AR head-mounted displays (HMDs), such as their ability to obscure parts of the user's field of view (FOV), including critical objects like traffic signs. To address these risks, they discussed and developed prototypes for safety mechanisms that control the visual output of AR applications.

The research mentioned applied PMs in AR[5] and redirected aware users using a custom video cropping technique[11]. However, they did not explore the possibility of physically manipulating unaware users using existing VR-based VPPMs in AR. SwitchAR addresses this gap by imperceptibly transitioning users into a virtual reconstruction of their environment while they believe they are still in AR. The steps of this process will be explained in the following section.

3 Concept

This section outlines the four key stages of SwitchAR's pipeline, as depicted in figure 3.1. The pipeline combines reconstructing the user's environment, switching from AR to this virtual reconstruction, and applying the desired PMs to enable PMs such as RDW in AR. An optional fourth stage, switching back to AR, can ensure a seamless experience.

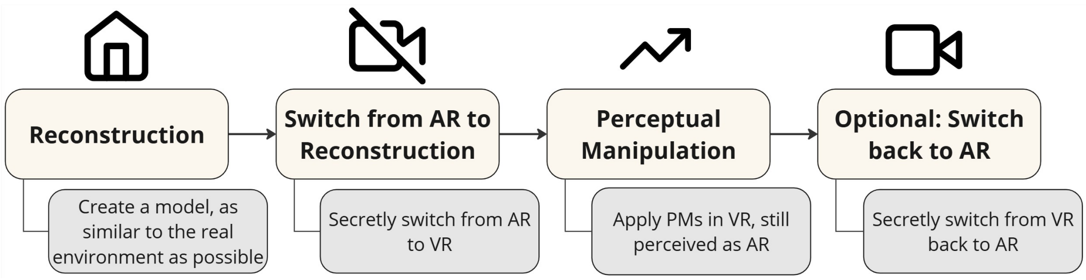

Figure 3.1: SwitchAR pipeline, illustrating the four key steps and their respective objectives.

3.1 Reconstruction

The first step is creating a virtual reconstruction of the environment where the application will be used. The reconstruction should closely resemble the real environment in terms of geometry, scale, alignment, and texture. This similarity is essential for utilizing inattentional blindness and change blindness to reduce the probability that the user will notice the subsequent switch.

3.2 First Switch

Once the reconstruction is ready, the next goal is to switch the HMD from AR to VR without the user noticing. Hiding this switch relies on inattentional and change blindness. A simple approach could be to instantly turn on the virtual reconstruction, occluding the real pass-through feed to switch to VR. However, such an abrupt change can attract attention and be more noticeable than gradual transitions[37, 38, 9]. An alternative, less intrusive approach could be to fade in the virtual reconstruction gradually, increasing its opacity until it fully covers the pass-through feed. Such an approach could minimize the abruptness and therefore the noticeability of the switch.

3.3 Perceptual Manipulation

Once the user is in VR, the full range of established PMs from VR research can be applied, such as haptic retargeting[1, 6] or redirected walking[25, 32]. During this phase, users should remain unaware that they are no longer in AR, so it is important not to exceed perceptual thresholds, which might alert them.
However, applying PMs by manipulating a virtual reconstruction introduces a side effect unique to SwitchAR, which does not occur in VR, even with the same PMs. Unlike in standard VR, where no alignment between the virtual and real world is necessary, SwitchAR requires the virtual reconstruction to initially match the real environment with regard to position and orientation for an imperceptible switch (step 1 in figure 3.2). Once the virtual scene is manipulated, for example using rotational gains in RDW, this alignment is disrupted, creating misalignment between the real and virtual environments (steps 2 and 3 in figure 3.2). For the manipulation to go unnoticed throughout, this misalignment needs to be corrected later on, which leads to the final step.

3.4 Optional Second Switch

In this final step, the HMD can switch back from VR to AR, relying on the same principles as the first switch and again without the user realizing. To achieve this, any misalignment caused by previous manipulations must first be corrected (steps 4 and 5

in figure 3.2). This can be done by either reversing the past manipulations or directly restoring the virtual reconstruction's alignment with the real world. As a result, the system can switch back to AR seamlessly, enabling the virtual reconstruction and the PMs to stay undetected.
This second switch is optional and depends on the intended application. If the manipulation only needs to happen once, or if it will eventually become apparent to the user anyway, there may be no need to resolve the misalignment or return to AR.

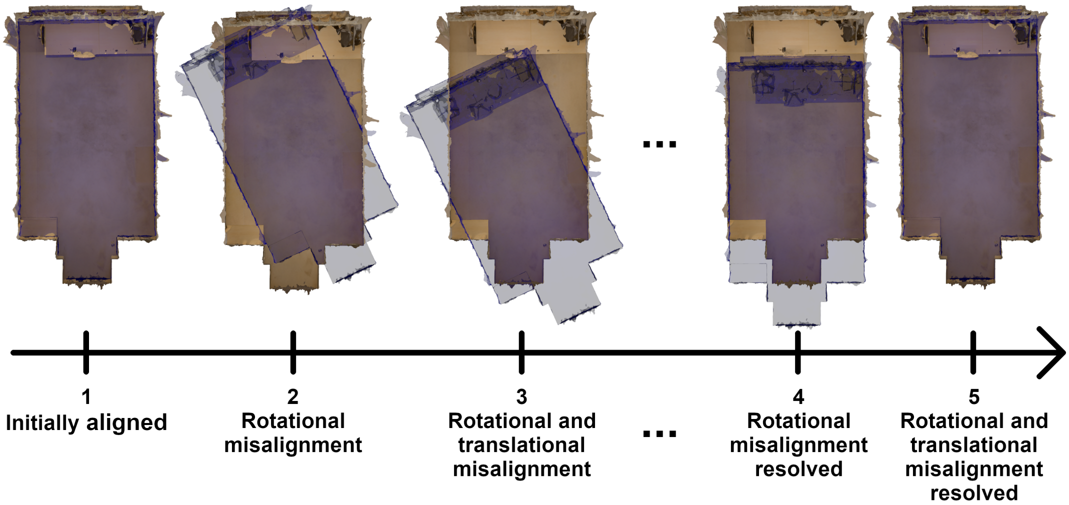

Figure 3.2: Sample timeline showing the emergence of misalignments (steps 1-3) between the physical environment and the 3D reconstruction (shown in blue for visualization purposes), and their resolution (steps 4-5).

4 Tricks

Name	Description	Goal
Style Transfer	Match the reconstruction's colour representation to the pass-through feed more closely	Visual similarity
Visual Noise	Compensate for missing camera noise on the reconstruction	Visual similarity
Virtual Cubes	Let user interact with virtual cubes in AR	Remind user of the AR scene they interacted with in the beginning
Body Representation	Cover the missing body in VR with a virtual representation	Visual similarity
Distraction Task	Distract the user, increase cognitive load	Distraction, cognitive load
FOV Covering	Reduce the amount of visible change by covering part of the FOV with a virtual painting	Reduce visibility of the changes

Table 4.1: Overview of the custom tricks applied in our SwitchAR implementation

Building on the conceptual pipeline, this section details the tricks employed at various stages of the implementation, as summarized in Table 4.1. Figure 5.1 visually maps these tricks to the corresponding pipeline stages.

Most of these tricks leverage two key principles. First, research on inattentional blindness shows that unexpected events are less likely to be noticed if they resemble objects that are purposefully being ignored [22, 8, 36]. In SwitchAR, this means it is advantageous to increase the similarity between the virtual reconstruction, which should not be noticed, and the real world, which users are ignoring while focusing on virtual objects.

Second, cognitive load introduced by distraction tasks can also reduce the rate of detection of unexpected changes. A distraction task with increased difficulty facilitates inattentional blindness[28]. Thus, SwitchAR can benefit from the implementation of a distraction task, contributing to inattentional blindness to make it harder for users to detect the switch from AR to the virtual reconstruction.

4.1 Style Transfer

To create a high-quality reconstruction, we built and used a camera rig[1] with three mirrorless cameras to take images of the room, rather than relying on frames from the HMD's pass-through capture. However, this introduced differences in colour representation between the captured images and the HMD's pass-through feed, reducing the visual similarity between the virtual reconstruction and the AR environment.
To address this, we applied a style transfer based on wavelet transforms [39] to the images obtained from the camera rig before processing them to produce the reconstruction. A set of screenshots from the Meta Quest 3's pass-through video feed served as the style reference. Before applying the style transfer on an image, it was assigned the most similar one of the reference images.
As real-time access to the Quest 3's full-resolution camera feed is currently not possible, we applied the style transfer on the input images in advance, rather than on the AR camera feed in real-time. Once real-time access is available, the style transfer could potentially be applied directly to the live camera stream for even better style alignment.
The style transfer reduced colour representation discrepancies, enhancing the reconstruction's texture so that it matched the AR pass-through's visuals more closely.
Figure 4.1 demonstrates the effect on a single input image, and figure 4.2 compares the pass-through AR with a reconstruction based on unmodified images and the final reconstruction after style transfer.

4.2 Visual Noise

When the final reconstruction could be viewed in the HMD, a further source of discrepancy between reconstruction and AR became apparent. The Quest 3's pass-through cameras introduced some grain and therefore visual movement into the pass-through

[1]https://rd.nytimes.com/projects/assembling-a-camera-rig-to-capture-complex-spaces-in-3d/, last visited: 25.09.2024

Figure 4.1: Comparison of one of the raw input images with the result of applying the style transfer to it

Figure 4.2: Comparison of the pass-through feed with reconstructions before and after the style transfer

AR. This made the virtual reconstruction, lacking any visual movement, appear static and artificial in comparison.
To address this, we added pseudo-random visual noise to the system by integrating an asset from the Unity Asset Store [2]. This noise, which was stronger than that from the pass-through cameras, was applied to both the AR and VR sections of the application. It was applied to everything, that should be perceived as part of the real world, therefore excluding purposely virtual elements like paintings and cubes.

[2]https://assetstore.unity.com/packages/vfx/shaders/fullscreen-camera-effects/old-movie-270021, last visited: 25.09.2024

The noise helped blend the reconstruction into the real world more seamlessly, and also masked minor defects in the 3D model and its texture.
Figures 4.3 and 4.4 illustrate how the noise affects the pass-through feed and the final reconstruction, enhancing the similarity between them.

Figure 4.3: Effect of the visual noise on the pass-through video feed

4.3 Virtual Cubes

To further enhance the sense of continuity between the AR environment and virtual reconstruction, users were encouraged to interact with a few virtual cubes in AR before the actual task and the switch to VR. Using the Quest 3's hand tracking, users could pick up, stack, and throw the cubes, creating unique structures and getting familiar with interacting with a pass-through AR application.
These personalized cube configurations were replicated in the virtual reconstruction after the switch, reinforcing the illusion that the virtual scene was the same as the AR environment. This trick aimed to strengthen the users' belief that they were still interacting with the AR scene whenever they saw their custom cube arrangement on the cupboard.
The virtual cubes are included in several figures, among them figure 4.4.

