Human Factors at Play: Understanding the Impact of Conditioning on Presence and Reaction Time in Mixed Reality

Yasra Chandio, Victoria Interrante, and Fatima M. Anwar

Fig. 1: *V*iew of the virtual scene in MR. In the conditioning stage view (left), there are two interactable virtual boxes (white boxes placed at the edge of the table), and a physical box is placed on the table (brown box at the back). With the grab pose in positive conditioning, hands do not go through the box (center-left); the grab pose in negative conditioning shows two see-through boxes improperly stacking and a hand going through the box, with the physical box and its sticker at the back visible through the box (right-left). In the trial stage (right) view, there is a 9-button 3×3 grid with a round-shaped push button having a blue button and grey rim (left-right). Buttons are then highlighted one at a time in red with all the white glow buttons and rims (center-right). A calming pattern in the break between rounds is shown through rims turning black one at a time (right-right).

Abstract—A prerequisite to improving the *presence* of a user in mixed reality (MR) is the ability to measure and quantify *presence*. Traditionally, subjective questionnaires have been used to assess the level of presence. However, recent studies have shown that presence is correlated with objective and systemic human performance measures such as reaction time. These studies analyze the correlation between presence and reaction time when technical factors such as object realism and plausibility of the object's behavior change. However, additional psychological and physiological human factors can also impact presence. It is unclear if presence can be mapped to and correlated with reaction time when human factors such as conditioning are involved.

To answer this question, we conducted an exploratory study ($N = 60$) where the relationship between presence and reaction time was assessed under three different conditioning scenarios: control, positive, and negative. We demonstrated that human factors impact presence. We found that presence scores and reaction times are significantly correlated (correlation coefficient of -0.64), suggesting that the impact of human factors on reaction time correlates with its effect on presence. In demonstrating that, our study takes another important step toward using objective and systemic measures like reaction time as a presence measure.

Index Terms—Radiosity, global illumination, constant time

1 Introduction

As Mixed Reality (MR) applications steadily permeate domains ranging from entertainment to education [50], understanding the user experience in this digital interface is important. Central to this experience is enabling and measuring the sense of presence, a user's immersive feeling of "being" within a virtual environment (VE). Traditionally, subjective questionnaires have been used to assess the level of presence. However, a recent study has shown that presence correlates with objective and systemic human performance measures such as reaction time [15]. This research examines the relationship between presence and reaction time, considering variations in technical factors such as object realism and plausibility of the object's behavior. While the technological aspects of MR influence the sense of presence [64], psychological [43] and physiological factors [48] also wield considerable influence [45, 57]. Therefore, any approach to measuring presence should be sensitive to these human factors.

An important human factor that influences presence is *conditioning* [16, 31, 80], which, in this context, delves into how a user's prior experiences or exposures (immediate or in the past) dictate their subsequent interactions and expectations in MR [45, 90]. A subtle yet impactful priming can be paired with more direct interventions such as training, which methodically imparts specific skills or knowledge to users. Our emphasis on conditioning is due to its unique characteristics compared to other human factors such as emotions [35], agency, and lifelong habitude [19, 26, 55], which are susceptible to external influences, lack intervention points, or are less adaptable over short time scales [55]. Also, conditioning allows more active influence over user behavior [31] and is adaptable through consistent exposure [11]. The additional factors, such as users' exposure to related technologies such as Virtual Reality (VR), Augmented Reality (AR), and gaming, can also be impactful.

Recognizing the multifaceted nature of user engagement, we conducted a comprehensive between-subjects study with 62 participants. These participants were evenly divided among three distinct conditioning scenarios: positive, negative, and control, with each group consisting of 20 individuals (2 participants were excluded from the analysis due to technical difficulties). Participants had diverse exposures to VR, AR/MR, and gaming within these groups, ranging from novices to experienced users.

The premise of our study is straightforward: while negative conditioning might lead to heightened presence and faster reaction times due to increased immersion, positive conditioning could curtail this sensation of presence, resulting in more sluggish response times. The control group, meanwhile, would serve as a comparative baseline. However, overlaying this was the variable of technological familiarity; would seasoned AR/VR and gaming enthusiasts react differently compared to their less-experienced counterparts?

- *Yasra Chandio is with University of Massachusetts Amherst. E-mail: ychandio@umass.edu*
- *Victoria Interrante is with University of Minnesota Twin Cities. E-mail: interran@umn.edu.*
- *Fatima M. Anwar is with University of Massachusetts Amherst Research. E-mail: fanwar@umass.edu.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

Our study navigates the intricate interplay between conditioning, presence, and response time in MR to bring forth insights into the effect of human factors on the presence and its correlation with reaction time. By investigating the effect of human factors on the correlation between presence and reaction time, we enable a richer, more nuanced understanding of objective performance measures in MR. As the MR landscape evolves, such insights can guide developers and researchers in crafting experiences that resonate with users, ensuring immersive and intuitive interactions.

To understand the relationship between *conditioning*, *presence*, and *reaction time*, this work makes the following contributions.

Contribution 1: We explore the relationship between human factors and presence. To do so, we design experiments that observe users' sense of presence using presence questionnaires and a prompt question in MR under various conditioning scenarios.

Contribution 2: We conduct an exploratory lab study (N = 60) to demonstrate that conditioning can alter presence but has the opposite effect, i.e., positive conditioning decreases presence, and the effect of conditioning is enhanced under prior exposure to gaming.

Contribution 3: We demonstrate that a change in presence correlates with a change in reaction time when human factors vary. We also show that presence and reaction time exhibit a negative correlation.

2 Background and Related Work

This section discusses prior work on measuring presence, including subjective questionnaires and human performance measures, and the relationship between *conditioning*, *presence*, and *reaction time*.

2.1 Measuring Presence in MR

Our definition of MR leans towards AR on Milgram's reality-virtuality spectrum [51], blending virtual objects with the real world. MR lets users engage with both realities and for simplicity, we will consistently refer to this as MR in our paper. Prior research highlights that the essence of the virtual experience hinges on the sense of *presence* — that feeling of "being there" within a VE. It is not just about high-fidelity graphics; the interaction should also mirror real-world engagements. Accurately measuring *presence* in any VE (AR/VR/MR) presents unique challenges. Traditional methods of measuring this immersion have relied on post-experience questionnaires [22, 29, 58, 63, 65, 66, 75, 94], capturing users' retrospective feelings about the VE to assess elements like sensory fidelity and emotional engagement. While they are easy to administer, their reliance on memory and the absence of dynamic real-time insights hinder comprehensive assessments [67].

Additionally, traditional tools for measuring presence were defined primarily for the fully immersive virtual environment (VR) [82]. While there have been attempts to tailor these tools specifically for AR/MR [25,53,59,81,82], they often measure auxiliary factors rather than the intrinsic sensation of *presence* itself [88]. However, prior studies suggest that the questionnaires developed for VR can still be useful when all users experience the same type of environment even if the environment is not fully immersive, such as AR/MR [86, 91]. Presence questionnaires are used to explore the subjective experience of presence rather than the link between perceived presence and aspects of technology; therefore, they can be employed anywhere on the virtuality continuum. Recent work by Chandio et al. establishing the relationship between presence and reaction time [15], when technical factors are manipulated, also strongly justifies using VR questionnaires in MR.

2.2 Presence and Reaction time

Researchers have explored objective measures for presence, recognizing the shortcomings of subjective measures [63, 67, 69]. In recent years, researchers have analyzed behavior-based metrics [36,76] such as examining automatic responses, including facial expressions [76], posture [23], and startle reflexes [93], among others. Prior work has explored leveraging physiological responses such as heart rate changes [28, 44] and skin conductance [48]. However, they can be inconsistent and sometimes unrelated to presence levels [48].

As the search for a reliable and objective measure continues [61], a human performance-based measure of *reaction time*[1] emerges as a promising candidate [5, 7, 8, 33, 47, 74, 75]. The reaction time captures the immediate response of a user while highlighting its potential as a real-time indicator of presence. Recent work by Chandio et al. [15] has demonstrated a correlation between presence and reaction time in MR; as the presence degraded, the reaction time went up and vice versa. Their experiment with two major technical factors related to the presence (*plausibility* and *realism* [73, 78]) also suggests that as users feel more immersed, their reactions become more swift. Thus, by combining reaction time with conventional measures, we can gain a richer, more holistic understanding of a user's *presence*, pushing the boundaries of what MR can achieve.

