

Effectiveness of Visual, Auditory, and Haptic Guidance Cues for Visual Targets in Virtual Exhibition Environments

Hila Sabouni¹, Jack Miller¹, