Inattentional Blindness in Virtual Reality Multisensory Tasks

Maryam Keyvanara
Robert S. Allison*
maryam.k1ara@gmail.com
allison@cse.yorku.ca
Department of EECS, York University
Toronto, Ontario, Canada

ABSTRACT

Inattentional blindness occurs when individuals fail to perceive visible elements in a scene because their attention is focused elsewhere. This phenomenon is particularly relevant in Virtual Reality (VR), where subtle scene changes can go unnoticed during task engagement, enabling techniques such as redirected walking and seamless environment manipulation. We investigated how multisensory engagement with a moving object, either by passively following it with gaze alone or by combining gaze with active hand movement, affects users' ability to detect rapid background changes in a VR scene. A shape identification task in each trial was used to confirm that participants maintained fixation on the moving object throughout. Our results show that participants were generally less sensitive to subtle scene shifts when performing an object displacement task, regardless of sensory engagement modality. However, participants using only gaze were significantly better at detecting larger shifts than those using both hand and gaze. These findings inform the design of immersive VR systems that aim to mask visual changes during task engagement.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**.

KEYWORDS

Inattentional Blindness, Virtual Reality, Haptics, Change Blindness, Multisensory Interaction

ACM Reference Format:
Maryam Keyvanara and Robert S. Allison. 2026. Inattentional Blindness in Virtual Reality Multisensory Tasks. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Graphics Interface)*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Human vision is a primary sense for interacting with the real and virtual worlds, providing essential information about our environment [6]. However, human perception is inherently multisensory, relying on the integration of various sensory inputs. In everyday life, we rarely rely solely on vision to understand and engage with our surroundings [12]. In addition to visual information, proprioception, the sense of the position and movement of the body, plays a critical role in enabling us to interact with our environment effectively. Proprioceptive feedback helps coordinate motor intent and execution, allowing individuals to adjust their actions in real time without relying exclusively on visual confirmation [31, 44].

This integration of proprioception, motor intent, and other sensory modalities such as haptics and auditory cues contributes to a more comprehensive perception of the world. In immersive technologies like virtual reality (VR), this multisensory integration becomes even more crucial, as it enables more natural and responsive interactions, enhancing the sense of presence and realism [49]. As such, understanding how proprioception and other senses integrate is essential for improving interaction design in VR environments [47]. In VR environments, users are typically engaged in goal-oriented tasks, which require varying amounts of attention and sensory engagement [48]. In many VR systems users interact with virtual objects through controllers, leap motion sensors, hand tracking, or virtual gloves [1, 4, 20, 23], offering intuitive and responsive interfaces. However, such systems often require significant cognitive and sensory resources. When users focus intensely on a specific task, they may experience phenomena such as inattentional blindness, in which important visible elements go unnoticed because the user's attention is occupied elsewhere [25, 26].

Despite growing interest in the limits of sensory perception in VR [16, 45], little is known about how combining multiple sensory modalities, such as gaze and hand interaction, affects perception when users are highly task-focused. Prior work has shown that increased task load reduces sensitivity to background changes, but it remains unclear whether the modality of engagement, manual+visual versus visual-only, systematically alters perceptual thresholds for detecting scene changes in immersive VR. Our study addresses this gap by directly comparing detection performance under these conditions, examining how dual-modality engagement (vision and haptics) versus single-modality engagement (vision-only) influence users' awareness of changes in a dynamic virtual environment while engaged in a secondary task.

By contrasting these sensory conditions, we aim to provide deeper insights into how multisensory processing both supports and challenges user performance in VR environments. The outcomes of this study can have practical implications for the design of VR systems that require sustained attention and situational awareness. In particular, applications such as training simulations, medical interventions, and industrial operations may benefit from

Graphics Interface, June 09–12, 2026, Waterloo, ON

ACM ISBN 978-1-4503-XXXX-X/2018/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

understanding how multisensory engagement shapes user responsiveness to environmental changes. Guided by prior literature on task load and sensory integration, we hypothesized that participants would detect fewer changes in the virtual environment (here scene shifts) in the hand-with-gaze condition than in the gaze-only condition, reflecting the higher cognitive and motor demands of the dual-modality task.

Building on these foundations, we make three key contributions:

- We empirically compare inattentional blindness under gaze-only versus gaze-plus-hand VR interactions.
- We present evidence that multisensory tasks impose higher cognitive load, reducing sensitivity to background changes.
- We discuss implications for designing subtle environmental manipulations in immersive VR systems.

2 RELATED WORK

Inattentional blindness is a perceptual experience in which we do not see highly visible objects that are in our field of view because our attention is elsewhere [25, 26]. Probably the most famous example of inattentional blindness is presented in [35], which involved participants watching a video of people passing a basketball and counting the passes. During the video, a person in a gorilla costume walks through the scene, but most participants failed to notice the gorilla, demonstrating how focused attention can cause people to miss unexpected yet obvious events. These findings showed that inattentional blindness is influenced by object similarity and task difficulty. This demonstration was later replicated in a 3D VR setting [34], where participants were more likely to notice the gorilla compared to traditional settings. One possible explanation is that the increased sense of presence and spatial proximity in VR altered attentional allocation, making the unexpected object more salient. Research has shown that the amount of workload affects the strength of inattentional blindness, and that the perceptual load of a task has an important role in determining conscious perception [7]. The detection of a stimulus depends on how much attention it can attract and the difficulty of the primary task. With more difficult tasks, more attentional resources are required [15, 35]. Wang et al. (2022) studied inattentional blindness in AR heads up displays, and found that a higher workload induced a stronger inattentional blindness than a low workload [46]. Despite this, the performance of the users on those tasks, i.e. individual differences, did not affect the amount of inattentional blindness [37].