2.3 Human Factors and Presence

A combination of technical and human factors influence the presence in MR. Prior work has explored how presence is impacted by technical factors such as visual realism [68, 71], field of view (FoV) [4, 12, 18], and level of details [95]. However, there is limited research that directly analyzes how presence is affected by human factors such as agency [39], emotion [19, 85], individual expectations [46, 89], genetic predispositions [13, 24], life-long habitude perceptions [19, 26, 55], and conditioning [45, 90]. Furthermore, only a few studies explore the roles of technical and human factors on presence in combination [38, 39]. However, the scope of prior work was limited to VR and leveraged only subjective questionnaires as a measure of presence.

2.4 Conditioning, Presence and Reaction Time

There is a gap in understanding how technical and human factors might interplay to cultivate presence and how this interrelation impacts the relationship between presence and reaction time. Our research is pioneering in its approach to delve into human factors systematically. Our investigation aims to discern the relationship between presence and response time, especially when viewed with human factors. We focus on *conditioning*, a psychological process where a response to a stimulus becomes more predictable or probable through reinforcement [16].

Our emphasis on conditioning is due to its unique characteristics compared to other human factors. While emotions [19], agency [35], individual expectations [46], genetic predispositions [13, 24], and life-long habitude perceptions [55] are significant, they differ fundamentally from *conditioning* in the context of presence [27, 57, 91]. For example, emotions are inherently transient and susceptible to numerous influences, posing challenges in consistent measurement [55]. In contrast, conditioning offers a systematic approach that can help identify response patterns to stimuli over time. Agency, another intrinsic factor, pertains to a user's perceived control in a VE, which is heavily influenced by scene design. Genetic predispositions provide insight into innate user tendencies but lack intervention points, whereas conditioning allows active behavioral influence [16, 27]. Life-long habitude perceptions, deeply ingrained over time, also influence virtual interactions. However, unlike the more dynamic nature of conditioning, these perceptions are less adaptable in short-term studies [26].

Conditioning, however, directly impacts how users interact with VEs, allowing more active influence over user behavior [31]. Conditioning can also adapt and potentially alter these perceptions through consistent exposure [11]. To summarize, while various factors contribute to the virtual experience, *conditioning* emerges as a pivotal research focus due to its adaptability and direct influence potential.

3 Approach

This section overviews our approach, outlines our hypothesis, and details our experiment design choices,

3.1 Overview

The intricate nature of human perception and response in MR can be significantly shaped by *conditioning* [80]. It is deeply rooted in psychological frameworks that can profoundly influence *presence* by setting

[1] Reaction time is the time a user takes to respond to a cue.

predefined emotional states or expectations in participants [78,90]. Two cornerstones of conditioning theory – classical and operant conditioning – were established by Pavlov [56] and Skinner [72], respectively. They demonstrate the responsiveness of individuals to stimuli when paired with either a positive or negative reinforcement. These principles in MR may help understand how participants might engage with and perceive the virtual environment.

The snapshot of how the *conditioning* will formulate the expectation of users and their *presence* is presented in Figure 2. When exposed to *positive conditioning* in MR, users will perceive scenes to be more realistic with intuitive interactions and appropriate feedback mechanisms. As suggested by Vroom's expectancy theory [87], this positive framing can raise their expectations, possibly resulting in heightened engagement and an enhanced sense of *presence* during the conditioning stage. An ancillary effect of this positive conditioning can be correlated with their response times. The alignment of the MR environment with the user's positive expectations might correlate with fluid interactions and quicker response times, reflecting their heightened state of immersion. However, during immediately subsequent interactions, any deviation from this positively conditioned experience, perhaps in the form of a more complex or different MR scene, might not only diminish the sense of *presence* due to a stark contrast from the initial conditioning but can correlate to slower reaction times. This delay can be attributed to the cognitive overhead of reconciling the disparity between their positive conditioning and the new, contrasting MR environment. This behavioral bias resonates with the effective priming theory [6], where prior positive exposure might shade subsequent interactions.

Conversely, *negative conditioning*, characterized by challenging interfaces or frustrating user experiences, prepares participants to expect more of the same difficulties. Such initial experiences might make them more alert and responsive, correlating to an enhanced sense of presence and quicker response times in the next MR encounter, especially if it's more realistic than the conditioning phase. Here, the cognitive dissonance theory [21] becomes particularly relevant, suggesting that users might experience discord if the actual MR experience contrasts sharply with their negative conditioning.

Control conditioning provides users with an unbiased experience [17]. In this setting, MR presence reflects the scene's design and interactivity, uninfluenced by any immediate prior experiences and pre-set bias, echoing the importance of minimizing cognitive bias in perception [40]. Response times in this condition represent a user's inherent interaction with the environment without the influence of prior conditioning [42]. Therefore, it serves as a baseline for our study, against which we contrast and evaluate the outcomes from the positive or negative conditioning.

3.2 Hypothesis

Given these observations, we hypothesize that the strategic application of conditioning within MR settings can influence the user's sense of *presence* and their *reaction times* to stimuli. Faster or slower response times can provide insights into a user's cognitive state and their level of immersion [52,84,92]. By calibrating initial experiences, researchers can anticipate and analyze subsequent user behavior and perceptions. Our hypothesised (**H**) are as follows:

H1: Manipulation in conditioning leads to a change in presence.

H2: **a** Prior `exposure` to gaming impacts presence.
- **b** Prior `exposure` to AR impacts presence.
- **c** Prior `exposure` to VR impacts presence.
- **d** Interaction between gaming `exposure` and `conditioning` impacts presence.

H3: During conditioning, if a participant interacts with highly realistic objects with plausible behaviors, their level of presence during the trial would be lower than the control group.

H4: During conditioning, if a participant interacts with less realistic objects with implausible behaviors, their level of presence during the trial would be higher than the control group.

H5: A change in the presence of a participant correlates to a change in the participant's reaction time.

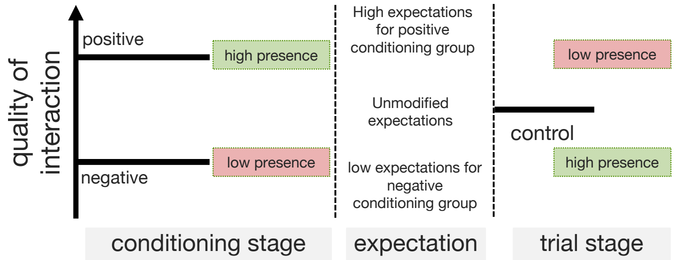

Fig. 2: Interplay between Conditioning and Presence.

H6: The sense of presence for a participant will change over time during a single session.

H7: Presence and reaction time are correlated.

3.3 Experimental Design

To evaluate our hypotheses on assessing the correlation between the presence and reaction time and how they are affected by human factors such as conditioning, we structured our experiment in two stages: the ***conditioning (C)*** stage and the ***trial (T)*** stage. During the conditioning stage, our objective is to shape participants' expectations for the subsequent trial, modulating their perception and response to stimuli and allowing us to observe the impact on presence under varied conditioning settings. In crafting our virtual scenes, we have striven for a balance, ensuring they are neither overly complex nor too simplistic, optimizing user engagement in both stages of the experiments [15, 33, 77]. To create positive and negative conditioning; and trials afterward, we use technical factors such as realism (appearance of objects), plausibility (behavior of objects), and interactions as priming knobs.

3.3.1 Conditioning Stage Design

To assess the effect of conditioning on presence and reaction times, we divide users into three groups: positive conditioning (***PC***), negative conditioning (***NC***), and a ***control*** group. We temper users' expectations in ***PC*** and ***NC*** groups to see how it would affect their presence and reaction times in the subsequent trial (***PT*** and ***NT***, respectively). The ***control*** serves as a baseline in the study, and the users do not engage in the conditioning stage experiment.

For conditioning experiments, we designed a scene with *one physical box*[2] and *two interactable virtual boxes [2]* visually represented in a solid white color to exude realism. In ***PC***, virtual boxes closely mimic real-world expectations: they can not float in the air and remain impervious to other objects passing through them. Interactions were managed through built-in hand tracking, enabling users to use their hands and fingers to grab, stack, and move the box consistent with real-world interactions (pointing or clicking gestures were excluded from the conditioning experiment interaction design) [9]. For instance, to move the box, one would need to grab it from the sides or bottom or push it by hand or with a machine/object. This precise hand-tracking is a feature and a bridge to a more immersive experience. Ensuring virtual interactions align closely with real-world expectations minimizes cognitive strain [34] and can enhance the presence as a byproduct.

In the ***NC*** stage, the virtual boxes retained their white color but were rendered transparent, allowing users' hands to pass through them, which visually deviates from real-world expectations. All interactive capabilities and physics models were removed: boxes hovered in the air and moved randomly in front of users without any physical initiation. The absence of colliders meant the boxes could intersect, further straying from real-world behavior. In essence, users' actions yielded no tangible results, and the boxes exhibited unrealistic appearance and implausible behavior.