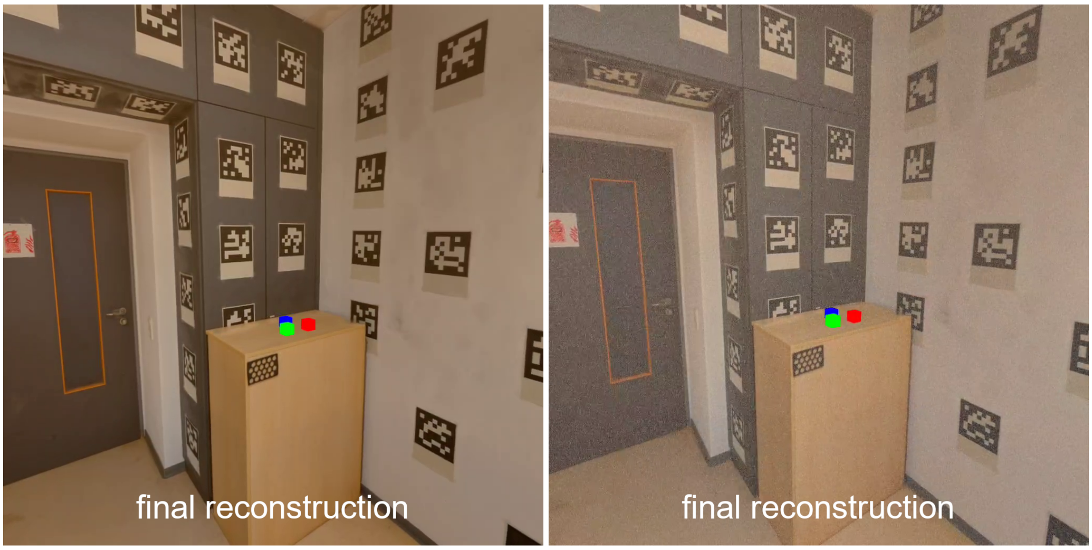

Figure 4.4: Effect of the visual noise on the reconstruction

4.4 Body Representation

After the switch from AR to the virtual reconstruction, the user is effectively using VR, which means that their physical body is no longer visible. The absence of the body might break the illusion of still using AR, especially when the body becomes the focus of attention. Whenever it is not the focus of attention, the missing body again reduces the similarity between the virtual reconstruction and the pass-through AR environment. To mitigate this source of dissimilarity and the risk of breaking the AR illusion, we added a virtual avatar representing the user's body both in AR and the virtual reconstruction. The avatar, sourced from Mixamo[3], was slightly larger than the user's actual body, particularly around the arms. This allowed it to cover large parts of the real body and thus obscure the fact that the body was not visible in VR.
The issue of the invisible body can be seen in figure 4.5, where the user's arms disappear in the virtual reconstruction. First, we added virtual hand visualizations provided by the Quest 3's hand tracking implementation (figure 4.6).

[3]https://www.mixamo.com/#/, last visited: 25.09.2024

Figure 4.5: Comparison of the pass-through feed and the reconstruction without any virtual body representation

Figure 4.6: Comparison of the pass-through feed and the reconstruction with a virtual hand representation

This was followed by a full-body avatar. With the Quest 3's inside-out tracking, upper body tracking and lower body prediction can be applied to rigged avatars. Figure 4.7 shows the combined hand and avatar representations, while figure 4.8 presents the user's view of their virtual body when looking down toward their feet.
The combination of hand visualization and full-body avatar reduced the effect of the missing body, potentially making the switch to the virtual reconstruction less noticeable.

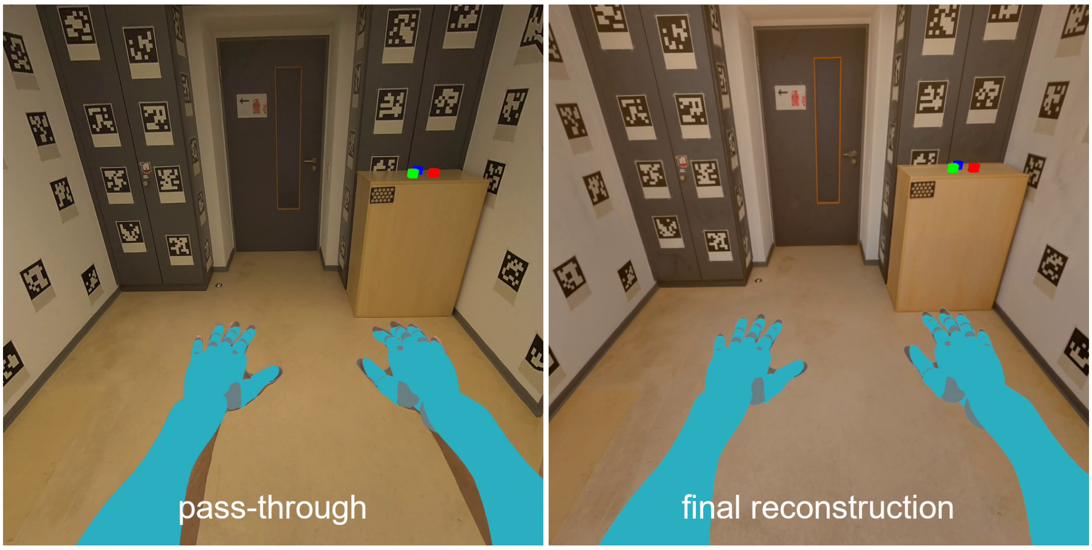

Figure 4.7: Comparison of the pass-through feed and the reconstruction with both the virtual hands and the avatar

4.5 Distraction Task

Without a task to focus on, users might explore their surroundings too closely, reducing opportunities to exploit inattentional blindness.
To counter this, we implemented a distraction task that required users to memorize the content of several virtual paintings, one after the other, within a limited time.
The time pressure while having to memorize paintings introduced cognitive load, which heightened inattentional blindness. Additionally, guiding the users' focus toward the

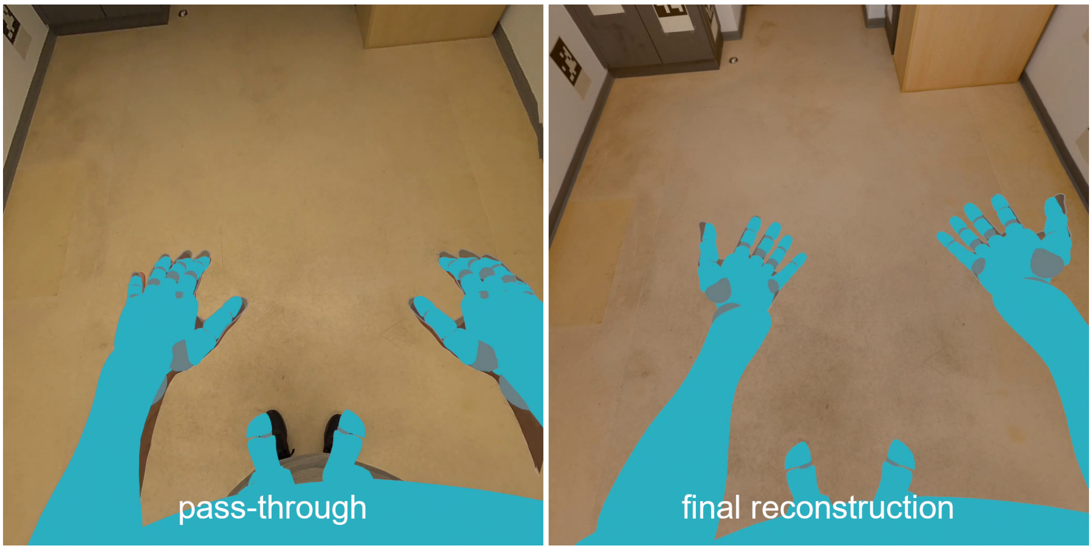

Figure 4.8: Comparison of the pass-through feed and the reconstruction when looking down on the feet in the final implementation

paintings and away from the environment decreased the opportunities for them to notice changes in the environment caused by the switch from AR to VR.

4.6 Covering a Large Part of the FOV During the Switch

Executing the switch while the user is actively looking at their environment increases the risk of visual changes between the virtual reconstruction and the real environment being detected.
To reduce this risk, we ensured that a large portion of the user's FOV was occupied by virtual content during the switch. This limited the exposure to visual discrepancies and consequently improved the chance of inattentional blindness preventing the user from noticing the switch.
This was achieved by incorporating the distraction task, which already involved looking at virtual paintings. Initially, these paintings were hidden behind a grey placeholder. The users had to center their FOV on the painting while standing close to it to reveal its content. While the painting occupied a large part of the FOV, the switch to the virtual reconstruction could then be performed in secret.

Figure 4.9 shows screenshots approximating the user's FOV immediately before (”pass-through”) and after (”final reconstruction”) the switch.

Figure 4.9: Comparison of looking at a painting in the pass-through feed with in the reconstruction

5 Implementation

Figure 5.1 provides an overview of the steps involved in implementing SwitchAR. This section is going to explain how we implemented each of them in Unity (version 2022.3.10f1) to enable RDW in AR.

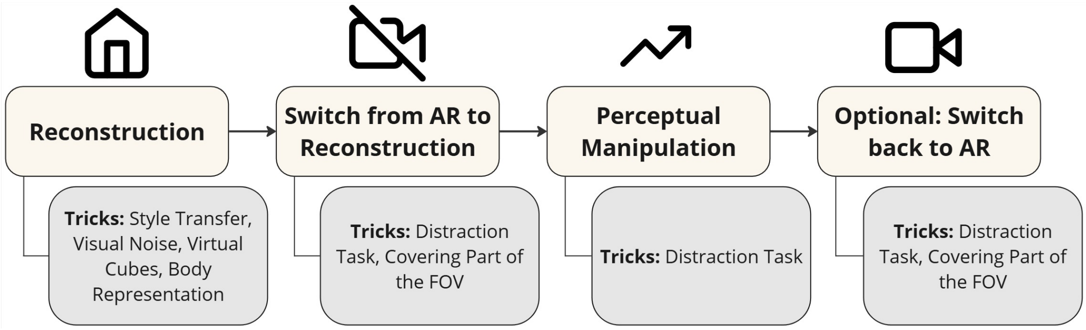

Figure 5.1: SwitchAR's pipeline, with the tricks applied at each of the four steps

5.1 Reconstruction

5.1.1 Using Photogrammetry To Create a Virtual Model

Since the Quest 3 currently does not provide real-time access to its pass-through camera feed, generating a virtual reconstruction of the environment at runtime is not feasible. To work around this, we prepared a room which stayed unchanged for the duration of the project. This allowed us to generate a photogrammetry-based reconstruction in advance, which could then be used in the application later.
To make the photogrammetry process easier, we took several preparatory steps. First,

we removed most of the furniture, simplifying the room's geometry. We then blocked out natural light from the windows to maintain consistent lighting from the ceiling lamps, independent of the time of day and weather. Additionally, we placed ArUco markers on large, featureless surfaces such as white walls and grey cupboards. These markers helped prevent picture alignment issues, which led to large holes in the mesh, by ensuring that even in featureless areas, there were enough visual cues for proper alignment.

After preparing the room, we assembled a custom camera rig equipped with three Sony Alpha 7R IV cameras and Tamron 28-200mm 1:2.8-5.6 Di III RXD lenses. We then used the rig to capture approximately 3000 images of the room. Before feeding these images into the photogrammetry software, we applied the style transfer described in section 4.1 to ensure that the colours in the final reconstruction would match the pass-through visuals from the Meta Quest 3.

Using RealityCapture[1], we processed these images to create both the 3D mesh and the texture for the virtual reconstruction. For performance optimization, we simplified the mesh to 250 000 triangles, ensuring that the Quest 3 could render it smoothly without compromising the frame rate. During the process, we encountered an issue with dark lines appearing at texture seams when combining multiple textures in Unity. To resolve this, we restricted the reconstruction to use a single high-resolution texture (16384 x 16384), which eliminated the seam artifacts.

The final reconstruction still contained a few defects, primarily in areas that users were unlikely to focus on, such as some holes in the ceiling (see figure 5.2). One more prominent defect was the bad reconstruction of a ceiling lamp, which we anticipated could occasionally be within the users' FOV. To address this, we manually replaced the lamp from the photogrammetry reconstruction with a custom 3D model (see figure 5.3 for a comparison of both versions).

Additionally, for our user study, we placed a virtual curtain in the environment to block off one end of the room. This curtain allowed the interviewer to remain physically present in the room but hidden, preventing the risk of breaking the AR illusion when the interviewer would disappear in the secret switch to VR. Figure 5.4 shows how the curtain was positioned to hide part of the room.