Virtual Reality heavily relies on sensory stimulation, leading researchers to explore perceptual illusions such as inattentional blindness. This phenomenon and its applications in VR have gained more attention in the recent years [16, 45], with applications including redirecting the users while navigating in VR [16], and also through auditory distractions. One study found that an auditory effect can degrade the detection of visual stimuli by VR users [27]. They recorded users' eye movements and found that their gaze behaviour did not change during this degradation, and concluded this degradation occurs even in the absence of saccades towards the sound source.

Change blindness is another failure of visual awareness in which an observer fails to notice an obvious change, and has been studied in many VR studies. One study compared how change blindness differed between vision and touch. They showed that unimodal attention during change blindness, visual only or haptic only, showed better performance compared to crossmodal conditions where both vision and touch were involved. Their results showed that change blindness is a multisensory process, and that vision and haptics differ in their encoding and memory limits [2].

The ability to detect small scene shifts has been studied in VR environments, where visual consistency is crucial for immersion [18, 29, 43]. Disruptions in these environments (e.g., through minor scene shifts or rotational changes) can be jarring or lead to disorientation or motion sickness. Display shifts in VR applications are often used to manipulate the user's perception of the virtual environment, enabling techniques like redirected walking or scene adjustments that subtly alter objects or perspectives without the user noticing [19, 33, 42]. Change blindness is another perceptual technique that can achieve redirected walking (RDW) [28, 39, 41], where users rely on the inability to notice changes to manipulate a 3D environment. Furthermore, Zenner et al. examined blink-suppressed hand redirection, showing that shifting the virtual hand during blinks and adjusting offsets while the eyes are open can enable unnoticeable redirection, with experiments that confirmed feasibility and provided detection thresholds [50].

To our knowledge, no prior VR studies have directly examined how the perception of the virtual environment changes when users perform a secondary task that requires focused attention, particularly when comparing conditions involving one sensory modality versus two. Our study is the first to systematically investigate how different combinations of scene shifts and an additional attention-demanding task influence user perception.

3 METHODS AND PROCEDURE

3.1 Stimuli and Apparatus

Our stimuli were 3D scenes designed and presented in Unity3D. For this experiment, we created a 3D room, furnished, and enclosed on all sides by walls and a ceiling. We placed a box on one of two identical cubical tables positioned side-by-side, simulating pedestals for the box displacement task. The sample scene can be seen in Figure 1. All 3D scenes were created and rendered in real-time on a 64-bit Windows 10 desktop computer with Intel Core i7 6700 CPU, and 16GB RAM. The visual environments were designed in Unity3D and programmed in C# scripts, and were presented on a Dual OLED 3.5-inch diagonal 3D HTC Vive Pro Eye headset display, with a resolution of 1440H * 1600V pixels per eye (2880 x 1600 pixels combined). The headset was connected to the desktop computer through a USB and DisplayPort video connection. The users' eye movements were recorded with a Tobii eye tracker, which was integrated inside the headset. The sampling frequency of the eye tracker was 120 Hz.

3.2 Experiment Design and Procedure

Unity3D and C# scripts were used for the design and execution of this experiment with support of the several utility packages. These were mainly SteamVR and the eye tracking API provided by Tobii eye tracking (SRanipal), supported by some other imported packages including SteamVR, SteamVRInput, SteamVRResources, TobiiXR and ViveSR.

Figure 1: Experiment Stimuli: The 3D scene used in all blocks of this experiment.

To compare how focusing on a task, such as moving an object, affects how we perceive changes applied to a scene we compared two interaction conditions that took place in the scene described above. The scene started with the box on the right pillar. During the trial this box was moved from one pillar to the other. The difference between the two conditions was how the box on the pillar was moved: it was either (a) displaced by the subjects' physical hand by grabbing it using the trigger button on a controller, and then moving it with the controller (*hand-with gaze* condition), or (b) it would be moved automatically through scripts (*gaze-only* condition). Importantly, in the gaze-only condition participants performed no direct object manipulation or gaze-contingent selection; they passively observed the box moving along the trajectory recorded from their own prior hand-with-gaze block. Thus, the two conditions were visually identical, but differed in the degree of sensory and motor engagement required.

Each session took approximately one hour and consisted of four blocks of trials, two blocks for each of the interaction conditions. We used an A-B-B-A block order, where "A" represents the hand-with-gaze condition and "B" the gaze-only condition. This design was chosen for two reasons: first, to reuse participant-specific motion trajectories captured during the first hand-with-gaze block in the subsequent gaze-only blocks, ensuring that any differences in detection performance between conditions could not be attributed to differences in object movement speed or path; and second, to balance potential order effects by symmetrically structuring exposure to each condition. While full counterbalancing was not feasible due to this trajectory dependency, the ABBA design mitigates linear learning trends and ensures each participant experienced both conditions twice in an alternating order.