3.3.2 Trial Stage Design

Drawing from cognitive load theory [79, 83] and embodied cognition [62], we devised a task based on pattern recall where user actions

[2]The shape, texture, and size of the physical box is not relevant to our experiment and is only used to influence the user subconsciously.

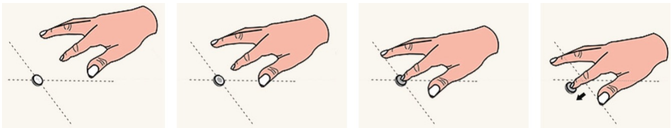

Fig. 3: Direct hand interaction poses for button press task [49].

Fig. 4: Direct hand interactions for grab, move, and rotate tasks [49].

mirror real-world interactions, commonly known as Corsi's 3D block-tapping test [10], a measure of short-term visuospatial memory. Our objective is not to measure cognitive strain or memory retention. Instead, we use this task to craft an engaging scene, enabling users to immerse themselves and allowing us to evaluate the effect of conditioning on presence and reaction times. Also, since we are not assessing the task's difficulty as part of our hypothesis, we do not increase the complexity throughout the trials.

3.4 Interaction Design

Interactions in our experimental task required users to use their hands continuously. We positioned elements with ergonomic considerations to prevent muscular fatigue [32]. We adhered to Hololens2 near interaction and placed primary interaction elements (boxes and buttons) within 35-50cm of the user's abdomen [2]. To minimize accidental selections, we introduced and instructed users to adopt hand poses aligned with the Hololens2 standard interaction paradigm. Pointing button press hand pose for the trial stage (shown in Figure 3) and a grab pose (shown in Figure 4) for the conditioning stage. In the trial stage, the interaction mechanism is a physics-based *pressable round push button* designed in Unity3D, simulating real-world button dynamics. This virtual button mimics the tactile feel of a real spring button, requiring a deliberate and physical push to register as a click, offering visual and auditory feedback upon interaction. Users can only use the index finger of either hand to complete the button press task, as shown in Figure 3. Auditory feedback is integrated through Microsoft spatializer [3] to create the perception that a sound originates from a specific button in a scene. Furthermore, virtual buttons were placed on the table surface to compensate for the haptic feedback.

3.5 Accounting for Prior Exposure

Research has shown that gamers possessing enhanced cognitive [30] and spatial capabilities [20] often experience heightened presence within the virtual environment. This amplified presence likely stems from their familiarity with digital interfaces and honed ability to navigate and process intricate virtual scenes [37, 94]. While we did not target individuals based explicitly on their gaming exposure during recruitment, we strategically placed them into groups as participants enrolled. This allocation ensured a balanced distribution rooted in their prior gaming experience. Such an approach becomes even more significant when considering the enduring and cumulative nature of conditioning. As elucidated by foundational works, conditioning, closely interwoven with experience, accumulates and augments user behavior with every repeated exposure [56, 60]. Through our sequential assignment methodology, we aimed to mitigate any potential confounding effects that the interplay between game familiarity and the accumulative aspect of conditioning might introduce to our study.

4 User Study

This section presents the details of our user study, experimental tasks, measures, and procedures.

Table 1: Demographic data and media usage (past 5 years) across all conditions and participants. The key for frequency: never/almost never; rarely (< 2times); occasionally (a few times); frequently in the past; frequently (> 2times/month).

demographics	# participants (from a total of 60 participants)
gender	24 female; 35 male; 1 preferred not to answer
age	mean = 22.7 years (STD = 4.43)
frequency of VR experience	16 never used; 20 rarely; 18 occasionally; 3 frequently; 3 frequently in the past
frequency of AR/MR experience	17 never used; 21 rarely; 19 occasionally; 2 frequently; 1 frequently in the past
frequency of Gaming	6 never used; 7 rarely; 16 occasionally; 25 frequently; 6 frequently in the past

4.1 Participants

Sixty-two participants took part in the study. Due to technical issues, two participants were excluded, resulting in 60 remaining participants for data analyses. Twenty participants were included in each condition (*positive, negative, and control*). All participants volunteered and provided written informed consent. They received $15 for their participation. All the participants had normal or corrected normal vision with contact lenses or glasses. The demographic distribution and media usage of participants can be seen in Table 1. The institution's ethics committee approved the study.

4.2 Material

We conducted this study on Hololens 2 [1], a self-contained device with a holographic processing unit with a Qualcomm Snapdragon 850 CPU, featuring eye, spatial, and hand-tracking capabilities, and dual displays with a 1440×936 pixels resolution with 110°FoV. We selected this device due to its direct view of the real world, which is crucial for safety-critical applications. Virtual scenes are crafted in Unity3D optimized for the Universal Windows Platform.

4.3 Experimental/Study Stages

The study is divided into two: the conditioning stage and the trial stage. Based on the group (`positive, negative, and control`), participants' experiences in the conditioning stage varied, but the trial stage remained the same for all the participants.

4.3.1 Conditioning Stage

In ***PC***, we immersed participants in the box scene with realistic appearance, behavior, and interaction (see more details in §3.3). Boxes were placed on the table in front of the participants, as shown in Figure 1 (left-left), and we asked them to interact with the boxes as if they were real, as illustrated in Figure 1 (left-center). We instructed participants on performing the grab, drag, move, and rotate interactions with the box. After 3 minutes of being immersed in the scene, participants were prompted by the study director to rate, on a scale of 1 to 10, their sense of connection to the virtual world, with the prompt question: *how connected do you feel to this virtual world?* and ended this stage. The prompt was explained to the participants during the initial briefing. In ***NC***, we immersed participants in box scene with no interactions and control, and boxes were not placed at the table in front of them, shown in Figure 1 (left-right). All the other steps remained the same as ***PC***. In ***control***, participants did not engage in the conditioning stage and directly advanced to the trial stage.

4.3.2 Trial Stage

Participants engaged in a memory puzzle task spanning ten trials for each condition. During each trial, the buttons lit up, and their rims were sequentially highlighted, constructing a distinct pattern from a predefined list, shown in Figure 1 (right-center). Once the pattern was shown, the buttons returned to their standard blue shade with grey rims (shown in Figure 1 (right-left)). Participants' task was to replicate the pattern, pressing the buttons in the precise order they were highlighted.

After replicating the pattern, participants received auditory feedback, seamlessly integrated using Microsoft spatializer plugin [3] to maintain their sense of presence, avoiding break-in presence [14]. A distinct `correct` sound played, signaling the correct identification of the button and its order in the sequence. However, if the selection was incorrect, an alternative `incorrect` sound was played to indicate the discrepancy. This audio feedback was to prime the participants to focus on correctness rather than quickly finishing the experiment. Participants were familiarized with these auditory cues during the initial briefing. A 25-second relaxation interval was initiated right after replicating the pattern. Here, participants observed a dummy calming pattern (shown in Figure 1 (right-right)) and were advised not to memorize it to memory (details in §4.6). Their attention was instead steered towards patterns shaped by the sequential red rims after the relaxation pattern. During relaxation, we asked the participants the same prompt question at the end of conditioning to assess the time-varying nature of presence.

Furthermore, the custom-developed application recorded various data points as participants engaged in the trial stage. To construct a comprehensive response time profile, timestamps of all events, such as button highlights, presses, relaxation pattern phases, and transitions, were captured concurrently in a separate thread. We also logged the sequence the participant entered and compared it against the predefined pattern highlighted to the participant. A comparison of the two determined the pattern accuracy: a perfect match indicated the correct pattern identification. An exact sequence match further categorized the input as "correct order and pattern". Otherwise, it was tagged as "correct pattern only". Although we are not evaluating memory retention in our study, this data was logged to disregard the response time of the wrong button presses to make it comparable across all users and conditions.

4.4 Experimental Task

The experimental task in the conditioning stage can be divided into two sub-tasks. The experiment objects are placed in the participant's FoV as a prerequisite. The first step for participants was to view the object (box) with their hand in a neutral position without raising their elbow. The second step for the participants was to lift their hands and perform drag or hold interaction with a five-finger grab, as shown in Figure 4. They could move, rotate, or stack the boxes with direct two-hand interaction [49].

For the trial stage, a pressable button with a holographic feature was designed for direct finger interaction [2]. To compensate for the absence of tactile feedback, it incorporated a depression mechanism. This mechanism gave the feeling of pressing down when the fingertip touches the button, allowing the button to move seamlessly with the fingertip's depth. Activation occurs when the button hits a designated depth (on press) or moves beyond that point and is released (on release). It adds a sound effect to provide feedback when the button is activated.