[1] https://www.capturingreality.com/, last visited: 25.09.2024

Figure 5.2: Example of some of the final reconstruction’s defects

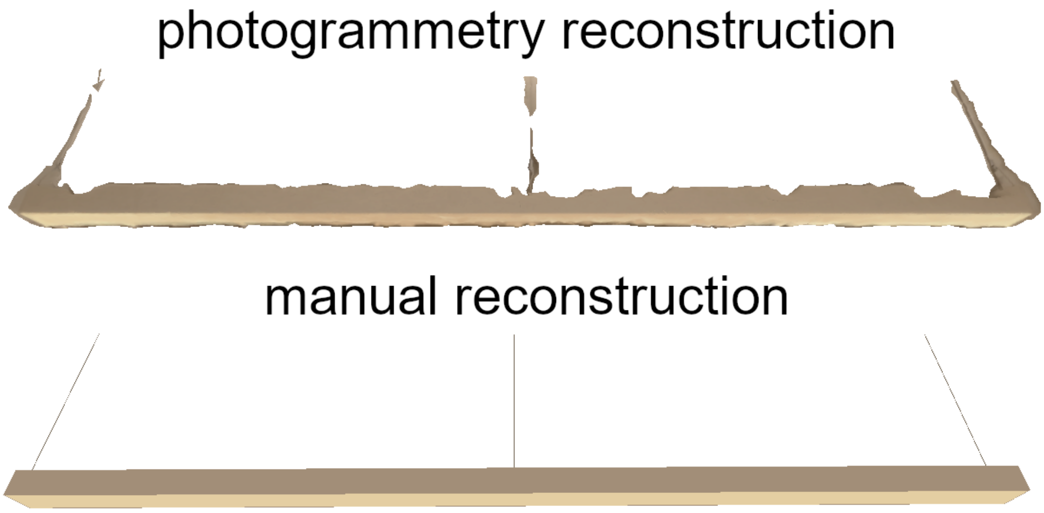

Figure 5.3: Comparison of the lamp from the photogrammetry reconstruction with the manually modelled one

Figure 5.4: A virtual curtain hiding one end of the room

5.1.2 Scale and Alignment

To ensure that the virtual reconstruction aligned with the real-world environment, we implemented a manual two-point alignment system. Using the Quest 3 controllers, we placed two anchor points in the physical environment, which allowed us to scale, position, and rotate the virtual reconstruction to match the real room. Once the alignment was satisfactory, we utilized the Quest 3's spatial anchor feature to store the alignment, so that it persisted across application sessions without needing to be recalibrated each time.

5.1.3 Perceptual Tricks

After completing the virtual reconstruction, we integrated some of the tricks described earlier. These included adding visual noise (section 4.2), incorporating virtual cubes for interaction (section 4.3), and providing a virtual body representation for the user (section 4.4).

5.2 Switching From AR to VR

Once the final virtual reconstruction of the room was complete, we were able to toggle between pass-through AR and VR. Although the small 2D screenshots cannot fully capture the experience inside the HMD, figure 5.5 provides a side-by-side comparison of the pass-through AR and the final virtual reconstruction.

Figure 5.5: Comparison of the pass-through feed and the reconstruction in the final implementation

5.2.1 Transition

For the AR to VR switch, there are several potential methods to consider. After internal testing, we settled on a gradual transition for the final implementation. Initially, we experimented with an instant switch between the pass-through AR and the virtual reconstruction. However, this caused noticeable visual movement in the users' peripheral vision. This effect can be explained by the higher density of rods in the periphery of human vision[7], which are highly sensitive to motion. The visual motion resulted from inaccuracies in the reconstruction's mesh, texture and particularly its alignment,

causing objects, such as the edge of a cupboard, to appear to jump in one direction slightly when switching to the reconstruction.
To mitigate this, we implemented a fade effect over three seconds, gradually increasing the opacity of the virtual reconstruction. This reduced the intensity of perceived visual movement by spreading any misalignments or changes over time, making the switch less noticeable.

5.2.2 Perceptual Tricks

During this transition, we also applied the distraction task discussed in section 4.5, as well as the strategy of covering a large part of the user's FOV, as detailed in section 4.6.

5.3 Perceptual Manipulation

Once users have been transitioned from AR to VR without their awareness, they can be manipulated with PMs that are only possible in VR, while they still perceive themselves to be in AR. The PM used in our implementation is RDW[25].
Initially, we tested curvature-based RDW, which performed well. However, for our purposes, we found rotation-based RDW to be more predictable, particularly in terms of anticipating where users would end up.
In scenarios where users were expected to turn 180 degrees, our implementation applied small rotation gains, subtly rotating the virtual environment around the user. This amplified or reduced their rotations by up to five percent, enabling us to guide them along a predictable, sawtooth-like trajectory instead of walking between the same two points repeatedly. Figure 5.6 illustrates this concept, with the green line representing the users' perceived straight path, while the red line shows the actual path taken by one of the participants during the user study.
In our implementation, we applied fixed relative rotation gains, without aiming for a specific predefined total gain. This meant that variations in how users turned and walked led to differing degrees of rotation manipulation, resulting in different levels of offset between the virtual reconstruction and the real-world environment.
While haptic retargeting[1] was not part of our user study, the level of control we achieved over users' rotations with this RDW implementation would have been sufficient to implement haptic retargeting if needed.

Figure 5.6: Top-down view of the room, including the paths one of the participants walked relative to the virtual reconstruction (green, straight) and relative to the real room (red, sawtooth-pattern)

5.4 Switching Back From VR to AR

Our implementation also included a method for mostly resolving the misalignment in the virtual reconstruction that occurred due to RDW, allowing to switch users back from VR to AR without them noticing it. This was achieved by applying the inverse of the rotation gains used before, in reverse order. This realigned the virtual reconstruction with the real world.
To ensure the second switch remained unnoticed as well, we again utilized the distraction task (section 4.5) and made sure a large part of the user's FOV was occupied by virtual content (section 4.6).

6 User Study

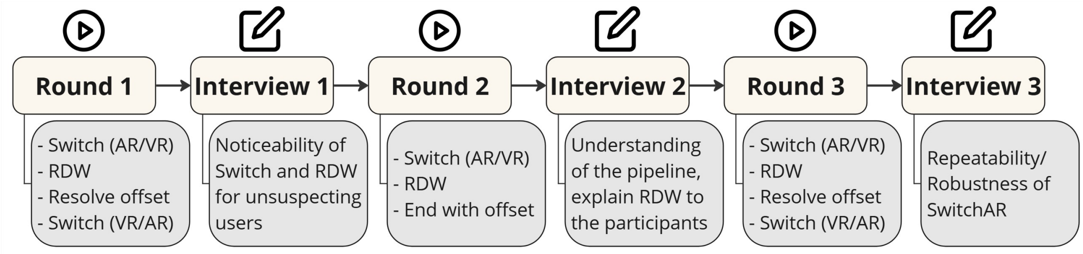

Figure 6.1: Procedure of the user study

6.1 Study Design

To validate our implementation of SwitchAR and therefore the feasibility of applying RDW in AR, we designed the following user study to explore three key aspects: the noticeability of the AR/VR switch, the noticeability of the RDW, and the repeatability of SwitchAR. Repeatability refers to whether the switch and RDW remain effective even after users have experienced the misalignment and are aware of the RDW manipulation.

As illustrated in figure 6.1, the study followed a sequence of three rounds, with each round consisting of participants interacting with the system and immediately being interviewed. While using the system, participants walked back and forth between two picture frames located on opposite walls. Upon reaching a frame, they had ten seconds to observe a new picture and memorize its content.
When the participants focused on the first picture, the application secretly switched from AR to VR in all three rounds. From that point onward, whenever participants turned around to walk toward the next picture, rotational gains were applied, gradually

creating misalignment between the virtual reconstruction and the real environment. This created an illusion that the picture frames stayed in their original positions, even though they were slowly shifting sideways relative to the physical space.
In the first and third round, the system then restored the original alignment of the virtual and real environments, allowing to switch participants back from VR to AR while still wearing the HMD, as described in section 5.4.
After the final picture in each round, participants were told to stand in front of one of the walls, facing a logo, and then take off the HMD. This setup helped them detect any misalignment by making offsets along the wall more noticeable when removing the HMD.

6.1.1 Initial Setup and Instructions

Before the first round began, participants were given a brief explanation of AR and the study's purpose. Since we could not disclose the hidden switch to VR at this stage, participants were told that the study focused on memory abilities in AR. We ensured they were debriefed at the end of the experiment, revealing the true nature of the manipulation.
Next, the participants signed a consent form, and the interviewer started a video recording and set up the application on the headset. Finally, the following process was explained, telling the participants the steps they should carry out once they put on the headset.

6.1.2 First Round: Unsuspecting Users

The first round began with participants interacting with the virtual cubes using hand tracking, as described in section 4.3. Following this, the system switched from AR to VR, applied RDW to induce an offset, then resolved the offset to finally return the participants to AR. This round served as a test of the system's functionality under optimal conditions, with completely unaware participants.
After this initial round, the first interview was conducted, aimed at gathering two types of insights. First, participants were asked open-ended questions about whether they noticed anything unusual, giving them an opportunity to report any unexpected occurrences.
Secondly, we assessed whether the RDW had been detected. Participants were asked to mark the positions of some of the paintings on a 2D top-down map of the room. If all

the paintings were marked in the same two positions, we could infer that participants did not notice the shifting positions. In cases where paintings were placed next to each other on the map, participants were asked to clarify whether they actually believed the paintings had moved, or whether they only placed them next to each other to prevent intersecting markings.

6.1.3 Second Round: Revealing the Offset

In the second round, the system did not restore the alignment of the virtual and real environments, meaning that there was no switch back to AR and the offset was preserved. As a result, we expected participants to notice the offset when they removed the HMD. This round was designed to explore participants' reactions and their explanations for being in an unexpected location in the room.
The second interview followed the same structure as the first, beginning with an open-ended question about anything participants had noticed. Afterward, participants were asked to explain how they thought the offset had occurred, which provided insights into their awareness of the different parts of SwitchAR's pipeline.
Following their explanation, participants were informed that RDW had been applied to subtly guide them to a different spot than they expected. At this point, they were not told about the switch to VR, though. Finally, participants were asked to mark the positions of a different set of paintings on the map.

6.1.4 Third Round: Repeatability Test

The third round followed the same process as the first, except that the cube interaction was omitted. As in the first round, the virtual world was misaligned through RDW and then realigned before switching back to AR. This round aimed to test the repeatability of SwitchAR after participants had experienced the misalignment in the second round and had learned about the RDW manipulation.
The final interview again began with the open-ended question about things noticed by the participants and the painting position task. This was followed by a new question. Participants were asked to explain why the offset occurred in the second round but not in the others.

Afterward, a set of four final questions was asked to determine participants' awareness of our pipeline:

1. Did you notice that the environment changed?
2. Have you been in AR for the whole duration of the study?
3. Did you notice a point in time when you changed from AR to VR?
4. Did you notice that there was a photogrammetry/virtual reconstruction of the room?

This sequence of increasingly specific and revealing questions was modelled after the methodology used by Simons et al.[28] in their inattentional blindness study.

6.1.5 Debriefing

Following the conclusion of the final interview, participants were debriefed. The interviewer explained the true nature of the experiment, including the switch from AR to VR, the purpose of the memorization task as a distraction, the application of RDW, and the actual goal of the study. Participants were also offered to revisit the virtual reconstruction using the HMD, where we highlighted specific cues and defects that differentiated it from the real-world pass-through camera feed.

6.2 Participants

We recruited 20 participants for the user study, with an average age of 24.58 years (SD = 3.92), ranging from 20 to 35 years old. One participant preferred not to disclose their age. Thirteen participants were female, and seven male.
Before the study, six participants were familiar with the concept of RDW, while three were aware of photogrammetry. However, none of the participants had previous experience using photogrammetry themselves.
The average duration of each experiment was 49 minutes, and participants were compensated €15 for their time.

6.3 Data Analysis

For the open-ended questions about whether participants noticed anything unusual, responses were coded to identify instances where participants explicitly mentioned the switch from AR to VR.
For the final four questions, two of the authors collaboratively reviewed and coded the responses, grouping participants into four categories based on their level of understanding of the SwitchAR pipeline. The transcribed responses, along with the codings and group classifications, are included in appendix B for reference.
When participants placed the painting positions on a map, most of them explicitly stated whether they believed the paintings remained in the same two positions or moved along the wall. In five cases, participants did not provide an explicit statement. But, as they placed their markings in clusters with a maximum distance of less than four millimeters (converted to actual space on the tablet used), these small discrepancies were attributed to inaccuracies in using the trackpad. Thus, the markings were considered to represent the same locations.

6.4 Results

6.4.1 Noticing the Switch or RDW During Usage

As shown in figure 6.2, no participants reported noticing the switch between AR and VR in any of the three rounds. Responses such as "*No. It's always the same*" (P6), "*I just followed the arrow, I didn't notice anything*" (P12), "*Ah, no*" (P13), "*No, nothing*" (P19), and "*Nope. Should anything be noticeable?*" (P20) were typical. Participants who mentioned noticing something did not refer to the switch, instead commenting on unrelated aspects, such as the duration for viewing each painting being too short (P3) or a change in the number of paintings (P10). The full responses are included in appendix A.
In the first round, 18 participants did not perceive any variation in the painting positions, with only P11 and P14 reporting differences. In the second and third rounds, P11 was the only participant to report changes in the positions of the paintings. Therefore, P11 was the only participant to consistently do so in all rounds.