Upon arrival, each participant was given background information and instructions about the study. Participants signed a written consent form approved by York University's ethics committee. At the start, we ran a couple of trials of the experiment to familiarize the participant until we made sure that they understood the experiment procedure and the task. These trials were excluded from our data analysis. Once they were familiarized with the procedure, we started the first block. For the duration of the experiment, the users were seated in a chair within the Vive's tracking space to keep the environment constant for all participants. We cleaned the headgear using sanitizing wipes after each participant. The setup of the participants can be seen in Figure 2.

Figure 2: A participant, seated on a chair with no wheels, completing the experiment.

3.2.1 Hand with Gaze Movement Blocks. In this condition, the participants used the controller to move the box from the green pillar on the right to the blue pillar on the left. They touched the box with the controller, and pressed the trigger button, shown in Figure 3, to pickup the object. They held onto the trigger button to maintain their grasp, and released the trigger button when they wanted to set the box on the blue pillar on the left. As the target object (the box) was being moved, its 3D position and orientation obtained from the Vive tracker were recorded in a file. They were used later (in the gaze-only block) for object displacement.

3.2.2 Gaze Only Blocks. In this experimental block, during each trial the participants only watched the recorded displacement of the box from the right pillar to the left pillar. During these trials, there was no hand interaction of the subjects with the scene objects. The discrimination tasks were identical to those in the hand blocks. They responded to two tasks, one about the direction of the scene shift and the other about the shape that appeared on the box while it was being moved. The shape task was demanding as to require the participants to track the block carefully with their eyes.

3.2.3 Visual Tasks. While the subjects were moving the box (or watching it move), they engaged in two simultaneous visually-demanding tasks:

(1) Scene shift discrimination. During the movement of the box there was an instantaneous (across one frame) scene rotation, at a random time, in one of the directions of pitch up, pitch down, yaw right or yaw left, with an amplitude of 0.75°, 1.5°, or 2.25°. The rotation amplitudes used in our study were selected based on perceptual thresholds established in prior redirected walking and scene manipulation research. Prior work has demonstrated that discrete scene manipulations introduced during visual suppression events (e.g., blinks or saccades) can go unnoticed when their magnitude is small. For instance, Bolte and Lappe [3] and Langbehn et al.[22]

Figure 3: Frame sequences of a sample trial in the experiment block, starting from the top right where the participant reaches for the yellow box, and ending with its displacement onto the left pedestal.

found that orientation changes of up to 2–5° can remain imperceptible when applied during a blink. Similarly, redirected walking research has identified rotational gains under 1.3× (equivalent to subtle direction shifts) as remaining below detection thresholds during locomotion[40].
Additionally, research into saccadic suppression of image displacement (SSID) [5] supports the notion that the visual system is relatively insensitive to small instantaneous shifts in visual input when they occur during eye movements. Using brief single-frame rotations, our design conceals the manipulations from user awareness while enabling precise testing of perceptual sensitivity across a near-threshold range. The 2AFC paradigm provides an unbiased measure of sensitivity by anchoring performance to a known chance level (50%), consistent with signal detection theory [13] and its established application in perceptual threshold research. This grounds our sensitivity measurements in a long tradition of forced-choice psychophysics while remaining consistent with perceptual literature on blink-masked and saccade-masked manipulations.
Participants were shown two options (e.g., "Left" vs. "Right", or "Up" vs. "Down") depending on the trial. These options were chosen based on the ground truth shift direction. For instance, if the scene rotated leftward, participants had to choose between "Left" and a randomly chosen incorrect direction.

(2) Shape identification. To make sure the participants kept their eyes on the target box while moving it, we briefly displayed a pattern on the box at some random time before or after the scene shift. The pattern was one of the four shapes shown in Figure 4. The pattern would appear on the box, remain for 100 ms, and then disappear. Participants were also asked to report this shape at the end of each trial. This was also a 2AFC task, with two buttons, one that showed the correct shape as well another one from the remaining three shapes.

When the participants selected a choice for both tasks, they proceeded to the next trial.

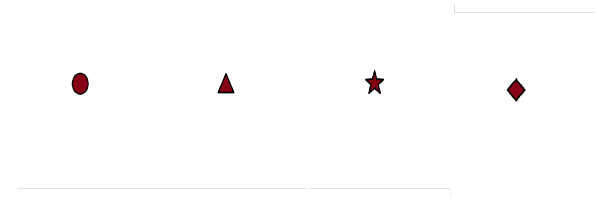

Figure 4: Shapes that appeared on the box while it was displaced.

3.2.4 Block Design and Procedure. The rest of the procedure was the same for both types of block: Each trial took 2300 ms, but subjects could take as long as needed to make their responses. The times taken to respond to the tasks were recorded for each trial. As explained, we ran the study in four blocks for each subject. Because of the design of the experiment, a block with the hand movement (box displacement by user) had to be completed first, followed by a gaze only block. Then we ran a gaze only block again (with the box displacement data of the first block), and then finished with another hand movement block.