4.5 Measures

We measured *presence* using three questionnaires: the Witmer and Singer presence questionnaire (PQ) [94], the Igroup presence questionnaire (IPQ) [65,66], and Slater-Usoh-Steed questionnaire (SUS) [75]. PQ provides the four factors: sensory fidelity, immersion, and interface quality. The SUS and IPQ are based on three factors: the sense of physically being in the virtual environment, realism, and involvement, which could be beneficial regarding causality [90]. Presence scores for the questionnaires were obtained using 39 items (6 SUS, 14 IPQ, 19 PQ) on a 7-point scale. We did not modify any of the questions. We asked an additional prompt question (*how connected do you feel to this virtual world?*) on a 10-point scale during the conditioning stage and between trials. The reaction time was recorded by our software on HoloLens 2. On average, we collected four reaction time measurements per trial, totaling 40 data points per participant for the ten trials. The reaction time is measured in milliseconds (*ms*).

4.6 Pilot Study

Before initiating the main study, we conducted a pilot involving six participants to refine the experiment's parameters, ensuring minimal environmental biases and optimal reliability. The talk-aloud protocol gathered feedback on participant comfort and observations without swaying the experiment in favor of any specific parameter.

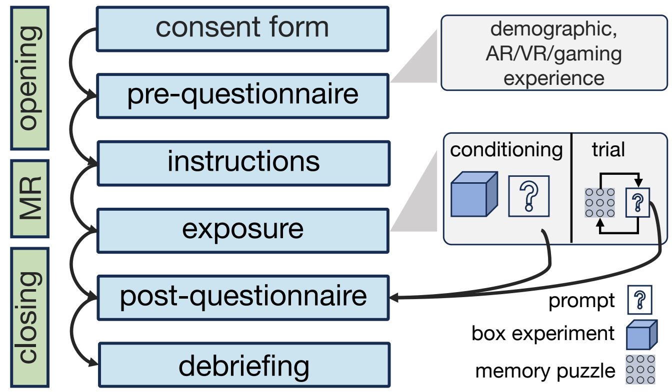

Fig. 5: The outline of the user study procedure.

Conditioning Stage Parameters In the 10-minute conditioning stage (boxes experiment), participants showed boredom by 5 minutes. By 3 minutes, 4 out of 6 reported lost interest. All preferred box placements close to the table's edge.

Pattern size and number of trails In Corsi Block-Tapping Test [10], a typical adult can correctly recall pattern size between 5 and 6 blocks (buttons). Patterns exceeding 7 blocks prove challenging for most, though individual capacities differ [41,54]. Our pilot assessed pattern sizes of 3 to 6 buttons, randomized for each participant. To evaluate the appropriate number of trials, we introduced feedback buttons "bored" and "too much now" and monitored both response times and correctness. Despite correctness being tangential to our study, its assessment ensured participants could complete sequences (most of the time). Our analysis indicated that the optimal structure involved a 4-button pattern size across 10 trials, as accuracy declined and response time variability surged beyond this for most participants.

Relax and reset pattern and duration Recognizing that cognitive tasks can deplete participants' cognitive resources and potentially impede presence, we deemed breaks between trials necessary [79]. We needed some downtime for participants and enough time between the rounds to ask the prompt question and for participants to answer it properly [41]. We experimented with a break duration of 10 seconds to a minute and alternated between three calming patterns and no pattern. In 4-button patterns of 10 trials, we randomized the calming patterns across 4 trials. Participants' feedback and performance metrics indicated a 25-second break as ideal, with a preference for slower, soothing patterns over rapid or no patterns.

4.7 Procedure

The study procedure is shown in Figure 5. Participants begin by reviewing and signing the consent form. They then provided demographic details such as gender, age, and familiarity with AR/VR/gaming. Based on the gaming familiarity of participants, we assigned the participants to the group as described in §3.5. Afterward, participants received a briefing about the study stages (experiments and questionnaires), scene apps, headset functionalities, expected interactions, hand poses to perform those interactions, prompt questions, and visual stimuli characteristics. Both Positive and negative conditioning groups are briefed the same about the conditioning stage not to influence the experience of ***NC*** intended to diminish presence. We did not explicitly tell participants about the physical box in the environment to subconsciously influence them to know what the physical box looks like in the MR. In the conditioning stage, participants put on the headset, and the study director started the box app. After 3 minutes, we asked them to answer a prompt question, and after that, they could take the headset off. Then, participants filled out presence questionnaires and an open-ended question about the experiment to record the baseline conditioning presence scores for the subsequent trial stage.

In the trial stage, participants from all groups played 10 trials of a memory puzzle game, each displaying a four-button pattern. After observing the pattern, participants replicated it, receiving audio feedback on their accuracy. In the briefing at the start of the session, all

Measure	control (μ, M, SD)	NC (μ, M, SD)	NT (μ, M, SD)	PC (μ, M, SD)	PT (μ, M, SD)
Presence Score (1-7), ALL	4.78, 4.75, 0.30	3.53, 3.48, 0.85	5.03, 4.98, 0.54	4.94, 4.96, 0.54	4.28, 4.40, 0.62
Presence Score (1-7), SUS	4.74, 4.75, 0.08	3.75, 3.65, 0.69	4.91, 4.89, 0.42	5.13, 5.22, 0.73	4.48, 4.43, 0.60
Presence Score (1-7), IPQ	4.62, 4.56, 0.41	3.36, 3.19 , 0.95	5.09, 5.19, 0.83	4.76, 4.76, 0.35	4.14, 4.06, 0.65
Presence Score (1-7), PQ	4.90, 4.98, 0.26	3.58, 3.53, 0.82	5.02, 4.99, 0.21	5.02, 5.11, 0.58	4.33, 4.12, 0.59
Prompt Score (1-10)	7.20, 7.55, 1.79	3.92, 3.56, 1.56	7.84, 7.95, 1.08	6.82, 7.15, 1.57	6.47, 6.25, 1.61
Reaction Time (ms)	4414, 4443, 1210	–	4130, 4232, 774	–	4984, 4600, 837

Table 2: Mean (μ), median (M), and standard deviation (SD) for the `presence score`, `prompt scores`, and `reaction time`. The control group (`control`), negative conditioning group during conditioning (`NC`) and during the trial (`NT`), and positive conditioning group during conditioning (`PC`) and during the trial (`PT`). Reaction time is not measured during the conditioning phase of the experiment.

Table 3: Shapiro-Wilk test statistics to determine the normality of `presence score`, `prompt score`, and `response time`.

Metric	Test Statistic	p-value
presence score	0.9753	0.2614
prompt score	0.9696	0.1391
response time	0.9883	0.8373

participants were primed to answer correctly and not hastily by iterating that they had to focus on the correctness of the pattern. This was to get as many data points as possible. Between each round of the puzzle, a prompt question is asked. After finishing all 10 trials and prompts, participants completed the same questionnaires as the conditioning stage and participated in a debriefing session. The entire process, encompassing the briefing, conditioning stage, questionnaires, trial stage, and feedback, lasted up to 35 minutes.

5 Results

This section presents quantitative and qualitative results that validate our hypothesis. We report high-level descriptive statistics (i.e., mean (μ), median (M), and standard deviation (SD)) for the three questionnaires, prompt scores, and reaction times for each condition in Table 2. We discuss these results in Section 5.3. In preprocessing, we had to discard an average of four reaction time values for each participant as they entered the wrong pattern for at least one trial (the maximum number of trials with the wrong pattern was 2).

Our experimental evaluation involved one independent variable, *conditioning*, and three dependent variables, *presence score*, *prompt score*, and *reaction time*. Our independent variable is categorical, while all dependent variables are continuous. The different observations of dependent variables were collected from different participants using a between-group study design. Furthermore, Table 3 presents the results for the Shapiro-Wilk normality test [70] for the three dependent variables, which suggests that all the variables follow a normal distribution.

The aforementioned characteristics of the results allow us to employ Analysis of Variance (ANOVA) tests. We use one-way ANOVA to assess the effect of *conditioning* on a single dependent variable. We use multivariate ANOVA (MANOVA) to assess the effect of *conditioning* on multiple independent variables. We also conducted t-tests to further confirm the direction of change in presence scores under various conditioning scenarios. Furthermore, we performed our subsequent statistical analyses using individual presence scores for each questionnaire. As all the questionnaires yield the same results, we only present the results using the aggregate presence scores.

5.1 Conditioning Effect on Presence and Prompt Score

Table 4 presents the results of a Multivariate Analysis of Variance (MANOVA) examining the effect of conditioning on presence measures, such as presence and prompt scores. We conducted multiple tests, including Wilks' lambda, Pillai's trace, Hotelling-Lawley trace, and Roy's greatest root, all of which indicate highly significant effects (all p-values are 0.0000). We only report the results for Wilks' lambda for brevity. The results under the "intercept" section suggest a significant multivariate effect of the conditioning variable on the combined dependent variables. The results for conditioning indicate significant univariate effects for conditioning on both presence and prompt scores.