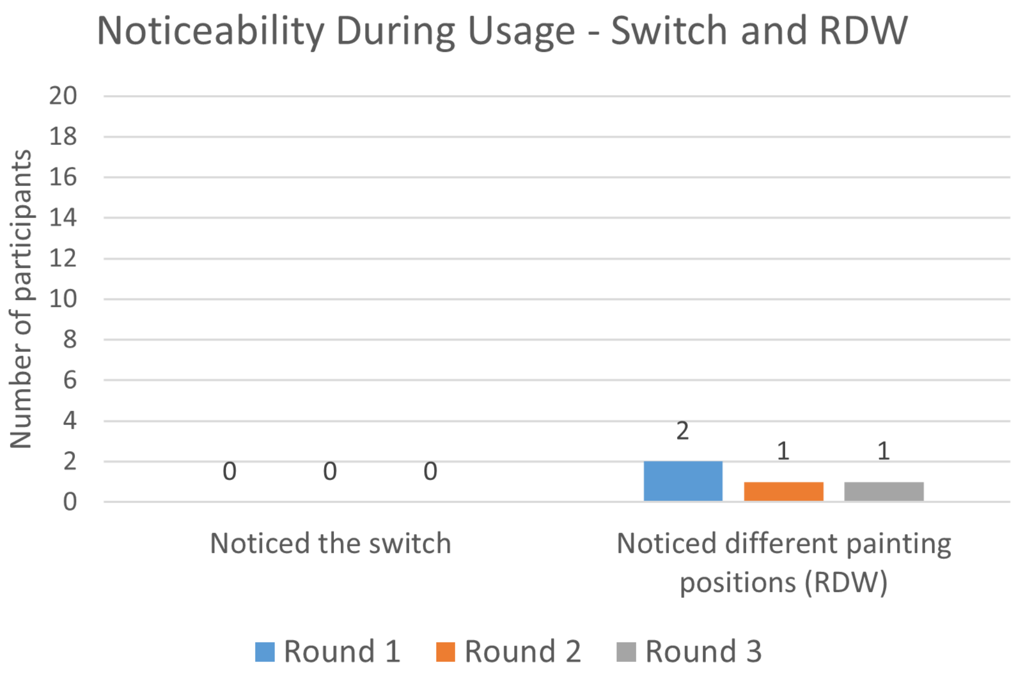

Figure 6.2: Number of participants noticing the switch and varying painting positions

6.4.2 Noticing the Offset at the End of Round Two

The intentionally revealed offset when participants removed the HMD at the end of the second round had a mean value of approximately 55cm, with a standard deviation of 21cm. 16 participants did not initially notice or react to the offset. They were prompted to look again at the logo on the wall, both through the HMD and without it, to help them recognize the displacement. Even after this, three participants failed to notice the offset. Notably, while the offsets for P13 and P19 were relatively small (17cm and 34cm, respectively), P3 had an offset of about 84cm, yet they still did not detect it. This suggests the occurrence of change blindness, potentially due to the brief visual interruption caused by removing the HMD.

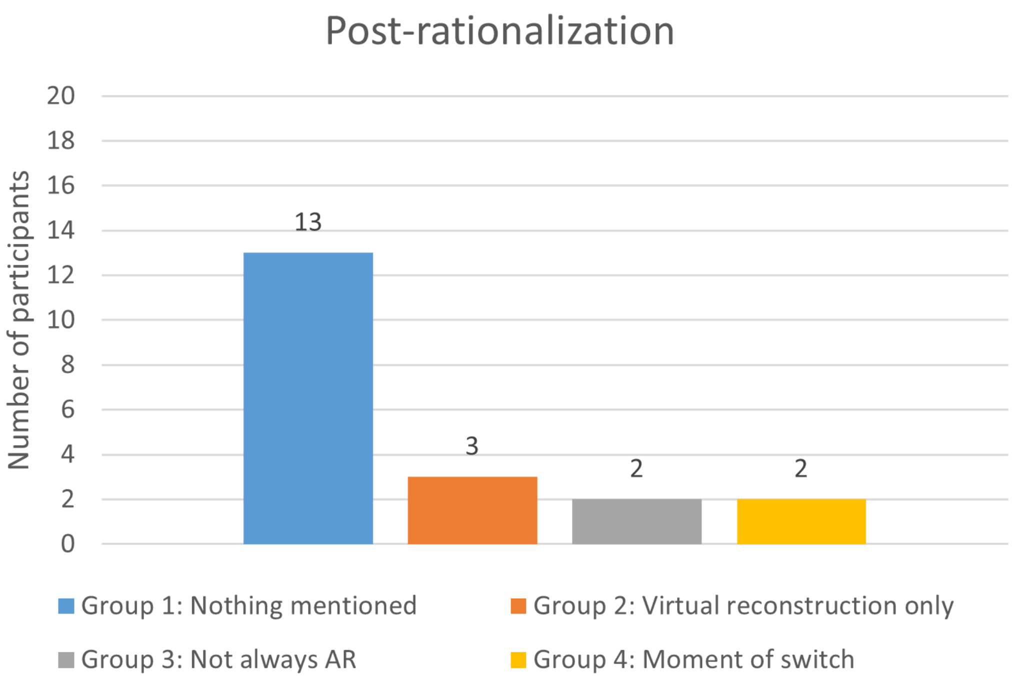

Figure 6.3: Post-rationalization: Level of understanding of SwitchAR's pipeline

6.4.3 Post-rationalization

We used participants' answers to the question "Did you notice anything?" after each round to assess their ability to detect the switch to VR. The final four questions gradually revealed details about the system's mechanics (e.g., asking about VR and the virtual reconstruction), which is why participants' responses were regarded as post-rationalizations. We were interested in whether participants, once prompted, could deduce how and when the switch could have occurred.

Based on their responses, participants were coded into four groups reflecting their level of understanding of the SwitchAR system (figure 6.3). The thirteen participants in group 1 exhibited the lowest level of understanding, reporting no awareness of the switch or the virtual reconstruction, e.g., "*Nope. I thought that was the room the whole time.*" (P3).

The three participants in group 2 (P5, P8, P19) mentioned the possibility of a virtual

reconstruction only after it was explicitly suggested to them, e.g., "*That was my theory about the shift. But I thought it wouldn't look that realistic.*" (P5). None of them reported anything related to the switch or the reconstruction at previous questions.
Group 3, containing two participants (P7, P15), suspected in hindsight that they were not always in AR but were unable to identify when the switch occurred, e.g., "*I'm not sure about round two. Because if I had seen the whole real world, I would have seen the logo in the right place. And I didn't.*" (P15).
Finally, the remaining two participants, assigned to group 4 (P12, P17) correctly deduced when the switch likely happened, e.g., "*Mh, probably when I looked at the pictures. And then focused on... the picture frames.*" (P17). However, even these participants did not notice the switch during the experiment itself. P17 for example answered with "*Mmm, yes*" when asked if they were using AR continuously before.
Altogether, no participant in any of the groups was able to notice the switch while wearing the HMD.

6.5 Anecdotes

Several interesting observations arose during the interviews with participants. In the second interview, after revealing the offset, 17 participants offered incorrect or no explanations for how it occurred. A common wrong explanation, given by four participants (P3, P5, P15, P20), was that the real logo on the wall was obscured by a virtual white rectangle, with a virtual logo displayed next to it, making it seem like the logo's position had shifted. Only three participants (P5, P7, P12) correctly hypothesized that the offset was realized with the usage of VR, though none of them indicated having noticed the switch from AR to VR.
In the final interview, when asked why the offset only appeared in the second round, participants offered a variety of theories. Some, like P16, suggested it was an error that did not occur consistently across rounds, while others, such as P2, believed the manipulation failed in the third round because they were paying more attention to their own movements. Two participants (P12, P15) explicitly stated that they were not using VR during the third round, despite the switch to VR occurring in all rounds. For example, P12 commented, "*I think this time, I see the real world. And last time, I, maybe I just see a virtual world, maybe*". Except for two participants (P7 and P11), most did not demonstrate a better understanding of the system by the end of round three compared to right after the second round. P7 and P11 displayed an increased understanding of the PM, with P11 even mentioning perceiving rotational manipulations in hindsight.

Even when directly asked in the final questions if they noticed any changes in their environment, none of the participants mentioned anything related to the virtual reconstruction of the room beyond the offset revealed in the second round. Some participants stated that they were too focused on the memorization task to observe their surroundings. Others mentioned actively looking for changes in the third round but still failing to notice anything unusual, like P15 stating "*I didn't find anything, even though I specifically looked around several times*".

7 Discussion

In the following, we will discuss our findings along three main topics: (1) *Feasibility*, (2) *Generalizability* and (3) *Misuse*.

7.1 Feasibility

Our study demonstrated the feasibility of SwitchAR, as no participant noticed the switch between AR and VR in any of the three rounds. This confirms that SwitchAR can imperceptibly transition users from an AR environment to a VR environment while maintaining their belief that they are interacting in the real world (AR). Furthermore, with only one participant detecting rotational gains, our results suggest that RDW can also be effectively applied within AR using SwitchAR.
The third round, which followed the participants' awareness of RDW after round two, tested the repeatability and robustness of the system. For 18 out of 20 participants, the knowledge of RDW did not diminish the effectiveness of SwitchAR or improve their ability to better understand it. This finding aligns with Simons' work in a follow-up project[29] of their original experiment about inattentional blindness[28], where expecting a specific unexpected event, in our case RDW, did not enhance participants' ability to notice other unexpected occurrences, like the switch between AR to VR in our study.
After the offset reveal in round two, participants were informed about RDW, and the third round ended without any offset. Only three participants (P7, P11, P19) correctly linked it to the RDW altering their walking paths. The remaining 17 participants offered no or incorrect explanations for this difference, thinking that it was caused by changes in their own behaviour or that the RDW was absent in the third round, despite its consistent application throughout the entire study. Two participants (P12, P15) who deduced the potential use of VR in the interviews later claimed there was no VR involved in the third

round. Despite actively paying attention to the environment in the final round, P15 did still not notice the switch or the reconstruction. This suggests that SwitchAR can also be applied in applications involving repeated usage by the same person.

7.2 Generalizability

We see SwitchAR as a foundational approach for enabling PMs in AR. By successfully transitioning users from AR to a fully virtual environment unnoticed, we argue that many PMs currently used in VR could be adapted to AR using SwitchAR. While our focus was on demonstrating the unnoticed switch from AR to VR, RDW served primarily as a demonstration and evaluation tool.
This capability opens new application scenarios, extending the benefits of VR-based PMs to AR environments. Beyond the established use cases of PMs in VR, the different AR context may unlock new, previously unexplored opportunities for applying these manipulations in unique ways.

7.3 Misuse

We acknowledge that the ability to covertly transition a user from AR to VR introduces potential for misuse, representing a novel security and safety vulnerability in video pass-through Extended Reality (XR) HMDs. The threats identified for VR by Tseng et al.[34], such as attackers using RDW to lead users into harmful situations (e.g., falling down stairs or interacting with dangerous objects), are equally applicable to AR through SwitchAR. Additionally, the risks may even be amplified in AR, as users believe they are still interacting with the real world and thus may be less cautious. In contrast, when using VR applications, users are aware that they are interacting with a virtual environment, which might cause them to be more careful with physical movements.
Related to this, a reason for the misuse potential of SwitchAR is that it might undermine the ”device-gap” described by Slater et al.[31]. They argue that *”the fact of putting on the HMD itself demarcates reality from virtual reality—so that unless participants are induced to somehow forget that they are wearing the HMD, they will not believe that the virtual scenario is a real one”*. This gap refers to the mental distinction users make between reality and virtuality when they put on an HMD. While SwitchAR does not make users forget they are wearing an HMD, it obscures the fact that they are no longer

perceiving the real world at certain moments, but instead a virtual reconstruction. This makes it harder for users to discern whether what they are seeing is real or virtual.

Our goal in this work was to demonstrate how concepts like change blindness and inattentional blindness can enable perceptual manipulations in AR. While SwitchAR has potential for misuse, we hope that by raising awareness of these capabilities, we can inspire the community to explore positive applications and encourage legislators and device manufacturers to implement stronger safety mechanisms to protect users and maintain their perceptual integrity.