Each block started with a calibration procedure and IPD setting, using Tobii's SDK for the HTC Vive Pro Eye headset. In this experiment, the calibration was performed successfully for all participants, except one. This participant removed their glasses, confirmed they could see clearly, and the calibration completed successfully. After calibration the first trial started. The trials in each block consisted of 12 different conditions that were presented to the participants in random order. The conditions included all combinations of 3 sizes of scene rotation by 4 directions of scene rotation. Every unique condition was randomly repeated 10 times for each participant, and therefore we had 120 trials in both the hand and gaze conditions for each subject across four blocks of 60 trials each.

3.3 Data Collection

The data collected included the responses of the participant for the two 2AFC tasks for the scene shift and the pattern on the box for each trial. We recorded the type and timing of scene shifts and target patterns, as well as participants' response times. During each trial we also continuously recorded the participant's eye movements, head pose (rotation and translation coordinates), and the hand controller pose on every frame.

3.4 Subjects

Eight participants completed the experiment, four female and four male, with an average age of 28.63 years (range [24-33]). Five participants had normal vision, and the rest had vision corrected to normal; none wore or needed glasses during the experiment. Five participants had prior experience with VR and were familiar with using a handheld controller and VR headset. All participants were graduate students at York University.

Although COVID-19 protocols constrained our sample to $N = 8$ participants, this is consistent with established practice in VR perceptual threshold studies [3, 29]. Critically, statistical power in this design derives primarily from trial-level observations rather than participant count alone, consistent with psychophysical methodology. A formal sensitivity analysis is reported in the Results Section.

4 RESULTS

We used MATLAB and Python for data processing. For scene shift detection, we computed accuracy rates across rotation magnitudes and directions. Given the 2AFC task, chance level was 50%. Chi-square tests and McNemar's test were applied to assess differences between interaction types, which perfectly captures overall performance.

In the gaze-only block, scene shifts occurred, on average, 990 ms into the trial, with a standard deviation of 56 ms, and with the shape appearing on the box at 1058±48 ms. For trials involving hand movements, the average timing of the scene rotations was 996±46 ms, and the shape appeared on the box at 1068±33 ms from the start of the trial.

4.0.1 Trials with success in the shape identification task. Participants were instructed to maintain their gaze on the box throughout all trials, as this was essential for successfully completing the shape task. The overall accuracy for the shape identification task was 92.60% during the gaze-only blocks and 91.67% during the blocks involving both gaze and hand movement. Figure 5 presents the percentage of correct shape identifications, as well as the percentage of correctly identified scene shifts, both with and without corresponding correct shape identifications.

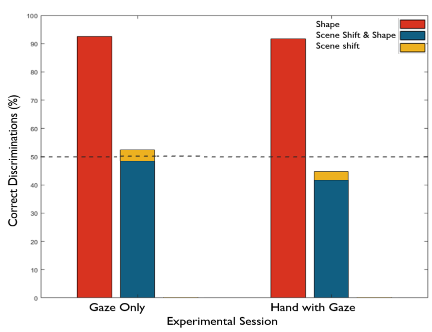

Figure 5: Percentage of correct identifications for shape on the box (red), scene shifts and shape on the box (blue), scene shifts only (yellow), during both experimental blocks. The dashed line represents chance (guessing) performance on the 2AFC tasks.

As this was a dual-task study, we also examined trials where participants correctly responded to both tasks simultaneously, indicating that they were focused on the scene shift task while maintaining their gaze and attention on the box. From Figure 5 we can see that for most of the trials where the scene shift was identified correctly, the shape was also identified correctly. This is expected as the participants performed well on the shape task in general and this did not appear to depend on the shift task. In contrast, on average participants performed poorly on the scene shift discrimination task with overall performance near chance.

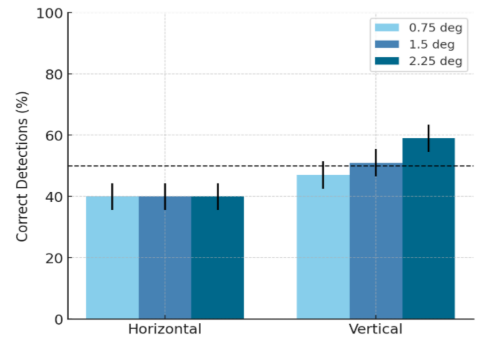

Figure 6: Detection accuracy for the hand-with-gaze condition. Percentage of correct scene shift discriminations across shift sizes and directions.

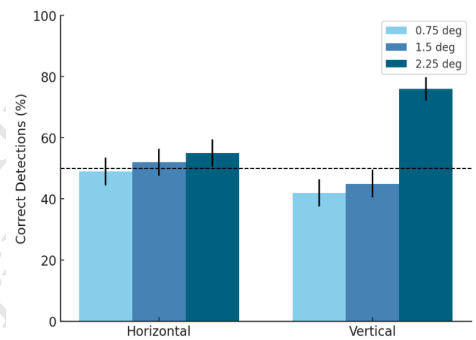

Figure 7: Detection accuracy for the gaze-only condition. Percentage of correct scene shift discriminations across shift sizes and directions.

4.0.2 Effect of interaction condition. In the hand-with-gaze multisensory engagement trials, participants used a controller and their physical hand to move a virtual object while keeping their eyes on it. The proportion of correct responses during horizontal scene shifts did not show significant differences across the three scene shift sizes. Figure 6 illustrates the percentage correct discriminations for hand-with-gaze trials. It can be seen that the participants were unable to correctly discriminate the direction of horizontal shifts as they were slightly below and not significantly different from chance level of 50%.