Table 4: Multivariate Analysis of Variance (MANOVA) using `conditioning` as an independent variable, and `presence score` and `prompt score` as the dependent variables measuring presence.

intercept	Test Stats	F-Value	p-value
Wilks' lambda	0.0276	985.07	0.0000
conditioning	**Test Stats**	**F-Value**	**p-value**
Wilks' lambda	0.7036	3.5404	< 0.005

Table 5: One-way ANOVA to analyze the effect of `conditioning` on `presence score` and `prompt score` metrics.

Metric	F	p-value	η^2
presence score	11.14	< 0.01	0.28
prompt score	3.99	0.02	0.12

Table 5 presents our analysis of which specific variables contribute to the overall multivariate difference using one-way ANOVA. The main effect of conditioning on the presence score is significant (F-value $= 11.14$, $p < 0.01$). The effect size, measured by η^2, is 0.28, indicating a large effect. This means that the different conditioning levels can explain around 28% of the variance in the presence score. For the prompt score, we observe a similarly significant result (F-value $= 3.99$, $p = 0.02$). The effect size η^2 is 0.12, indicating a moderate effect. Based on these results, we accept H1.

Table 6 presents the results for `t-tests` further to prove the effect of conditioning on the presence and determine the direction of change. The positive conditioning in the paired samples t-test led to a significant increase in presence scores (`t-stat` $= 4.79$, p-value $= 1.27 \times 10^{-4}$). The one-sample t-test on the differences confirms that the mean difference is significantly greater than zero (`t-stat` $= -4.79$, p-value $= 6.33 \times 10^{-5}$). The negative conditioning in the paired samples t-test led to a significant decrease in presence scores (`t-stat` $= -6.45$, p-value $= 3.49 \times 10^{-6}$). The one-sample t-test on the differences confirms that the mean difference is significantly less than zero (`t-stat` $= 6.45$, p-value $= 1.74 \times 10^{-6}$). These results suggest that positive and negative conditioning significantly impacted presence scores in the expected directions. We can accept **H3** and **H4**.

5.2 Effect of Exposure on Presence

Table 7 shows the results of two-way ANOVA of gaming `exposure` and `conditioning` on presence. The main effect of `conditioning` is significant, but the main effect of gaming `exposure` is not ($p > 0.05$, $\eta^2 = 0.027$). However, the interaction between gaming `exposure` and `conditioning` has a significant effect ($p < 0.01$, $\eta^2 = 0.29$). Therefore, we reject **H2a** but accept **H2d**. While we collected AR and VR exposure data, we did not have enough participants in all subcategories for those variables and decided not to analyze them (**H2b** and **H2c** were not evaluated).

5.3 Reaction Time vs. Presence Score

In Table 2, we observe that the control group has values in the middle for μ and M for all the dependent variables: presence score, prompt score, and reaction time. During the trial, the negative conditioning has the highest μ and M for presence and prompt scores and vice versa. For reaction time, the trend is the opposite; positive conditioning

Table 6: `T-tests` for the effect of `conditioning` on `presence` (paired samples) and the direction of change (one-sample).

Conditioning Type	positive	negative
Paired Samples T-Test		
t-statistic	4.7925	-6.4514
p-value	1.27×10^{-4}	3.49×10^{-6}
One-Sample T-Test		
t-statistic	-4.7925	6.4514
p-value	6.33×10^{-5}	1.74×10^{-6}

Table 7: Two-ANOVA to determine the effect of gaming `exposure` and `conditioning` on presence.

Effect	F	p-value	η^2
`exposure`	0.67	0.62	0.027
`conditioning`	3.03	< 0.01	0.24
`exposure`×conditioning	14.41	< 0.01	0.29

has the highest value. Since reaction time is not measured during the conditioning stage, we cannot draw any conclusions about the relationship between presence score and reaction time. However, we can accept **H5** based on the statistically significant results during the trial stage.

Figure 7 shows participants' presence scores and reaction times scatter plots for individual conditioning groups and all participants. We observe that reaction time is inverse correlated with presence score, albeit with different slopes across different conditioning groups. The correlation values lie in the modest range between -0.6 and -0.65. This means that further analysis is needed to fully establish that reaction time can be used to quantify presence. However, we can accept **H7** based on our results.

5.4 Reaction Time and Prompt Scores over Time

The presence scores are collected using questionnaires at the end of the trial stage and do not have a time-varying nature. Therefore, we collected additional presence scores using prompts. Figure 6 shows each trial's prompt scores and reaction times over time. We observe that participants took significantly longer to complete the trials at the start of the trial stage. The reaction time and prompt scores dropped and settled to a steady state within the first 3-4 trials. Our results show that the presence (increases) and the reaction time (decreases) have an expected trend over time. Therefore, we can accept **H6**.

5.5 Quantitative and Thematic Analysis of Subscales

While our aggregate results demonstrate that conditioning affected presence scores, we wanted to investigate the factors that contributed to the change in presence. To do so, we disaggregated the questionnaires into their subscales: Involvement, Realism, Possibility to Act, Interface Quality, Possibility to Examine, and Self-Evaluation of Performance. We do not present the results for general and spatial presence subscales as they are highly correlated with overall scores. The results are reported in Table 8. We also present a thematic analysis of the open-ended participant responses.

Involvement. Transitioning from ***NC*** to ***NT***, participants felt a heightened sense of involvement, as observed in our quantitative as well as quantitative results. One noted, "This virtual experience felt somehow real. It was easy to control the 3D object". Similarly, those moving from ***PC*** to ***PT*** remarked, "It was amazing; it felt like being in a real-life video game". In ***control***, comments like "I thought the interface was very cool!" indicated an engaged experience.

Realism Similarly, the shift from ***NC*** to ***NT*** highlighted an enhanced sense of realism, which was observed in the realism subscale and participant responses. For instance, one participant said, "The buttons looked and reacted in a way that seemed real". Moving from ***PC*** to ***PT***, some described the virtual experience as more lifelike. In ***control*** group, feedback varied, from affirmation of realism to suggestions for improved immersion.

Subscale	`control`	`NC`	`NT`	`PC`	`PT`
Involvement	4.63	2.45	5.55	4.99	3.81
Realism	4.85	3.04	5.86	5.54	4.43
Possibility to Act	4.41	3.43	4.61	4.15	3.80
Interface Quality	4.77	4.53	4.80	4.65	4.58
Possibility to Examine	4.73	3.07	4.66	4.49	3.77
Self-Evaluation of Performance	5.17	3.39	4.43	5.17	4.68

Table 8: Presence scores for various subscales of questionnaires. The scores for each subscale are aggregated across all questionnaires.

Possibility to Act The ***NC*** to ***NT*** transition conveyed greater agency, as illustrated by remarks like, "The buttons in the VE were responsive". For those transitioning from ***PC*** to ***PT***, challenges were highlighted, such as the less control. Feedback was mixed in ***control***, with few suggesting they desired a more intuitive experience. Our quantitative results corroborated with the qualitative findings.

Interface Quality Feedback from ***NC*** to ***NT*** emphasized certain aspects of the headset. Some participants pointed out the limitations of FoV. Transitioning from ***PC*** to ***PT***, participants pointed out areas for improvement. One comment encapsulated a common sentiment for the ***control*** group: "The colors are not always consistent. They change when you look from different angles or adjust the headset". However, since the interface quality did not change across groups, there was no significant difference in presence scores for this subscale.

Possibility to Examine Curiosity underpinned the shift from ***NC*** to ***NT***, reflected in comments like "This time was more immersive as I could interact with the buttons". Those transitioning from ***PC*** to ***PT*** discussed adjusting to the environment. ***control*** group participants provided varied feedback, from appreciation of consistency to suggestions for deeper interaction. Our quantitative results corroborated the findings of our thematic analysis.

Self-Evaluation of Performance Participants transitioning from ***NC*** to ***NT*** noted improved confidence in navigating the VE, which manifested itself in the questionnaire subscale scores. Those from ***PC*** to ***PT*** provided balanced self-assessments. Feedback from ***control*** covered a range, including positive remarks like "It was great, and I enjoyed it".

6 Discussion

We measured *presence* across three groups: `positive, negative, and control`. With a notable effect size of 0.29 for the total presence score, it is evident that our cognitive and sensory/perceptual manipulations substantially influenced participants' immersion in the MR. This significant effect underscores the pivotal role of conditioning in shaping an individual's sense of connection in VE. In the following discussion, we delve deeper into these findings.