8 Limitations and Future Work

SwitchAR operates by applying the PMs in what is perceived as AR, but is technically VR. However, we argue that this technical distinction does not affect the user experience. Since the system begins and ends in AR, with a temporary switch to VR only when PMs are needed, users do not detect these switches and thus perceive the entire interaction as AR. Their experience is indistinguishable from using a system that remains in AR throughout.

A technical limitation of our SwitchAR implementation is its reliance on pass-through AR. This is due to the need to fully occlude the real world within the user's entire FOV in order to replace the physical environment with a manipulable virtual reconstruction. Currently, this level of occlusion is only achievable with pass-through AR. However, once optical see-through AR technology progresses, giving HMDs the ability to block light from the physical environment across the whole FOV, SwitchAR can operate on optical AR HMDs using the same principles as described in this paper.

In theory, the switch between AR and VR could be executed without the need for distraction tasks or perceptual tricks, provided the virtual reconstruction is highly similar to the real environment with precise alignment. In such a scenario, the visual change during the switch would be minimal, allowing to potentially switch users from AR to VR without them noticing it, even while they are actively observing their surroundings. However, any defects in the virtual reconstruction or inaccuracies in alignment increase the need for distractions and perceptual tricks to mask the switch. While the tricks we employed in our implementation seemed to effectively conceal the switch in the user study, we do not yet know the individual contribution of each trick, the necessary intensity of their application, or whether some could have been omitted altogether. Therefore, future research could focus on quantifying the impact of these factors.

9 Conclusion

In this work, we introduced SwitchAR, a perceptual manipulation technique for video pass-through AR that leverages change blindness and inattentional blindness to imperceptibly switch between the real-world camera feed and a 3D reconstruction of the environment, enabling VR redirection techniques in AR. We outlined its core concept, which involves (1) creating a virtual reconstruction, (2) switching from AR to the virtual reconstruction, (3) applying PMs, and (4) switching back to AR. A sample implementation for RDW in AR was also presented.
We evaluated the SwitchAR implementation through a user study (n=20) that involved three rounds of switching from AR to VR followed by redirection. The feasibility of SwitchAR was confirmed by the results. None of the participants noticed the switch, and only one consistently recognized the RDW manipulation. Even after being made aware of the manipulation in the second round, participants remained unaware of the switch during the third round, demonstrating the system's robustness and repeatability. This work enables transferring both the benefits and potential risks of VR-based PMs to AR, making SwitchAR a foundational approach to enabling perceptual manipulations like redirected walking in Augmented Reality environments.

Remarks

Die vorliegende Arbeit wurde eigenständig erstellt und unter Zuhilfenahme von ChatGPT sprachlich überarbeitet.

This thesis was written independently and was linguistically revised with the assistance of ChatGPT.

Bibliography

[1] Mahdi Azmandian et al. "Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 1968–1979. ISBN: 9781450333627. DOI: 10.1145/2858036.2858226. URL: https://doi.org/10.1145/2858036.2858226.

[2] Elise Bonnail et al. "Memory Manipulations in Extended Reality". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394215. DOI: 10.1145/3544548.3580988. URL: https://doi.org/10.1145/3544548.3580988.

[3] Peter Casey, Ibrahim Baggili, and Ananya Yarramreddy. "Immersive Virtual Reality Attacks and the Human Joystick". In: *IEEE Transactions on Dependable and Secure Computing* 18.2 (2021), pp. 550–562. DOI: 10.1109/TDSC.2019.2907942.

[4] Kirsten Cater, Alan Chalmers, and Colin Dalton. "Varying rendering fidelity by exploiting human change blindness". In: *Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*. GRAPHITE '03. Melbourne, Australia: Association for Computing Machinery, 2003, pp. 39–46. ISBN: 1581135785. DOI: 10.1145/604471.604483. URL: https://doi.org/10.1145/604471.604483.

[5] Kaiming Cheng et al. "Exploring User Reactions and Mental Models Towards Perceptual Manipulation Attacks in Mixed Reality". In: *32nd USENIX Security Symposium (USENIX Security 23)*. Anaheim, CA: USENIX Association, Aug. 2023, pp. 911–928. ISBN: 978-1-939133-37-3. URL: https://www.usenix.org/conference/usenixsecurity23/presentation/cheng-kaiming.

[6] Aldrich Clarence et al. "Stacked Retargeting: Combining Redirected Walking and Hand Redirection to Expand Haptic Retargeting's Coverage". In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '24. Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 9798400703300. DOI: 10.1145/3613904.3642228. URL: https://doi.org/10.1145/3613904.3642228.

[7] Christine A. Curcio et al. "Human photoreceptor topography". In: *Journal of Comparative Neurology* 292.4 (1990), pp. 497–523. DOI: https://doi.org/10.1002/cne.902920402. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.902920402. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cne.902920402.

[8] Yifan Ding et al. "Similarity of an unexpected object to the attended and ignored objects affects noticing in a sustained inattentional blindness task". In: *Attention, Perception, & Psychophysics* 85 (Oct. 2023). DOI: 10.3758/s13414-023-02794-2.

[9] Leah Findlater et al. "Ephemeral adaptation: the use of gradual onset to improve menu selection performance". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, MA, USA: Association for Computing Machinery, 2009, pp. 1655–1664. ISBN: 9781605582467. DOI: 10.1145/1518701.1518956. URL: https://doi.org/10.1145/1518701.1518956.

[10] Matt Gottsacker et al. "Diegetic Representations for Seamless Cross-Reality Interruptions". In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2021, pp. 310–319. DOI: 10.1109/ISMAR52148.2021.00047.

[11] Akira Ishii et al. "Optical Marionette: Graphical Manipulation of Human's Walking Direction". In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. UIST '16. Tokyo, Japan: Association for Computing Machinery, 2016, pp. 705–716. ISBN: 9781450341899. DOI: 10.1145/2984511.2984545. URL: https://doi.org/10.1145/2984511.2984545.

[12] Mohamed Kari, Reinhard Schütte, and Raj Sodhi. "Scene Responsiveness for Visuotactile Illusions in Mixed Reality". In: *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. UIST '23. San Francisco, CA, USA: Association for Computing Machinery, 2023. ISBN: 9798400701320. DOI: 10.1145/3586183.3606825. URL: https://doi.org/10.1145/3586183.3606825.

[13] Kiron Lebeck, Tadayoshi Kohno, and Franziska Roesner. "How to Safely Augment Reality: Challenges and Directions". In: HotMobile '16. St. Augustine, Florida, USA: Association for Computing Machinery, 2016, pp. 45–50. ISBN: 9781450341455. DOI: 10.1145/2873587.2873595. URL: https://doi.org/10.1145/2873587.2873595.

[14] Kiron Lebeck et al. "Securing Augmented Reality Output". In: *2017 IEEE Symposium on Security and Privacy (SP)*. 2017, pp. 320–337. DOI: 10.1109/SP.2017.13.

[15] David Lindlbauer and Andy D. Wilson. "Remixed Reality: Manipulating Space and Time in Augmented Reality". In: CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–13. ISBN: 9781450356206. DOI: 10.1145/3173574.3173703. URL: https://doi.org/10.1145/3173574.3173703.

[16] Arien Mack and Irvin Rock. "Inattentional Blindness: Perception Without Attention". In: *Visual Attention* 8 (Jan. 1998).

[17] Sebastian Marwecki et al. "Mise-Unseen: Using Eye Tracking to Hide Virtual Reality Scene Changes in Plain Sight". In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. New Orleans, LA, USA: Association for Computing Machinery, 2019, pp. 777–789. ISBN: 9781450368162. DOI: 10.1145/3332165.3347919. URL: https://doi.org/10.1145/3332165.3347919.

[18] Daniel Medeiros et al. "The Benefits of Passive Haptics and Perceptual Manipulation for Extended Reality Interactions in Constrained Passenger Spaces". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023.

ISBN: 9781450394215. DOI: 10.1145/3544548.3581079. URL: https://doi.org/10.1145/3544548.3581079.

[19] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays". In: *IEICE Trans. Information Systems* vol. E77-D, no. 12 (Dec. 1994), pp. 1321–1329.

[20] Paul Milgram et al. "Augmented reality: A class of displays on the reality-virtuality continuum". In: *Telemanipulator and Telepresence Technologies* 2351 (Jan. 1994). DOI: 10.1117/12.197321.

[21] Roberto A. Montano Murillo, Sriram Subramanian, and Diego Martinez Plasencia. "Erg-O: Ergonomic Optimization of Immersive Virtual Environments". In: UIST '17. Québec City, QC, Canada: Association for Computing Machinery, 2017, pp. 759–771. ISBN: 9781450349819. DOI: 10.1145/3126594.3126605. URL: https://doi.org/10.1145/3126594.3126605.

[22] Steven B. Most et al. "How not to be Seen: The Contribution of Similarity and Selective Ignoring to Sustained Inattentional Blindness". In: *Psychological Science* 12.1 (2001). PMID: 11294235, pp. 9–17. DOI: 10.1111/1467-9280.00303. eprint: https://doi.org/10.1111/1467-9280.00303. URL: https://doi.org/10.1111/1467-9280.00303.

[23] Fabian Pointecker et al. "Bridging the Gap Across Realities: Visual Transitions Between Virtual and Augmented Reality". In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2022, pp. 827–836. DOI: 10.1109/ISMAR55827.2022.00101.

[24] Fabian Pointecker et al. "From Real to Virtual: Exploring Replica-Enhanced Environment Transitions along the Reality-Virtuality Continuum". In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. CHI '24. Honolulu, HI, USA: Association for Computing Machinery, 2024. ISBN: 9798400703300. DOI: 10.1145/3613904.3642844. URL: https://doi.org/10.1145/3613904.3642844.

[25] Sharif Razzaque. *Redirected walking*. The University of North Carolina at Chapel Hill, 2005.

[26] Franziska Roesner and Tadayoshi Kohno. "Security and Privacy for Augmented Reality: Our 10-Year Retrospective". In: 2021. URL: https://api.semanticscholar.org/CorpusID:236167221.

[27] Franziska Roesner, Tadayoshi Kohno, and David Molnar. “Security and privacy for augmented reality systems”. In: *Commun. ACM* 57.4 (Apr. 2014), pp. 88–96. ISSN: 0001-0782. DOI: 10.1145/2580723.2580730. URL: https://doi.org/10.1145/2580723.2580730.

[28] Daniel Simons and Christopher Chabris. “Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events”. In: *Perception* 28 (Feb. 1999), pp. 1059–74. DOI: 10.1068/p2952.

[29] Daniel J Simons. “Monkeying around with the Gorillas in Our Midst: Familiarity with an Inattentional-Blindness Task Does Not Improve the Detection of Unexpected Events”. In: *i-Perception* 1.1 (2010). PMID: 23397479, pp. 3–6. DOI: 10.1068/i0386. eprint: https://doi.org/10.1068/i0386. URL: https://doi.org/10.1068/i0386.

[30] Daniel J. Simons and Daniel T. Levin. “Change blindness”. In: *Trends in Cognitive Sciences* 1.7 (1997), pp. 261–267. ISSN: 1364-6613. DOI: https://doi.org/10.1016/S1364-6613(97)01080-2. URL: https://www.sciencedirect.com/science/article/pii/S1364661397010802.

[31] Mel Slater et al. “The Ethics of Realism in Virtual and Augmented Reality”. In: *Frontiers in Virtual Reality* 1 (2020). ISSN: 2673-4192. DOI: 10.3389/frvir.2020.00001. URL: https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2020.00001.

[32] Evan Suma Rosenberg et al. “Leveraging change blindness for redirection in virtual environments”. In: Apr. 2011, pp. 159–166. DOI: 10.1109/VR.2011.5759455.

[33] Lin Sun et al. *Switching realities for better task efficiency*. US Patent 10,783,711. Sept. 2020. URL: https://patents.google.com/patent/US10783711B2/en.

[34] Wen-Jie Tseng et al. “The Dark Side of Perceptual Manipulations in Virtual Reality”. In: *CHI Conference on Human Factors in Computing Systems*. CHI ’22. ACM, Apr. 2022. DOI: 10.1145/3491102.3517728. URL: http://dx.doi.org/10.1145/3491102.3517728.

[35] Andrew D. Wilson and Shane Williams. “Autopager: exploiting change blindness for gaze-assisted reading”. In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ETRA ’18. Warsaw, Poland: Association for Computing Machinery, 2018. ISBN: 9781450357067. DOI: 10.1145/3204493.3204556. URL: https://doi.org/10.1145/3204493.3204556.