In the gaze-only trials, participants simply observed the box moving from the start to the end position. The box's motion followed the recorded positions and rotations from when the subject had manually moved the virtual box in a previous trial. Figure 7 presents the results for these gaze-only trials, where participants tracked the moving object with their gaze. The performance pattern was similar to that of the hand-with-gaze trials: horizontal shifts and small vertical shifts were not significantly different from chance. However, larger scene shifts (2.25°) were significantly easier to discriminate compared to smaller shifts of 1.5° and 0.75° ($\chi^2(2) = 4.46$, $p < 0.05$; Cramér's V = 0.32), indicating a medium effect size.

We then analyzed how the proportion of correct scene shift discriminations varied between the gaze-only and hand-with-gaze movement blocks. For horizontal scene rotations, the proportion of correct responses differed significantly between the hand-with-gaze and gaze-only conditions ($p < 0.01$, repeated-measures McNemar test). Participants responded "Right" more frequently during gaze-with-hand movement than gaze-only conditions. However, no significant differences were found for vertical rotations. We also compared the correct responses for different magnitudes of scene rotations between the hand-with-gaze and gaze-only conditions. The McNemar test revealed no significant differences between the interaction conditions for either the 0.75° or 1.5° scene rotations. However, for the larger 2.25° scene rotations, participants made significantly more correct discriminations during the gaze-only condition ($\chi^2(1) = 6.88$, $p < 0.01$) compared to the hand-with-gaze condition. A sensitivity analysis based on this McNemar test ($\chi^2(1) = 6.88, df = 1, \alpha = .05$) on 960 paired trial-level observations yields an achieved power of 0.75, sufficient to detect the primary effect of interaction modality. This confirms that despite the limited participant count, the trial-level design provides statistical sensitivity equivalent to much larger between-subjects studies.

4.0.3 Trials with failure in the shape identification task. Across participants, shape discrimination was highly accurate. On trials where participants did not correctly identify the shape (on average 8% of trials across participant), scene shift discrimination remained above chance: the mean percentage of correct responses was 63% for gaze-only trials and 47% for hand-with-gaze trials. These values were computed by averaging each participant's performance in trials with incorrect shape discrimination, ensuring that variability across participants was accounted for. Figure 8, shows the correct discriminations for each size of scene rotation, for the two conditions. We found the rate of correct discriminations in this case to be significantly different between the gaze-only condition and the hand-with-gaze condition ($\chi^2(1)$=6.88, p<0.05). The pattern in this subset followed the general pattern as seen for the full data set in Figures 6 and 7 with hand-with-gaze discrimination near chance levels and significantly better for large shifts under gaze-only conditions.

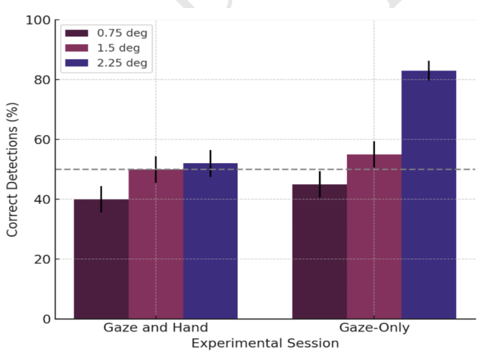

Figure 8: Percentage of correct scene shift discriminations during trials in which the shape on the box was not identified correctly (Failure in shape on the box identification).

5 CONCLUSIONS AND DISCUSSION

Our results demonstrate that multisensory task engagement reduces sensitivity to rapid background scene changes in VR, with participants showing near-chance detection performance regardless of interaction modality. When only gaze was involved, participants were significantly better at detecting larger scene shifts than those using both hand and gaze, suggesting that the additional cognitive and motor demands of manual interaction further suppress perceptual awareness of environmental changes.

Results further show that when users were engaged in a primary task, their sensitivity to rapid viewpoint changes in VR was reduced, regardless of whether they used gaze alone or both gaze and hand. Users were more likely to detect larger shifts of the scene that occurred orthogonally to the direction of the box's movement when they relied only on vision, which aligns with findings of saccadic suppression [3, 18, 19, 21]. This suggests that participants were able to detect salient changes in the 3D scene, consistent with previous research on attentional focus and reorientation in dynamic environments [32]. When both hand and gaze were involved in object tracking, participants became less sensitive to image rotations compared to when they only used their gaze. Note that gaze shifts were conspicuous when there was no attention task, and prior work has reported discrimination rates of about 90% for horizontal and vertical 1.0° shifts in the background while following a translating object [17, 19]. This relatively robust visibility of scene shifts, compared to the near-chance performance in this experiment with even larger 1.5° shifts, suggests that attention to the shape task affected the ability to notice the normally salient scene shifts. Participants in the hand-with-gaze (active manipulation) condition were required to integrate both haptic (hand) and visual (gaze) inputs to complete the task, engaging in a multisensory processing condition that likely increased cognitive load. This demand on multisensory integration may have resulted in reduced sensitivity to changes in the background environment, as attention was primarily focused on the task-relevant hand-eye coordination. In contrast, participants in the gaze only (passive observation) condition engaged only their visual sense, a task that imposed a lower cognitive load, allowing for greater attentional resources to detect subtle changes in the scene, such as rotations [10, 11].