Overall Statistics and Normality The observation that the control group has median and mean values in between those of the two conditioning groups for all measured variables reinforces the notion that the control group acts as a neutral benchmark. This positioning highlights the opposing effects of positive and negative conditioning. Furthermore, the normal data distribution for these variables ensures the validity of subsequent parametric tests and analyses.

6.1 Presence

The Conditioning Effect H1 postulated that manipulation in conditioning would lead to a change in presence. The strong significance shown in the MANOVA and one-way ANOVA results emphasizes the robust impact of conditioning on both presence and prompt scores. Notably, the variance in presence score can be attributed to conditioning up to 28%. This was profoundly evident, with both positive and negative conditioning groups exhibiting shifts in presence scores, solidifying the importance of initial experiences. The gradient of effects between positive, negative, and control groups further underscored this.

Gaming Exposure's Complex Interplay on Presence While **H2** explored the interplay between gaming exposure and conditioning, our results revealed that mere exposure was not as critical as its interaction with conditioning. The primary takeaway here is that while gaming exposure in itself does not notably influence presence (leading to the rejection of **H2a**), the combination of gaming exposure with conditioning

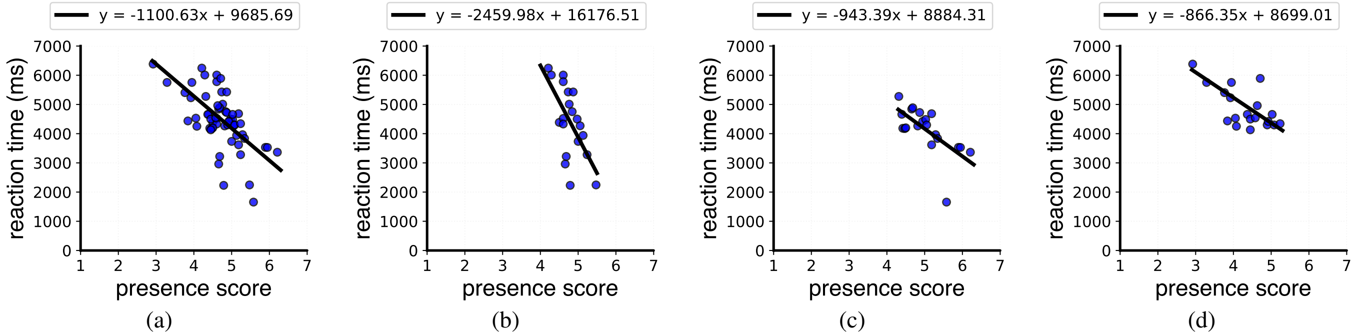

Fig. 6: Presence vs. Reaction Time: Presence decreases as reaction time increases. Reaction time and presence also show a modest correlation: (a) overall (-0.64), (b) control (-0.62), (c) positive (-0.66), and (d) negative (-0.64). Each blue circle represents a study participant. The black line is the linear regression fit for the data.

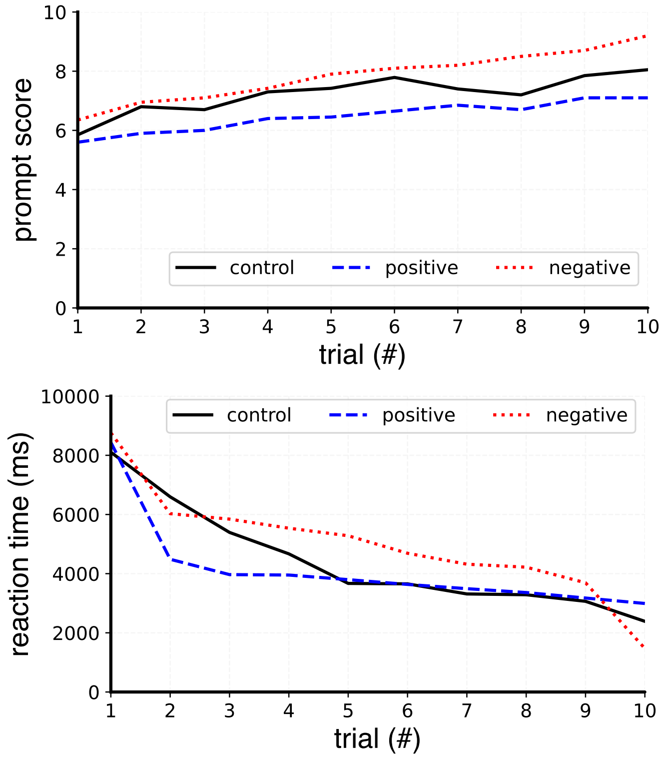

Fig. 7: Average prompt score and user `reaction time` for different `conditioning` groups. User `reaction time` recovers, and `prompt score` decreases over time.

does. This aligns with our discussion on the conditioning continuum, suggesting that our pre-existing biases from gaming do not operate in isolation but dynamically interact with new conditioning experiences. The acceptance of **H2d** underscores the complexity of these interactions and emphasizes that it is not just the individual factors but their synergistic effects that play a pivotal role in shaping MR experiences.
Counterintuitive Interplay of Conditioning and Presence The results from the t-tests offer a nuanced view of how different conditioning scenarios (positive and negative) specifically impact presence. **H3** and **H4**'s predictions about positive and negative conditioning impacting presence scores were validated. The heightened presence scores post-negative conditioning, as discussed, might stem from a mix of heightened awareness and adaptive response, while the positive conditioning experience aligns with the traditionally understood immersion dynamics. This inverse relationship in positive and negative conditioning outcomes leads us to accept **H3** and **H4**.
Dynamic Shifts in Presence H6 considered the possibility of presence evolving over time within a single session. The drop and subsequent stabilization of presence scores and reaction times in our results reinforced this, indicating that a participant moved from initial acclimatization to eventual comfort. The data showcasing reaction time and prompt scores over time is particularly revealing. The noticeable trend of participants taking longer during the initial trials, followed by a drop and stabilization, highlights the learning curve ending at acclimatizing to the virtual environment. This aligns with **H6**, emphasizing that their response times decrease as users become more acclimatized (or their presence deepens). It also indicates participants transitioning from conscious effort to a more automated, instinctual interaction pattern as the become familiar with the virtual environment.

6.2 Reaction Time

Presence and Reactions Time With Hypothesis **H5**, the suggestion that change in presence would correlate with a participant's reaction time was empirically validated. As participants felt more "present", their reactions streamlined, reflecting their comfort and ease in the VE.
Reaction Time as an Indicator of Presence The inverse correlation between presence scores and reaction times resonates with our hypothesis **H7**. While the correlation values are modest, the consistent negative trend across conditioning groups suggests that as participants feel more 'present' in the environment, their interactions become swifter, more intuitive, and more fluid. This potentially cements reaction time as an objective, quantifiable proxy for presence, augmenting the traditionally subjective nature of presence measurements.

6.3 Implications

Reinforcement of Conditioning Effects While reaction time was not directly measured during the conditioning stage, its nuanced manifestations during the trial stage, particularly its inversely proportional relationship with presence scores, accentuates the conditioning's residual effects. The fact that negative conditioning led to heightened presence scores but longer reaction times, while positive conditioning had the opposite effect, further underscores the conditioning's role in framing and influencing subsequent MR interactions.
Qualitative Observations The qualitative insights derived from participants echo the quantitative findings, offering deeper contextual grounding for our hypotheses. The heightened sense of involvement and the nuanced perception of realism, particularly when transitioning from ***NC*** to ***NT*** or from ***PC*** to ***PT***, provide empirical evidence supporting H1, which posited that conditioning influence presence. This congruence between subjective narratives and objective measurements underscores the validity of the hypothesized relationship. Furthermore, comments about the quality of the interface and the participant's ability to act or examine within the virtual space are particularly revealing. These feedback points reflect the embodied effects of conditioning, which, according to H3 and H4, would manifest in enhanced or diminished presence based on the nature of the conditioning. The participant self-evaluations, indicating varying levels of confidence and self-efficacy in navigating the MR environment, further strengthen the link between conditioning and subsequent user behavior, solidifying the foundations of our hypotheses. Participants' qualitative feedback dovetails with our hypothesized outcomes, reinforcing the importance of conditioning in modulating presence in MR environments.

By triangulating our insights with the established hypotheses, it becomes evident that while our predictive framework was robust, human interactions with MR, influenced by conditioning, are layered and multifaceted. The hypotheses provided a structure, but the richness of the data, coupled with our detailed discussions, truly brings to light the depth and complexity of the human experience within MR environments.