[36] Katherine Wood and Daniel J. Simons. "The role of similarity in inattentional blindness: Selective enhancement, selective suppression, or both?†". In: *Visual Cognition* 25.9-10 (2017), pp. 972–980. DOI: 10.1080/13506285.2017.1365791. eprint: https://doi.org/10.1080/13506285.2017.1365791. URL: https://doi.org/10.1080/13506285.2017.1365791.

[37] Steven Yantis and John Jonides. "Abrupt visual onsets and selective attention: Evidence from visual search". In: *Journal of experimental psychology. Human perception and performance* 10 (Oct. 1984), pp. 601–21. DOI: 10.1037/0096-1523.10.5.601.

[38] Steven Yantis and John Jonides. "Abrupt Visual Onsets and Selective Attention: Voluntary Versus Automatic Allocation". In: *Journal of experimental psychology. Human perception and performance* 16 (Feb. 1990), pp. 121–34. DOI: 10.1037/0096-1523.16.1.121.

[39] Jaejun Yoo et al. *Photorealistic Style Transfer via Wavelet Transforms*. 2019. arXiv: 1903.09760 [cs.CV]. URL: https://arxiv.org/abs/1903.09760.

List of Figures

List of Tables

A Interview Responses – Asking for Observations

Table A.1: Participants' responses to being asked if they noticed anything in the first interview, with our coding for it

Participant	**Transcription of response** (first interview)	**Categorized as**
p1 (translated to English)	Not in the room, no.	Nothing relevant noticed
p2 (translated to English)	No, no	Nothing relevant noticed
p3 (translated to English)	Not really, just the time was short	Nothing relevant noticed
p4 (translated to English)	No	Nothing relevant noticed
p5 (translated to English)	No	Nothing relevant noticed
p6	No, nothing	Nothing relevant noticed
p7 (translated to English)	Nope	Nothing relevant noticed

Table A.1: Participants' responses to being asked if they noticed anything in the first interview, with our coding for it

Partici-pant	**Transcription of response** (first interview)	**Categorized as**
p8 (translated to English)	Them: Oh, because I was just here, eh, here. So I was looking for the blue sign, and then I saw that the three pieces there, the cubes, were no longer on the table, but more or less in the air. And there was also a blue person standing next to me, so I was here, and then the blue person is... [talked about how activating the paintings got easier over time, supposedly not only because of practice, but also because a change in the app]	Nothing relevant noticed (only some virtual objects like the cubes and the avatar, no hint to the switch)
p9 (translated to English)	No. But sometimes, when I stand here in front of the picture, there is this white thing as if it were growing, but if I turn my head a little bit, it becomes like this... interrupt.	Nothing relevant noticed (talked about the interaction technique for activating the paintings)
p10	Them: I don't know, if these stickers are for the experiment, or… [pointing at the aruco markers]. Interviewer: Because they look strange? Them: Ah yes, they look like some, eh, [?]. Looks a bit like QR code, maybe they [convert?] some information. Interviewer: Okay, yeah it was mainly for the tracking […]. But apart from this, nothing? Them: No	Nothing relevant noticed
p11 (translated to English)	No, not particularly. Just the pictures.	Nothing relevant noticed
p12	I just followed the arrow, I didn't notice anything	Nothing relevant noticed
p13	Ah, no	Nothing relevant noticed

Table A.1: Participants' responses to being asked if they noticed anything in the first interview, with our coding for it

Partici-pant	**Transcription of response** (first interview)	**Categorized as**
p14 (translated to English)	Apart from the picture I did not notice	Nothing relevant noticed
p15 (translated to English)	I was way too focused on that green blob. Yes, I noticed the curtain, I noticed the dice, but nothing else.	Nothing relevant noticed
p16 (translated to English)	Erm, yes, there are just film grain everywhere. It was a bit uncomfortable. And, eh, once it was as if a ghostly hand remained standing in the room. […] But otherwise…	Nothing relevant noticed
p17 (translated to English)	Eh, no	Nothing relevant noticed
p18 (translated to English)	Them: What do you mean? So abnormal then, right? Interviewer: Exactly Them: Mh, no	Nothing relevant noticed
p19 (translated to English)	No, nothing	Nothing relevant noticed
p20 (translated to English)	Nope. [?] No, no, no, I didn't notice	Nothing relevant noticed

Table A.2: Participants' responses to being asked if they noticed anything in the second interview, with our coding for it

Partici-pant	**Transcription of response** (second interview)	**Categorized as**
p1 (translated to English)	Them: Well, the HCI thing was further to the right than in the… Interviewer: So the logo has shifted? Them: Yes	Nothing relevant noticed (apart from purposely provoked offset)

Table A.2: Participants' responses to being asked if they noticed anything in the second interview, with our coding for it

Partici-pant	**Transcription of response** (second interview)	**Categorized as**
p2 (translated to English)	Them: That the hands were frozen from time to time. Interviewer: Lost the tracking? Them: Yes. Interviewer: But nothing else? Them: no	Nothing relevant noticed
p3 (translated to English)	No, nothing bothered me	Nothing relevant noticed
p4 (translated to English)	No	Nothing relevant noticed
p5 (translated to English)	Them: That's moved? Interviewer: The picture is moved? Them: The HCI Lab [logo]. It was about half a meter further back. And when I took it off [the headset], it was not where I expected it to be. Interviewer: Apart from that, did you notice anything in the room while walking around? Them: No	Nothing relevant noticed (apart from purposely provoked offset)
p6	No	Nothing relevant noticed

Table A.2: Participants' responses to being asked if they noticed anything in the second interview, with our coding for it

Partici-pant	**Transcription of response** (second interview)	**Categorized as**
p7 (translated to English)	Apparently the room was narrower, I would say, in VR. No, that doesn't fit. If I was standing there and the logo was there, then the room should... No, it was. I first went over there to go to the logo, and then it was... No, in VR it was [?]. Well, somehow, either the HCI Lab logo shifted, but it didn't look like that. It looked more like you had to cover more distance in VR than in the real world. Interviewer: So you were in VR or what? Them: Yes, that you walked in VR, um, and covered less distance in VR than you did in the real world. I don't know if that makes sense. Interviewer: So you couldn't see the room around you anymore, so there was no AR anymore? Them: Yes. The two didn't match. VR and the real world. Interviewer: Okay. Them: In VR I still had the feeling that I could walk further, even though I was already standing at the closet. Interviewer: And were you in VR the whole time, or what was it like? Them: What, what was in VR the whole time? Interviewer: So you used AR, or VR? Them: I'd say it was VR, but I also just look at pictures that... Yeah, I don't know. Could also have a simulated area in here. I'd say it was VR, not AR.	Nothing relevant noticed (apart from purposely provoked offset)
p8 (translated to English)	Not this time	Nothing relevant noticed

Table A.2: Participants' responses to being asked if they noticed anything in the second interview, with our coding for it

Partici-pant	**Transcription of response** (second interview)	**Categorized as**
p9 (translated to English)	Yes..., no, everything.... But only when I [looked for/saw] this HCI lab, then I was too close to this wall. And then I found this, this word jumped to the left. Yes, and then I went a little to the left, and then I saw this HCI lab in the middle again and then I removed it [the headset].	Nothing relevant noticed (apart from purposely provoked offset)
p10	Them: Eh, why am I in the.... I was kind of like.... I thought I was somewhere in the middle, but [it seemed so wide?] [pointing towards direction of the door and cupboards]. And I finished at here [pointing at the second painting position], I don't know why. Last time I finished at here [pointing at the first painting position]. Interviewer: Okay. Yeah so the number of pictures was different, this time? Them: I think so Interviewer: And what was the first thing you said? You were in the middle? Them: I feel like I'm more close to the door this time. And when I see HCI, I feel like, when I take off glasses, I thought it's in front of me, but actually, it's on the left side of me a bit. Interviewer: Okay. So, the position of this, eh, logo changed? Them: Yeah, I feel like.	Nothing relevant noticed (apart from purposely provoked offset)
p11 (translated to English)	Them: Yeah, well, that was slightly offset here. Interviewer: What was offset? Them: Um. Well, I don't know what it looks like, but the logo is definitely a little bit further to the right in the real world, uhm further to the left.	Nothing relevant noticed (apart from purposely provoked offset)
p12	Them: The positions are different Interviewer: What positions? Them: The logo Interviewer: The logo moved? Them: Yeah	Nothing relevant noticed (apart from purposely provoked offset)

Table A.2: Participants' responses to being asked if they noticed anything in the second interview, with our coding for it

Participant	**Transcription of response** (second interview)	**Categorized as**
p13	No	Nothing relevant noticed
p14 (translated to English)	Them: Eh, I noticed that the picture from both sides is always in the same position. Interviewer: Okay. And was that different before, or uncertain? Them: I don't remember. Do you mean in comparison to the first round? Interviewer: Yes. Them: Eh, you can't estimate. Only notice the second round, um, the two pictures are in the same position.	Nothing relevant noticed
p15 (translated to English)	No. Well, I could certainly imagine that you change the QR codes. But I wouldn't be able to recognize that. I would have to remember them.	Nothing relevant noticed
p16 (translated to English)	Um, no. Nothing else.	Nothing relevant noticed
p17 (translated to English)	No	Nothing relevant noticed
p18 (translated to English)	Mh, pretty normal, like for the first time	Nothing relevant noticed
p19 (translated to English)	No	Nothing relevant noticed
p20 (translated to English)	Nope. Should anything be noticeable?	Nothing relevant noticed

Table A.3: Participants' responses to being asked if they noticed anything in the third interview, with our coding for it

Partici-pant	**Transcription of response** (third interview)	**Categorized as**
p1 (translated to English)	Them: No. But can I add something about the one? Interviewer: One? The first run through? Them: Yes. Well, I had the feeling that with the one I had to get closer and closer to the green so that... Interviewer: And that wasn't the case with the second and third? Them: That wasn't the case with the second and third.	Nothing relevant noticed (talked about the interaction technique for activating the paintings)
p2 (translated to English)	Them: Maybe the picture frames were bigger from the [pointed at one of the walls, where the virtual paintings were shown] Interviewer: Bigger? So this one was bigger than that one? [probably pointing at the two walls] Them: Yes. I noticed that. And yes, things were still shifted, but not always, strangely enough. For example, here I realized that what I was seeing was shifted slightly to, er, to the left. [The metal pegboard he touched when wearing the HMD] When I took the glasses off here, it wasn't shifted. Although, maybe because I was paying attention to my steps, I turned 180° and then 90. I'm just not sure, orientation in space was... Then I realized that maybe orientation in space is difficult. I don't know if I'm really walking in a straight line. So I think I'm just walking backwards, but then I see things shifted, so I think maybe it's something in the glasses.	Nothing relevant noticed (for the switch)
p3 (translated to English)	No	Nothing relevant noticed
p4 (translated to English)	No	Nothing relevant noticed

Table A.3: Participants' responses to being asked if they noticed anything in the third interview, with our coding for it

Participant	**Transcription of response** (third interview)	**Categorized as**
p5 (translated to English)	No	Nothing relevant noticed
p6	No. It's always the same	Nothing relevant noticed
p7 (translated to English)	Them: Not really. Interviewer: But? Them: I just had a guess as to whether you walk in a square. But that... can't happen here... I wouldn't say here. That's too vague for me.	Nothing relevant noticed
p8 (translated to English)	Them: So I saw the curtains, that it is not a line, but a bit like this, um, round. [Drew something to explain the different distances between the curtain and the two painting positions] Interviewer: Okay, so the paintings are not at the same distance from the curtain? Them: Yes, yes, yes. And this time the, [...] my body appears again, the blue body again... Interviewer: Ah, okay, it froze again... Them: Yes. And um, sometimes I see my hand, well, I always held it here [on the sides of the headset], and then I see my hand, a bit behind. Well, actually not so behind, I held my hand here, but I saw my hand on both sides here	Nothing relevant noticed
p9 (translated to English)	[Talked about the paintings contents] Interviewer: And outside of the paintings, did you notice anything about the room? Them: Outside of the paintings? Interviewer: Exactly, here in the room, was there anything unusual? Them: Mh, no, but when I stand here, I see the, eh, the hand of a person. Interviewer: The blue one? Them: Blue, yes.	Nothing relevant noticed

Table A.3: Participants' responses to being asked if they noticed anything in the third interview, with our coding for it