These findings suggest that interaction modality (gaze vs. hand + gaze) modulates attentional resources differently, and that VR designers should tailor system feedback based on expected user engagement and sensory load.

These findings have important implications for various VR applications, such as gaming, rehabilitation, interaction design, and object manipulation [24, 30, 38]. In scenarios that require multitasking, users may not notice smaller background changes, especially when their attention is focused on tasks that require significant sensory or cognitive engagement, such as hand interactions. This suggests design options for using subtle reorientations or viewpoint adjustments in VR environments without disturbing users. Conversely, large changes in the scene were more likely to be noticed, particularly when users are not actively engaged in multitasking, suggesting that designers must consider the size and nature of scene alterations to maintain immersion without causing disorientation.

Our study showed that participants frequently missed image transitions while performing a challenging shape identification task, whether using one or two senses. Scene rotations were easier to detect during gaze-only trials compared to hand-and-gaze trials, but even the largest rotations went unnoticed in some cases. This reinforces the idea that when users are deeply focused on a demanding task, their sensitivity to background changes decreases, a phenomenon also observed in earlier work by Simons and Jensen (2009), who reported reduced detection of unexpected objects during more difficult tasks [37]. Our findings support this: task difficulty plays a crucial role in inattentional blindness, while others suggest that individual differences in attentional capacity may not significantly affect detection rates [36].

Interestingly, failure in the primary task (shape identification) did not guarantee success in detecting scene shifts, and vice versa. This suggests that errors in one task were not necessarily due to participants reallocating attention to the other. Instead, it may be that both tasks competed for attention, leading to occasional failures in both domains. The timing of events, such as the appearance of the shape or the occurrence of image shifts, could have further compounded this challenge, as participants were not always able to anticipate when critical events would occur, causing lapses in attention. These findings are also consistent with theories of visual suppression during blinks or saccades, which could have contributed to missed detections. Recent research has explored the potential of using cognitive or physical tasks to mask user redirection in VR. For example, Coelho et al. (2022) found that user redirection was less detectable during physically demanding tasks, although other studies suggested that moderate exercise can improve cognitive performance and reduce inattentional blindness, up to a certain intensity threshold [8, 9, 14]. Similarly, Schmelter et al. (2021) demonstrated that interactive tasks such as 'focus, pick up, throw, shoot, and fight' could be used as distractors to implement discrete manipulations during redirected walking [33]. These findings highlight the possibility of subtly guiding VR users by engaging them in attention-demanding tasks, offering new directions for designing immersive, distraction-tolerant virtual environments.

This study has several limitations worth acknowledging. The sample was constrained to $N = 8$ participants due to COVID-19 recruitment restrictions; however, as discussed in Subjects Section, statistical power in this design derives primarily from trial-level observations consistent with psychophysical methodology, and a formal sensitivity analysis confirmed adequate power for the primary comparison. Additionally, all participants were graduate students at York University, which limits demographic diversity. The use of a seated, controller-based setup also bounds the generalizability of the findings, as discussed in the future work directions below.

Future work could examine whether the inattentional blindness effects observed here generalize to controller-less hand tracking setups (e.g., pinch-to-grasp), where proprioceptive and attentional demands may differ from controller-based interaction. Additionally, as this study was conducted with seated participants, it remains an open question whether these findings transfer to standing or locomoting users, where vestibular and whole-body movement signals may further modulate sensitivity to background scene changes, particularly in redirected walking applications.

In conclusion, this study contributes to our understanding of how multisensory engagement affects attentional focus and perception in VR. By identifying the conditions under which users are more or less likely to notice changes in the environment, we can better inform the design of VR systems that balance immersion with user awareness, ensuring that critical environmental cues are effectively communicated without overwhelming the user. Future research could explore further how different types of sensory input or task difficulty levels interact to influence user perception and attentional changes in VR.

REFERENCES

[1] Monthir Ali and Rogelio E Cardona-Rivera. 2020. Comparing gamepad and naturally-mapped controller effects on perceived virtual reality experiences. In *ACM Symposium on Applied Perception 2020*. 1–10. https://doi.org/10.1145/3385955.3407923

[2] Malika Auvray, Alberto Gallace, Hong Z Tan, and Charles Spence. 2007. Crossmodal change blindness between vision and touch. *Acta Psychologica* 126, 2 (2007), 79–97. https://doi.org/10.1016/j.actpsy.2006.10.005

[3] Benjamin Bolte and Markus Lappe. 2015. Subliminal reorientation and repositioning in immersive virtual environments using saccadic suppression. *IEEE transactions on visualization and computer graphics* 21, 4 (2015), 545–552. https://doi.org/10.1109/TVCG.2015.2391851

[4] Doug A Bowman and Larry F Hodges. 1997. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*. 35–ff. https://doi.org/10.1145/253284.253301

[5] Bruce Bridgeman, David Hendry, and Lawrence Stark. 1975. Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research* 15, 6 (1975), 719–722. https://doi.org/10.1016/0042-6989(75)90290-4