7 Limitations and Future Work

While our research underscores the significance of conditioning in modulating the sense of presence in MR, it's essential to recognize inherent limitations and avenues for future exploration. The study's participant cohort was limited to 62 individuals, and the results might exhibit more variations with a larger sample. Additionally, the design of the conditioning scenarios (positive, negative, and control) could be expanded upon. For instance, the nuance of mixed or alternating conditioning stimuli remains unexplored. Aspects such as object realism and plausibility, which we have considered, could be redefined or extended as technology advances. The current study was also limited in scope by primarily relying on the relationship between conditioning, presence, and reaction time. A holistic understanding would benefit from including other cognitive, psychological, and physiological measures.

Further, given the intricate interplay between technical familiarity and conditioning, it might be beneficial to probe deeper into the specificities of prior technological exposure. How might a user, seasoned in VR but naive to AR, react compared to one familiar with both? Future research could also explore more human factors, such as memory retention, cognitive load, agency, genetic predisposition, lifelong-habitude or emotional resilience, and their impact on presence in MR and its correlation with reaction time.

8 Conclusion

In this explorative study ($N = 60$), our research has illuminated the intricate relationships between `conditioning, presence, and reaction time` in MR. The way users are conditioned has profound implications on their sense of "being" within a virtual space and how swiftly they react within it. Through between-subjects study, we have derived critical insights into how prior experiences (`conditioning`), both positive and negative, shape user immersion and interactions (*presence*) and responsiveness (*reaction time*). We identified a notable effect size (0.28) between `presence and conditioning`, underlining the influence of a user's prior experiences on their sense of presence. Additionally, the correlation between `presence and reaction time` `(-0.64)` was indicative, suggesting that as users feel more "present," their response times may become more reflexive. These findings not only enhance our current understanding of MR interactions but also suggest that reaction time might emerge as a valuable indicator of presence in future MR applications with more exploration.

References

[1] Hololens 2. https://www.microsoft.com/en-us/hololens/hardware. 4

[2] Intractable objects. https://docs.microsoft.com/en-us/windows/mixed-reality/design/interactable-object. 3, 4, 5

[3] Microsoft spatializer. https://learn.microsoft.com/en-us/windows/mixed-reality/develop/unity/spatial-sound-in-unity. 4

[4] K. W. Arthur. *Effects of Field of View on Performance with Head-Mounted Displays*. PhD thesis, 2000. 2

[5] W. Barfield and S. Weghorst. The sense of presence within virtual environments: A conceptual framework. *Advances in Human Factors Ergonomics*, 19, 1993. 2

[6] J. A. Bargh and P. Pietromonaco. Automatic information processing and social perception: the influence of trait information presented outside of conscious awareness on impression formation. *Journal of personality and Social psychology*, 43(3):437, 1982. 3

[7] C. Basdogan, C.-H. Ho, M. A. Srinivasan, and M. Slater. An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction*, 7, 2000. 2

[8] T. Baumgartner, L. Valko, M. Esslen, and L. Jäncke. Neural correlate of spatial presence in an arousing and noninteractive virtual reality: An eeg and psychophysiology study. *Cyberpsychology & Behavior : the Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society*, 9, 2006. 2

[9] S. Benford, C. Greenhalgh, G. Reynard, C. Brown, and B. Koleva. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on computer-human interaction (TOCHI)*, 5(3):185–223, 1998. 3

[10] D. B. Berch, R. Krikorian, and E. M. Huha. The corsi block-tapping task: Methodological and theoretical considerations. *Brain and cognition*, 38(3):317–338, 1998. 4, 5

[11] C. Botella, J. Fernández-Álvarez, V. Guillén, A. García-Palacios, and R. Baños. Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current psychiatry reports*, 19:1–13, 2017. 1, 2

[12] C. C. Bracken, G. Pettey, T. Guha, and B. E. Rubenking. Sounding out small screens and telepresence. *Journal of Media Psychology*, 2010. 2

[13] M. Bruce and H. Regenbrecht. A virtual reality claustrophobia therapy system-implementation and test. In *2009 IEEE Virtual Reality Conference*, pp. 179–182. IEEE, 2009. 2

[14] L. Brübach, F. Westermeier, C. Wienrich, and M. E. Latoschik. Breaking plausibility without breaking presence - evidence for the multi-layer nature of plausibility. *IEEE Transactions on Visualization and Computer Graphics*, 28, 2022. 4

[15] Y. Chandio, N. Bashir, V. Interrante, and F. M. Anwar. Investigating the correlation between presence and reaction time in mixed reality. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–13, 2023. doi: 10.1109/TVCG.2023.3319563 1, 2, 3

[16] J. Colomb and B. Brembs. The biology of psychology: simple conditioning? *Communicative & integrative biology*, 3(2):142–145, 2010. 1, 2

[17] T. D. Cook, D. T. Campbell, and A. Day. *Quasi-experimentation: Design & analysis issues for field settings*, vol. 351. Houghton Mifflin Boston, 1979. 3

[18] Y. A. de Kort, A. L. Meijnders, A. A. Sponselee, and W. A. IJsselsteijn. What's wrong with virtual trees? restoring from stress in a mediated environment. *Journal of environmental psychology*, 26(4):309–320, 2006. 2

[19] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in Psychology*, 6, 2015. 1, 2

[20] J. Feng, I. Spence, and J. Pratt. Playing an action video game reduces gender differences in spatial cognition. *Psychological science*, 18(10):850–855, 2007. 4

[21] L. Festinger. Social comparison theory. *Selective Exposure Theory*, 16:401, 1957. 3

[22] J. Freeman, S. Avons, D. Pearson, and W. Ijsselsteijn. Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence*, 8, 1999. 2

[23] J. Freeman, S. E. Avons, R. Meddis, D. E. Pearson, and W. IJsselsteijn. Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators & Virtual Environments*, 9, 2000. 2

[24] A. Gaggioli, M. Bassi, and A. Fave. Quality of experience in virtual environments. *Emerging Communication*, 5:121–136, 2003. 2

[25] M. Gandy, R. Catrambone, B. MacIntyre, C. Alvarez, E. Eiriksdottir, M. Hilimire, B. Davidson, and A. C. McLaughlin. Experiences with an ar evaluation test bed: Presence, performance, and physiological measurement. In *IEEE International Symposium on Mixed and Augmented Reality*, 2010. 2

[26] M. Giesel, A. Nowakowska, J. Harris, and C. Hesse. Perceptual uncertainty and action consequences independently affect hand movements in a virtual environment. *Scientific Reports*, 10: 22307, 12 2020. 1, 2

[27] E. Glotzbach-Schoon, M. Andreatta, A. Mühlberger, and P. Pauli. Context conditioning in virtual reality as a model for pathological anxiety. *e-Neuroforum*, 4:63–70, 2013. 2

[28] F. Graham. Attention: The heartbeat, the blink and the brain. *Attention and Information Processing in Infants and Adults: Perspectives from Human and Animal Research (Eds: B.A. Campell, H. Hayne, R. Richardson)*, 1992. 2

[29] S. Grassini and K. Laumann. Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, 11, 2020. 2

[30] C. S. Green and D. Bavelier. Action video game modifies visual selective attention. *Nature*, 423(6939):534–537, 2003. 4

[31] C. Grillon, J. M. Baas, B. Cornwell, and L. Johnson. Context condition-

ing and behavioral avoidance in a virtual reality environment: effect of predictability. *Biological psychiatry*, 60(7):752–759, 2006. 1, 2

[32] J. T. Hansberger, C. Peng, S. L. Mathis, V. Areyur Shanthakumar, S. C. Meacham, L. Cao, and V. R. Blakely. Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions. In *Virtual, Augmented and Mixed Reality: 9th International Conference, VAMR 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 9*, pp. 505–520. Springer, 2017. 4

[33] W. A. IJsselsteijn. Presence in depth. *Technische Universiteit Eindhoven Eindhoven, The Netherlands*, 2004. 2, 3

[34] N. Jamalian, M. Gillies, F. F. Leymarie, and X. Pan. The effects of hand tracking on user performance: an experimental study of an object selection based memory game. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 768–776. IEEE, 2022. 3

[35] Y. Jang and E. Park. An adoption model for virtual reality games: The roles of presence and enjoyment. *Telematics and Informatics*, 42:101239, 2019. 1, 2

[36] T. F. I. J.D. Prothero, D.E. Parker. Towards a robust, quantitative measure for presence. *Proceedings of the Conference on Experimental Analysis and Measurement of Situation Awareness*, 1995. 2

[37] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9):641–661, 2008. 4

[38] C. Jicol, C. Clarke, E. Tor, R. M. Dakin, T. C. Lancaster, S. T. Chang, K. Petrini, E. O'Neill, M. J. Proulx, and C. Lutteroth. Realism and field of view affect presence in vr but not the way you think. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. 2