Partici-pant	**Transcription of response** (third interview)	**Categorized as**
p10	Them: I think, I saw the, eh, hand track again. Interviewer: What do you mean, you saw it again? Them: No no, I saw the cubes again. Interviewer: Okay. And they haven't been there before, or? Them: Second round, I didn't notice. [...]. I didn't notice [something different?].	Nothing relevant noticed
p11 (translated to English)	Them: Um, I had the feeling that the pictures were, um, more offset this time. Interviewer: Offset in which direction?Them: It was as if one picture was hanging here and the other a bit further back. Interviewer: Okay, the one over there was always further to the right and that was [to the?] left? Them: Exactly. Yes, exactly. Interviewer: So the position of the two pictures opposite was shifted? Okay. Them: Yes. Interviewer: Okay, but did you notice anything else? Them: No.	Nothing relevant noticed (for the switch)
p12	Them: Now, this time, the logos are in the same positions Interviewer: Okay, so no offset this time? Them: [right?]	Nothing relevant noticed
p13	No	Nothing relevant noticed
p14 (translated to English)	Them: Eh, there is no more logo deviation. Interviewer: Okay. But nothing else? Them: Yes, a lot. First of all, there are actually eleven paintings. [started talking about the paintings' contents]	Nothing relevant noticed
p15 (translated to English)	I don't think so. You have to tell me later whether there was anything. I didn't find anything. I even looked around several times	Nothing relevant noticed
p16 (translated to English)	Them: Eh, yes. This time loading the images didn't work so well. Interviewer: [...] But apart from that? Them: Eh, no.	Nothing relevant noticed

Table A.3: Participants' responses to being asked if they noticed anything in the third interview, with our coding for it

Partici-pant	**Transcription of response** (third interview)	**Categorized as**
p17 (translated to English)	Them: Eh, the cubes were a bit further away at the end [Showing a leftward movement with her hand] Interviewer: Anything else? Noticed anything? Them: No	Nothing relevant noticed
p18 (translated to English)	Them: [?] it is exactly in the middle? Interviewer: This time it has not moved? Them: No	Nothing relevant noticed
p19 (translated to English)	[Shaking head]	Nothing relevant noticed
p20 (translated to English)	Them: Hmm, I don't know, I had the feeling that I wasn't standing right in the middle of the pictures, but a little bit to the side. So that the loading bar starts, so to speak. But I'm not sure if it was the same for the others. But I just noticed that I had to stand a little bit further to the side. Interviewer: Otherwise, uhm, but nothing? Them: Nothing else, no.	Nothing relevant noticed (talked about the interaction technique for activating the paintings)

B Interview Responses – Final Four Questions

Table B.1: Participants' responses to the first of the final four questions, "Did you notice that the environment changed?", with our coding for it

Participant	**Transcription of response** (final question 1/4)	**Categorized as**
p1 (translated to English)	No.	Nothing relevant noticed.
p2 (translated to English)	Hm, no, not really. So the pictures on [one] wall are bigger.	Nothing relevant noticed.
p3 (translated to English)	Not particularly.	Nothing relevant noticed.
p4 (translated to English)	No.	Nothing relevant noticed.
p5 (translated to English)	Them: No. Interviewer: Okay, and before that? Them: Well, I've already looked through it a few times, and I didn't notice anything, actually. Maybe the color of the curtain, but I didn't pay attention to that either.	Nothing relevant noticed.

Table B.1: Participants' responses to the first of the final four questions, "Did you notice that the environment changed?", with our coding for it

Partici-pant	**Transcription of response** (final question 1/4)	**Categorized as**
p6	Them: I didn't pay attention to that. Yeah, I think in the first and second I always noticed that three boxes, but the third time I didn't notice. I don't know, if it's still there or no more. Interviewer: But you don't know? Them: Yeah I don't know, I didn't pay attention.	Nothing relevant noticed.
p7 (translated to English)	I tried to pay attention to things like that, like the fire alarm symbol over there, but it was there. It's possible that the sticker wasn't there, but I could be wrong. I... I noticed the fire alarm, the sign there, in particular.	Nothing relevant noticed.
p8 (translated to English)	Them: Yes, so as I said before, even though I take it off right here in front of the..., when I took it off, in front of the [logo], and so [I assume?] the virtual [?] is shifted. So it's not as strong as before. Interviewer: So it depends on the distance from this logo, right...? Them: Oh, I don't know exactly. But yes, but the position is of course a little bit, a little bit, I think it could have an influence, yes.	Nothing relevant noticed.
p9 (translated to English)	Them: Hm, yes. Interviewer: In what way? What has changed? Them: Excuse me? Interviewer: Did the space around you change when you were…? Them: Changed? Interviewer: Yes. Them: Do you mean the space, the paintings, or…? Interviewer: No, the real space around you here. Them: Maybe in the end, eh, the distance between the paintings and this HCI-Lab a bit closer. I walked less, I think. Interviewer: Oh, when you walked there at the end? Them: No, I mean from here to here [pointing towards one painting position and then the HCI-Lab logo]	Nothing relevant noticed.

Table B.1: Participants' responses to the first of the final four questions, "Did you notice that the environment changed?", with our coding for it

Participant	**Transcription of response** (final question 1/4)	**Categorized as**
p10	Oh, I didn't notice at all, I was focused on the paintings.	Nothing relevant noticed.
p11 (translated to English)	No.	Nothing relevant noticed.
p12	Them: Ah, I found a strange thing. Eh, in the last round, when I look from here to there [pointing roughly at the two painting positions], that there is a virtual hand. It's horrible. Interviewer: Oh, so it scared you? Them: I saw the hand two times, or three times. Interviewer: [explained the frozen hand tracking] But apart from this, nothing? Them: No. Because I try my best to remember things, I…	Nothing relevant noticed.
p13	No.	Nothing relevant noticed.
p14 (translated to English)	Them: Eh, I just gave my attention to the picture. Maybe now and then with this cupboard, as a memory to help with positioning. Otherwise I didn't notice any others. Interviewer: But nothing has changed on the cupboard either? Them: No	Nothing relevant noticed.
p15 (translated to English)	Except that the logo was offset, no. Um, I'm not sure anymore if at some point, I suddenly thought in the middle of the study: wasn't the HCI logo red earlier? And then I wasn't sure anymore, and since then I've been stuck. Since then I've been stuck on this: was it red at some point. It could also be that I simply noticed the fire extinguisher on the door and mapped the red onto the logo. But I could have sworn it was red at some point and was like: why is it blue?	Nothing relevant noticed.

Table B.1: Participants' responses to the first of the final four questions, "Did you notice that the environment changed?", with our coding for it

Partici-pant	**Transcription of response** (final question 1/4)	**Categorized as**
p16 (translated to English)	Hmm, not really. I didn't really pay much attention to the surroundings... It was always roughly the same [?]. But if there had been a different pattern in the pictures, I wouldn't have noticed.	Nothing relevant noticed.
p17 (translated to English)	Them: Erm, slightly in the last round, but otherwise... no. Interviewer: What changed in the last round? Them: So, just the cubes, and [nothing else?]. Interviewer: But nothing else? Them: I didn't notice any change.	Nothing relevant noticed.
p18 (translated to English)	Oh, I didn't notice that. I'm concentrating on remembering the pictures. Has there been any change?	Nothing relevant noticed.
p19 (translated to English)	Mhmh [shakes head]	Nothing relevant noticed.
p20 (translated to English)	No.	Nothing relevant noticed.

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Partici-pant	**Transcription of response** (final question 2/4)	**Categorized as**
p1 (translated to English)	Yes.	Always in AR.
p2 (translated to English)	Hmm, yeah? What's the question? I was in augmented reality the whole time. Or what do you mean?	Always in AR.

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Partici-pant	**Transcription of response** (final question 2/4)	**Categorized as**
p3 (translated to English)	Them: I guess so. I saw the arrows, all the paintings that suddenly opened up. Interviewer: And your environment too? Them: Yes, exactly. The floor, the walls, the lamps... I think as well.	Always in AR.
p4 (translated to English)	Yes.	Always in AR.
p5 (translated to English)	I think.	Always in AR.
p6	Them: I think it is..., the whole round I think it is. Because I can always see that..., [arrows]. Interviewer: Okay. But you could also see the real world all the time as well, right? Them: I can always see those symbols [pointing at the aruco markers] Interviewer: So it was.... All the time you had like real world and virtual stuff on top of each other? Them: Yeah, yeah.	Always in AR.
p7 (translated to English)	Especially not in the second scene.	Not always in AR (deduced in hindsight)
p8 (translated to English)	Them: So here is the curtain, behind the curtain not, but the other side yes. Interviewer: Have you always seen? Them: Yes.	Always in AR.
p9 (translated to English)	Them: See the real world? Yes, I can. Interviewer: All the time? Them: Yes, all the time, I can.	Always in AR.

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Partici-pant	**Transcription of response** (final question 2/4)	**Categorized as**
p10	Them: Eh, no, I don't think it's always AR. I think it's AR, but I think the curtains, I think it's a bit VR. Interviewer: Okay, but apart from what the curtain is hiding, you could see the real world all the time? Them: Yeah, yeah yeah yeah. Yes, I think so, yes.	Always in AR.
p11 (translated to English)	Yes.	Always in AR.
p12	[They talked about only the second round having been VR earlier] Interviewer: So in the first and last round, you've been in Augmented Reality all the time, right? You could see the real world around you all the time? Them: I think so. Interviewer: And the second round, you couldn't, it was virtual, right? That was your explanation, right? Them: Yeah. Interviewer: And, like was it virtual for the whole round, or did it change? Them: I think, in the beginning it is real world, but when I focus myself on remember things, it changed. And I didn't notice the change, but I, I [?] last, I think. Because the change of the logo. Interviewer: You found it what, sorry? Them: So, I said I think, uhm, in the beginning, it is real world, but when I focus on the other things, it changed, but I didn't notice it, until I found the positions and [/of?] the logos are changed. Interviewer: Okay, that's the first moment you noticed it, basically? When you saw… Them: Not the first time, but in the process I, I think, I think it's not that real, but it just, it just a feeling, I, yeah I….	Not always in AR (deduced in hindsight)

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Partici-pant	**Transcription of response** (final question 2/4)	**Categorized as**
p13	Them: Oh, when I wear this? [pointing at the headset] Yeah. Interviewer: Okay, so you could see the real room all the time? Them: Yeah, yeah, yeah. Just a bit different resolution. Eh, not reso..., I mean the, like the.... If I use my eyes to see, it's very clear. But I used the camera, a little, eh, not very totally clear.	Always in AR.
p14 (translated to English)	Them: Eh, I think so. I didn't notice anything. Interviewer: Didn't notice what? Them: Eh, that is, when I walk around, I only think about the order of the image or the content, or I only notice the image. And then I don't see, or I don't notice much of the surroundings. Interviewer: Okay. But basically you could always see the room, the real room here? Them: I can always see the room. So a small detail, maybe not that important. I saw my, my hand. [Interviewer explains the hand tracking getting stuck causing some frozen hands] Them: When I want to look down, I see my foot. Interviewer: Okay. Er, the, the foot was normal, or not? Them: Blue. Is it like this, this side is blue, this side is grey.	Always in AR.

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Partici-pant	**Transcription of response** (final question 2/4)	**Categorized as**
p15 (translated to English)	Them: Eh, I think so. Interviewer: So you could see the real world around you the whole time? Them: I'm not sure about round two. Because if I had seen the whole real world, I would have seen the logo in the right place. And I didn't. So I'm not sure about round two. By round three, I'm pretty sure that I saw the real world around me. Especially because I did interact with the real world briefly by touching this thing [they touched the perfboard to check if it was really there in the real world]. Yes, as I said, by round two, I'm not really sure because the logo was moved and... yes. Interviewer: But in the other two rounds, yes? Them: Yes. Well, I don't know about round one, I didn't really pay much attention to it. I just don't know for sure. By round three, I'm pretty sure that it was the real world. With the first one I was still overwhelmed by all the QR codes and stuff. And the glasses didn't work. [problems with activating the paintings]	Not always in AR (deduced in hindsight)
p16 (translated to English)	At least I had the feeling that I could still see this room.	Always in AR.
p17 (translated to English)	Mmm, yes.	Always in AR.
p18 (translated to English)	Mmm, yes.	Always in AR.
p19 (translated to English)	Yes.	Always in AR.

Table B.2: Participants' responses to the second of the final four questions, "Have you been using AR for the whole duration of the study?", with our coding for it

Participant	**Transcription of response** (final question 2/4)	**Categorized as**
p20 (translated to English)	Yes.	Always in AR.