[6] Eric Burns, Sharif Razzaque, Abigail T Panter, Mary C Whitton, Matthew R McCallus, and Frederick P Brooks. 2005. The hand is slower than the eye: A quantitative exploration of visual dominance over proprioception. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* IEEE, 3–10. https://doi.org/10.1109/vr.2005.77

[7] Ula Cartwright-Finch and Nilli Lavie. 2007. The role of perceptual load in inattentional blindness. *Cognition* 102, 3 (2007), 321–340. https://doi.org/10.1016/j.cognition.2006.01.002

[8] Daniel Neves Coelho, Frank Steinicke, and Eike Langbehn. 2022. Design of Mentally and Physically Demanding Tasks as Distractors of Rotation Gains. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 620–621. https://doi.org/10.1109/vrw55335.2022.00161

[9] Karen Davranche and Terry McMorris. 2009. Specific effects of acute moderate exercise on cognitive control. *Brain and cognition* 69, 3 (2009), 565–570. https://doi.org/10.1016/j.bandc.2008.12.001

[10] Alyanne M de Haan, Miranda Smit, Stefan Van der Stigchel, and H Chris Dijkerman. 2016. Approaching threat modulates visuotactile interactions in peripersonal space. *Experimental brain research* 234 (2016), 1875–1884. https://doi.org/10.1007/s00221-016-4571-2

[11] Marc O Ernst and Heinrich H Bülthoff. 2004. Merging the senses into a robust percept. *Trends in cognitive sciences* 8, 4 (2004), 162–169. https://doi.org/10.1016/j.tics.2004.02.002

[12] Yi Gao, Kai Xue, Brian Odegaard, and Dobromir Rahnev. 2025. Automatic multisensory integration follows subjective confidence rather than objective performance. *Communications psychology* 3, 1 (2025), 38. https://doi.org/10.1101/2023.06.07.544029

[13] David Marvin Green, John A Swets, et al. 1966. *Signal detection theory and psychophysics*. Vol. 1. Wiley New York.

[14] Stefanie Hüttermann and Daniel Memmert. 2012. Moderate movement, more vision: Effects of physical exercise on inattentional blindness. *Perception* 41, 8 (2012), 963–975. https://doi.org/10.1068/p7294

[15] Melinda S Jensen, Richard Yao, Whitney N Street, and Daniel J Simons. 2011. Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science* 2, 5 (2011), 529–546. https://doi.org/10.1002/wcs.130

[16] Yashas Joshi and Charalambos Poullis. 2020. Inattentional Blindness for Redirected Walking Using Dynamic Foveated Rendering. *IEEE Access* 8 (2020), 39013–39024. https://doi.org/10.1109/access.2020.2975032

[17] Maryam Keyvanara. 2022. Perceptual Manipulations for Hiding Image Transformations in Virtual Reality. (2022).

[18] Maryam Keyvanara and Robert Allison. 2019. Transsaccadic awareness of scene transformations in a 3d virtual environment. In *ACM Symposium on Applied Perception 2019*. 1–9. https://doi.org/10.1145/3343036.334312

[19] Maryam Keyvanara and Robert Allison. 2020. Effect of a constant camera rotation on the visibility of transsaccadic camera shifts. In *ACM Symposium on Eye Tracking Research and Applications*. 1–8. https://doi.org/10.1145/3379155.3391318

[20] Maryam Khademi, Hossein Mousavi Hondori, Alison McKenzie, Lucy Dodakian, Cristina Videira Lopes, and Steven C Cramer. 2014. Free-hand interaction with leap motion controller for stroke rehabilitation. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 1663–1668. https://doi.org/10.1145/2559206.2581203

[21] Eike Langbehn, Gerd Bruder, and Frank Steinicke. 2016. Subliminal reorientation and repositioning in virtual reality during eye blinks. In *Proceedings of the 2016 symposium on spatial user interaction*. 213–213. https://doi.org/10.1145/2983310.2989204

[22] Eike Langbehn, Frank Steinicke, Markus Lappe, Gregory F Welch, and Gerd Bruder. 2018. In the blink of an eye: leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–11. https://doi.org/10.1145/3197517.3201335

[23] Jaeyeon Lee, Mike Sinclair, Mar Gonzalez-Franco, Eyal Ofek, and Christian Holz. 2019. TORC: A virtual reality controller for in-hand high-dexterity finger interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[24] Po-Wei Lee, Han-Yu Wang, Ying-Chao Tung, Jhe-Wei Lin, and Andries Valstar. 2015. TranSection: hand-based interaction for playing a game within a virtual reality game. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 73–76.

[25] Arien Mack. 2003. Inattentional blindness: Looking without seeing. *Current directions in psychological science* 12, 5 (2003), 180–184.

[26] Arien Mack and Irvin Rock. 1998. Inattentional blindness: Perception without attention. *Visual attention* 8 (1998), 55–76.