[39] C. Jicol, C. H. Wan, B. Doling, C. H. Illingworth, J. Yoon, C. Headey, C. Lutteroth, M. J. Proulx, K. Petrini, and E. O'Neill. Effects of emotion and agency on presence in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021. 2

[40] D. Kahneman, P. Slovic, and A. Tversky. *Judgment under uncertainty: Heuristics and biases*. Cambridge university press, 1982. 3

[41] M. J. Kane, A. R. Conway, T. K. Miura, and G. J. Colflesh. Working memory, attention control, and the n-back task: a question of construct validity. *Journal of Experimental psychology: learning, memory, and cognition*, 33(3):615, 2007. 5

[42] A. E. Kazdin and M. K. Nock. Delineating mechanisms of change in child and adolescent therapy: Methodological issues and research recommendations. *Journal of child psychology and psychiatry*, 44(8):1116–1129, 2003. 3

[43] S. J. Kim, T. H. Laine, and H. J. Suk. Presence effects in virtual reality based on user characteristics: Attention, enjoyment, and memory. *Electronics*, 10, 2021. 1

[44] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30, 1993. 2

[45] M. E. Latoschik and C. Wienrich. Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. *Frontiers in Virtual Reality*, 3, 2022. 1, 2

[46] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx. The relationship between individual characteristics and experienced presence. *Computers in Human Behavior*, 29(4):1519–1530, 2013. 2

[47] M. Lombard and T. Ditton. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3, 1997. 2

[48] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks. Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics*, 21, 2002. 1, 2

[49] Microsoft. Direct hand manipulation in hololens2. https://learn.microsoft.com/en-us/windows/mixed-reality/design/direct-manipulation, 2023. Accessed September. 4, 5

[50] Microsoft. Types of Mixed Reality Apps. https://learn.microsoft.com/en-us/windows/mixed-reality/discover/types-of-mixed-reality-apps, 2023. Accessed September. 1

[51] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented Reality: A Class of displays on the Reality-Virtuality Continuum. In *Telemanipulator and Telepresence Technologies*. SPIE, 1995. 2

[52] K. Nesbitt, S. Davis, K. Blackmore, and E. Nalivaiko. Correlating reaction time and nausea measures with traditional measures of cybersickness. *Displays*, 48, 2017. 3

[53] D. Nunez. How is presence in non-immersive, non-realistic virtual environments possible? In *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, 2004. 2

[54] A. M. Owen, K. M. McMillan, A. R. Laird, and E. Bullmore. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, 25(1):46–59, 2005. 5

[55] D. Paes, J. Irizarry, and D. Pujoni. An evidence of cognitive benefits from immersive design review: Comparing three-dimensional perception and presence between immersive and non-immersive virtual environments. *Automation in Construction*, 130:103849, 2021. 1, 2

[56] I. P. Pavlov. *Conditioned reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.* Oxford University Press, London, 1927. 3, 4

[57] D. Pimentel and C. Vinkers. Copresence with virtual humans in mixed reality: The impact of contextual responsiveness on social perceptions. *Frontiers in Robotics and AI*, 8, 2021. 1, 2

[58] S. Putze, D. Alexandrovsky, F. Putze, S. Höffner, J. D. Smeddinck, and R. Malaka. Breaking the experience: Effects of questionnaires in vr user studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020. 2

[59] H. Regenbrecht and T. W. Schubert. Measuring presence in augmented reality environments: Design and a first test of a questionnaire. *ArXiv*, abs/2103.02831, 2021. 2

[60] R. A. Rescorla. Informational variables in pavlovian conditioning. In *Psychology of learning and motivation*, vol. 6, pp. 1–46. Elsevier, 1972. 4

[61] G. Riva, F. Davide, and W. IJsselsteijn. Measuring presence: Subjective, behavioral and physiological methods. *Being there: Concepts, effects and measurement of user presence in synthetic environments*, 5, 2003. 2

[62] G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192, 2004. 3

[63] S. Safikhani, M. Holly, A. Kainz, and J. Pirker. The influence of in-vr questionnaire design on the user experience. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, 2021. 2

[64] M. Sanchez-Vives and M. Slater. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6, 2005. 1

[65] T. Schubert. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15, 2003. 2, 5

[66] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10, 2001. 2, 5

[67] V. Schwind, P. Knierim, L. Chuang, and N. Henze. "where's pinky?": The effects of a reduced number of fingers in virtual reality. In *ACM CHI Conference on Human Factors in Computing System in Play*, 2017. 2

[68] V. Schwind, P. Knierim, N. Haas, and N. Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019. 2

[69] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. "these are not my hands!": Effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017. 2

[70] S. S. SHAPIRO and M. B. WILK. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3-4):591–611, dec 1965. 6

[71] T. B. Sheridan. Musings on telepresence and virtual presence. *Presence Teleoperators & Virtual Environment*, 1, 1992. 2

[72] B. F. Skinner. *The Behavior of organisms: An experimental analysis*. Appleton-Century, New York, 1938. 3

[73] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 2009. 2

[74] M. Slater, V. Linakis, M. Usoh, and R. Kooper. Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 1996. 2

[75] M. Slater and A. Steed. A virtual presence counter. *Presence*, 9, 2000. 2, 5

[76] M. Slater, M. Usoh, and Y. Chrysanthou. The influence of dynamic shadows on presence in immersive virtual environments. *Eurographics Workshops on Virtual Environments*, 1999. 2

[77] M. Slater, M. Usoh, and A. Steed. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction*, 2, 1995. 3

[78] M. Slater and S. Wilbur. A framework for immersive virtual environ-

ments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6, 1997. 2, 3

[79] A. Steed, Y. Pan, F. Zisch, and W. Steptoe. The impact of a self-avatar on cognitive load in immersive virtual reality. In *IEEE Virtual Reality*, 2016. 3, 5

[80] J. Steuer, F. Biocca, M. R. Levy, et al. Defining virtual reality: Dimensions determining telepresence. *Communication in the age of virtual reality*, 33:37–39, 1995. 1, 2

[81] B. Stevens and J. Jerrams-Smith. The sense of object-presence with projection-augmented models. In *International Workshop on Haptic Human-Computer Interaction*, 2000. 2

[82] B. Stevens, J. Jerrams-Smith, D. Heathcote, and D. Callear. Putting the virtual into reality: Assessing object-presence with projection-augmented models. *Presence: Teleoperators & Virtual Environments*, 11, 2002. 2

[83] J. Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2):257–285, 1988. 3

[84] K. Szczurowski and M. Smith. Measuring presence: hypothetical quantitative framework. In *23rd International Conference on Virtual System & Multimedia*, 2017. 3

[85] X. d. Tinguy, C. Pacchierotti, M. Emily, M. Chevalier, A. Guignardat, M. Guillaudeux, C. Six, A. Lécuyer, and M. Marchal. How different tangible and virtual objects can be while still feeling the same? In *2019 IEEE World Haptics Conference (WHC)*, pp. 580–585, 2019. 2

[86] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence*, 9, 2000. 2

[87] V. H. Vroom. Work and motivation. *Wiley*, 1964. 3

[88] I. Wagner, W. Broll, G. Jacucci, K. Kuutii, R. McCall, A. Morrison, D. Schmalstieg, and J.-J. Terrin. On the role of presence in mixed reality. *Presence*, 18, 2009. 2

[89] S. Weech, S. Kenny, M. Lenizky, and M. Barnett-Cowan. Narrative and gaming experience interact to affect presence and cybersickness in virtual reality. *International Journal of Human-Computer Studies*, 138:102398, 2020. 2

[90] F. Westermeier, L. Brübach, M. E. Latoschik, and C. Wienrich. Exploring plausibility and presence in mixed reality experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2680–2689, 2023. 1, 2, 3, 5

[91] C. Wienrich, P. Komma, S. Vogt, and M. E. Latoschik. Spatial presence in mixed realities–considerations about the concept, measures, design, and experiments. *Frontiers in Virtual Reality*, 2, 2021. 2

[92] M. Wiesing, G. Fink, and R. Weidner. Accuracy and precision of stimulus timing and reaction times with unreal engine and steamvr. *PLoS ONE*, 15, 2020. 3

[93] J. R. Wilson, S. Nichols, and C. Haldane. Presence and side effects: Complementary or contradictory? In *HCI (2)*. Elsevier, 1997. 2

[94] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators & Virtual Environments*, 7, 1998. 2, 4, 5

[95] P. Zimmons and A. Panter. The influence of rendering quality on presence and task performance in a virtual environment. In *IEEE Virtual Reality, 2003. Proceedings.*, pp. 293–294, 2003. 2