Table B.3: Participants' responses to the third of the final four questions, "Did you notice a point in time when you changed from AR to VR?", with our coding for it

Participant	**Transcription of response** (final question 3/4)	**Categorized as**
p1 (translated to English)	No, I was too busy memorizing the pictures.	Nothing relevant noticed.
p2 (translated to English)	Hmm, no.	Nothing relevant noticed.
p3 (translated to English)	No.	Nothing relevant noticed.
p4 (translated to English)	No.	Nothing relevant noticed.
p5 (translated to English)	Them: So maybe when it moved that thing [the logo], but… Interviewer: So basically the moment you took it off [the headset]? Them: So it could be that it was the second time, VR not AR. Interviewer: Okay, and then from the beginning when you put the headset on? So did you notice a point in time when it changed, from AR to VR? Them: [head-shaking and a negating "Mhm"]	Nothing relevant noticed.
p6	No, I never thought that. Did it?	Nothing relevant noticed.

Table B.3: Participants' responses to the third of the final four questions, "Did you notice a point in time when you changed from AR to VR?", with our coding for it

Partici-pant	**Transcription of response** (final question 3/4)	**Categorized as**
p7 (translated to English)	Them: Mhm [negative]. Interviewer: So you would say you were in VR from the start? Them: That's difficult. Um. Because we had a scene change... Interviewer: In what way? Them: Exactly, when I took the glasses off and you checked again to see if everything was still OK, it could of course have been that you switched completely. But I don't think you switched from VR to AR within one run. That would be crazy. Interviewer: Okay, but you also tried hand tracking, looked at your hands and so on. They were there too. Them: Yes, I don't think I really did it in the second scene. I just kept walking around and didn't look at my hands. And in the third scene I looked at my hands again. Um. But I didn't notice anything then either. Interviewer: So the hands were there? Them: They were there in the third scene, but I didn't notice them in the second, so I don't know. Interviewer: Okay, so it was probably VR in the second one and the other two in AR? Them: You can see your body in an AR passthrough environment, so.	Nothing relevant noticed.
p8 (translated to English)	No, no, no.	Nothing relevant noticed.
p9 (translated to English)	Them: The last time is virtual realization? Interviewer: I ask whether it changed to the virtual at some point. Them: No, I didn't notice.	Nothing relevant noticed.
p10	Oh, I didn't notice. Switch from AR to VR? No, I didn't notice.	Nothing relevant noticed.

Table B.3: Participants' responses to the third of the final four questions, "Did you notice a point in time when you changed from AR to VR?", with our coding for it

Participant	**Transcription of response** (final question 3/4)	**Categorized as**
p11 (translated to English)	No.	Nothing relevant noticed.
p12	Them: No. Interviewer: So it was more just a feeling afterwards? [the explanation for the time of change in the previous question] Them: Yeah.	Deduced in hindsight
p13	No, I don't notice this.	Nothing relevant noticed.
p14 (translated to English)	I didn't notice. Eh, what I notice is that there is a time, a time window, eh, the second runs out.	Nothing relevant noticed.
p15 (translated to English)	No.	Nothing relevant noticed.
p16 (translated to English)	Them: Hm, no. Except maybe, eh, in the second environment, where I was offset. But otherwise I didn't notice it directly. Interviewer: Okay. In the second run, then, because there was an offset...? Them: Yes. And because I now imagine that there could have been less grain, but...	Nothing relevant noticed.
p17 (translated to English)	Them: Mh, probably when I looked at the pictures. And then focused on... the picture frames. Interviewer: Okay, but did you notice that at some point, or is that the most likely explanation in retrospect? Them: No, I noticed. That I was more focused on it [?] and then no longer perceived the surroundings. So [?] other pictures. And focused more on the virtual. Interviewer: Okay, but did you then notice that the rest, that the rest was switched, and that the room was basically...? Them: no, no, no.	Deduced in hindsight

Table B.3: Participants' responses to the third of the final four questions, "Did you notice a point in time when you changed from AR to VR?", with our coding for it

Partici-pant	**Transcription of response** (final question 3/4)	**Categorized as**
p18 (translated to English)	Mh, no, I never... I didn't notice.	Nothing relevant noticed.
p19 (translated to English)	No.	Nothing relevant noticed.
p20 (translated to English)	No.	Nothing relevant noticed.

Table B.4: Participants' responses to the fourth of the final four questions, "Did you notice that there was a virtual reconstruction of the room?", with our coding for it

Partici-pant	**Transcription of response** (final question 4/4)	**Categorized as**
p1 (translated to English)	No.	Nothing relevant noticed.
p2 (translated to English)	Them: Hmm, no, I didn't realize that at the time. Interviewer: Do you mean while you were using it? Them: Yes. Now I was just thinking about what it could be. And then this question, um, compared, okay, it could be something different.	Nothing relevant noticed.
p3 (translated to English)	Nope. I thought that was the room the whole time.	Nothing relevant noticed.
p4 (translated to English)	No.	Nothing relevant noticed.

Table B.4: Participants' responses to the fourth of the final four questions, "Did you notice that there was a virtual reconstruction of the room?", with our coding for it

Partici-pant	**Transcription of response** (final question 4/4)	**Categorized as**
p5 (translated to English)	Them: That was my theory about the shift. But I thought it wouldn't look that realistic. Interviewer: And that's why you thought it would still be AR, all the time? Them: Yes	Had the theory because of the offset
p6	Them: The setup of this room is absolutely changed, I think. Interviewer: The setup has changed? Them: yeah, um.... [looked up a word] Curtain. Yeah, the curtain... Interviewer: Yeah, what about it? Them: It's here, always, the curtain. Interviewer: Yes, but was it there all the time? Them: Yes. Interviewer: Okay, but apart from that you didn't notice any virtual model? Them: Um... [looking around] I don't know, I don't pay attention. I was always focusing on that... [pointing at the virtual painting's position]. And also that countdown. Yeah, I always watched that.	Nothing relevant noticed.
p7 (translated to English)	Them: Not actively. Well, later on, yes. So I was standing somewhere else than I thought. It must be a reconstruction. Interviewer: OK, but didn't you see it? Them: No, I actually didn't see it. I just thought that sometimes the UI elements of this bar, the white bar and the green thing were a bit pixelated, but it was everywhere.	Had the theory because of the offset

Table B.4: Participants' responses to the fourth of the final four questions, ”Did you notice that there was a virtual reconstruction of the room?”, with our coding for it

Participant	**Transcription of response** (final question 4/4)	**Categorized as**
p8 (translated to English)	Them: Well, I saw it [pointing at the whiteboard], because they weren't covered by the curtains. Well, roughly like this, and from here, but I can see it's not that big. Actually, well, actually only about this big, so my view isn't that big. So, when I see it in the real world, I'm taller than when I see it through the glasses, yes. I didn't notice that. Interviewer: Did you only notice that afterwards? Them: Yes, yes, yes, [more like this?]. Because in the virtual world I noticed that it's not as high. It's like a frame, roughly the same height.	Noticed something
p9 (translated to English)	Them: If you didn't ask, then no. But if you did ask, maybe, as I recall, sometimes yes. Interviewer: Okay. Can you remember any moment when you noticed that? Them: Um, maybe at the beginning? And then I didn't notice anything. Interviewer: Okay. At the beginning, did you notice anything in particular, or...? Them: At the beginning, when I look there, then maybe it looks like, not like the real, real world. Interviewer: When you picked up the cubes? Them: Yes, yes. Interviewer: Okay. And why was it maybe not like the real world? Them: Eh, I'm not sure. Maybe the cubes and stuff.... Maybe the effect of those cubes, and then I think they're not like the real world. Because the cubes are so... Interviewer: They're virtual? Them: Yes, mhmh.	Nothing relevant noticed.

Table B.4: Participants' responses to the fourth of the final four questions, "Did you notice that there was a virtual reconstruction of the room?", with our coding for it

Partici-pant	**Transcription of response** (final question 4/4)	**Categorized as**
p10	[Oy my god?] No, I didn't notice. Because I was thinking the, this, I mean, it was obviously AR, because I thought, there are posters on the wall. And the green, the green sign, which means that's, ehm, AR. So I was thinking, it's AR. Pretty sure I was in AR. I didn't notice some time to change to VR. Interviewer: Okay, so you thought you could see the real world all the time? Them: Yeah.	Nothing relevant noticed.
p11 (translated to English)	No.	Nothing relevant noticed.
p12	Them: In the end of the round two. Interviewer: Because the offset…? Them: Yeah	Had the theory because of the offset
p13	Them: Uhm, yeah. [pointing towards where the virtual cubes were] Interviewer: Okay, there were cubes, yes. The cubes were virtual. But did you notice that the whole room was virtual? Them: Uhm, no. You mean the whole room is virtual? Interviewer: Did you notice it at any point? Them: Uhm, I think I just focused my…. Maybe the cubes, I think this one is, is virtual. But around is, is real one.	Nothing relevant noticed.

Table B.4: Participants’ responses to the fourth of the final four questions, ”Did you notice that there was a virtual reconstruction of the room?”, with our coding for it

Partici-pant	**Transcription of response** (final question 4/4)	**Categorized as**
p14 (translated to English)	Them: Eh, I think I’m not sure of the terms. Eh, I think when I put the glasses on, of course it’s virtual... abstract. Interviewer: Exactly, but through cameras, first of all. Ehm, well, you can see the room through the cameras. And, exactly, did you notice that there was a virtual model that worked without the cameras? So, like, those cubes were completely virtual, right? Them: Yes. Interviewer: Yes. And did you notice, there was a completely virtual model of this room, that was used, in the headset? Them: Eh, I didn’t notice. I just got an idea. Maybe the question when you asked, eh, if the logo, eh, displacement. Maybe the question depends on whether you use camera. Interviewer: Okay. What does that mean? [Had a discussion about what AR and VR actually mean] Them: Maybe with the camera off, you can see the VR room and image. Interviewer: Okay. But did you notice it at some point? Them: No, I didn’t notice at all. Eh, I didn’t notice, but I just learned with this question, I just realized, maybe, maybe you get offset because you depict VR, VR space. Maybe that’s where displacement comes from. Or vice versa.	Nothing relevant noticed.
p15 (translated to English)	The logo was offset. I didn’t notice it consciously, but it’s a deduction from the fact that it was in a different place, and so logically it must have been VR. Otherwise I don’t think I would have noticed. Especially because I was focusing on the pictures.	Had the theory because of the offset

Table B.4: Participants' responses to the fourth of the final four questions, "Did you notice that there was a virtual reconstruction of the room?", with our coding for it

Participant	**Transcription of response** (final question 4/4)	**Categorized as**
p16 (translated to English)	Hm, no. Except for the curtain that was in the room...	Nothing relevant noticed.
p17 (translated to English)	Them: Hm, no. Well, maybe [a few things?] were sharper, but otherwise... Interviewer: A few things were sharper? Them: Yes, well, um, I noticed that the, eh, instructions looked a bit sharper, or slightly different, than in [reality?]. Interviewer: That, the note on the door, right? Them: Yes, exactly. Eh, [that one there?], exactly. Interviewer: When did you notice that? Them: In the second round, very briefly [?] Interviewer: Okay, but not otherwise? Them: No.	Noticed something
p18 (translated to English)	Them: Eh, yes, so there are a bit of the curtains here and so on... Interviewer: Okay, but of the room itself, so to speak? The walls, and the whole room, that there was a virtual model of it? Them: Mh. There are a few moments, because I can actually see my hands when I touch the glasses. And then there is a delay, that my hands [show themselves or are actually?] already here, and then apparently these virtual hands are still in front of them, right. Yes, and otherwise I don't really notice, otherwise I don't really notice, yes.	Nothing relevant noticed.
p19 (translated to English)	Them: I thought so, but I didn't notice it. Interviewer: Why did you think so? Them: Because otherwise the offset can't really happen.	Had the theory because of the offset
p20 (translated to English)	No.	Nothing relevant noticed.

Group	**Level of understanding/what was mentioned**	**Assigned Participants**
Group 1	Nothing mentioned	P1, P2, P3, P4, P6, P9, P10, P11, P13, P14, P16, P18, P20
Group 2	Virtual reconstruction	P5, P8, P19
Group 3	Not always AR	P7, P15
Group 4	Moment of switch	P12, P17

Table B.5: Grouping of the participants based on their responses to the final four questions