[27] Sandra Malpica, Ana Serrano, Diego Gutierrez, and Belen Masia. 2020. Auditory stimuli degrade visual performance in virtual reality. *Scientific reports* 10, 1 (2020), 1–9. https://doi.org/10.1038/s41598-020-69135-3

[28] Daniel Martin, Xin Sun, Diego Gutierrez, and Belen Masia. 2023. A study of change blindness in immersive environments. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2446–2455. https://doi.org/10.1109/tvcg.2023.3247102

[29] Sebastian Marwecki, Andrew D Wilson, Eyal Ofek, Mar Gonzalez Franco, and Christian Holz. 2019. 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*. 777–789. https://doi.org/10.1145/3332165.3347919

[30] Margarida F Pereira, Cosima Prahm, Jonas Kolbenschlag, Eva Oliveira, and Nuno F Rodrigues. 2020. Application of AR and VR in hand rehabilitation: A systematic review. *Journal of Biomedical Informatics* 111 (2020), 103584. https://doi.org/10.1016/j.jbi.2020.103584

[31] Uwe Proske and Simon C Gandevia. 2012. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews* (2012). https://doi.org/10.1152/physrev.00048.2011

[32] Anne Richards, Emily M Hannon, and Melanie Vitkovitch. 2012. Distracted by distractors: Eye movements in a dynamic inattentional blindness task. *Consciousness and cognition* 21, 1 (2012), 170–176. https://doi.org/10.1016/j.concog.2011.09.013

[33] Thereza Schmelter, Levente Hernadi, Marc Aurel Störmer, Frank Steinicke, and Kristian Hildebrand. 2021. Interaction based redirected walking. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 4, 1 (2021), 1–16.

[34] Benjamin Schöne, Rebecca Sophia Sylvester, Elise Leila Radtke, and Thomas Gruber. 2021. Sustained inattentional blindness in virtual reality and under conventional laboratory conditions. *Virtual Reality* 25, 1 (2021), 209–216. https://doi.org/10.1007/s10055-020-00450-w

[35] Daniel J Simons and Christopher F Chabris. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *perception* 28, 9 (1999), 1059–1074. https://doi.org/10.1068/p281059

[36] Daniel J Simons, Connor M Hults, and Yifan Ding. 2024. Individual differences in inattentional blindness. *Psychonomic Bulletin & Review* (2024), 1–32. https://doi.org/10.3758/s13423-023-02431-x

[37] Daniel J Simons and Melinda S Jensen. 2009. The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review* 16, 2 (2009), 398–403. https://doi.org/10.1167/8.6.459

[38] Avinash K Singh, Klaus Gramann, Hsiang-Ting Chen, and Chin-Teng Lin. 2021. The impact of hand movement velocity on cognitive conflict processing in a 3D object selection task in virtual reality. *NeuroImage* 226 (2021), 117578. https://doi.org/10.1016/j.neuroimage.2020.117578

[39] Frank Steinicke, Gerd Bruder, Klaus Hinrichs, and Pete Willemsen. 2010. Change blindness phenomena for stereoscopic projection systems. In *2010 IEEE Virtual Reality Conference (VR)*. IEEE, 187–194. https://doi.org/10.1109/vr.2010.5444790

[40] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2010. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2010), 17–27. https://doi.org/10.1109/TVCG.2009.62

[41] Evan A Suma, Seth Clark, David Krum, Samantha Finkelstein, Mark Bolas, and Zachary Warte. 2011. Leveraging change blindness for redirection in virtual environments. In *2011 IEEE Virtual Reality Conference*. IEEE, 159–166. https://doi.org/10.1109/vr.2011.5759455

[42] Qi Sun, Anjul Patney, and Frank Steinicke. 2020. Redirected walking in vr. *Real VR–Immersive Digital Reality: How to Import the Real World into Head-Mounted Immersive Displays* (2020), 285–292. https://doi.org/10.1007/978-3-030-41816-8_12

[43] Jochen Triesch, Brian T Sullivan, Mary M Hayhoe, and Dana H Ballard. 2002. Saccade contingent updating in virtual reality. In *Proceedings of the 2002 symposium on eye tracking research & applications*. 95–102. https://doi.org/10.1145/507089.507092

[44] John C Tuthill and Eiman Azim. 2018. Proprioception. *Current Biology* 28, 5 (2018), R194–R203.

[45] Cindy V Vasquez-Caballero. 2020. inattentional Blindness: A Virtual Reality Experiment. (2020).

[46] Yuwei Wang, Yimin Wu, Cheng Chen, Bohan Wu, Shu Ma, Duming Wang, Hongting Li, and Zhen Yang. 2022. Inattentional blindness in augmented reality head-up display-assisted driving. *International Journal of Human–Computer Interaction* 38, 9 (2022), 837–850. https://doi.org/10.1080/10447318.2021.1970434

[47] Chyanna Wee, Kian Meng Yap, and Woan Ning Lim. 2021. Haptic interfaces for virtual reality: Challenges and research directions. *IEEE access* 9 (2021), 112145–112162. https://doi.org/10.1109/access.2021.3103598

[48] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240. https://doi.org/10.1162/105474698565686

[49] Alexander Vladimirovich Zakharov, Alexander Vladimirovich Kolsanov, Elena Viktorovna Khivintseva, Vasiliy Fedorovich Pyatin, and Alexander Vladimirovich Yashkov. 2021. Proprioception in Immersive Virtual Reality. In *Proprioception*. IntechOpen. https://doi.org/doi.org/10.5772/intechopen.96316

[50] André Zenner, Kora Persephone Regitz, and Antonio Krüger. 2021. Blink-suppressed hand redirection. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 75–84. https://doi.org/10.1109/vr50410.2021.00028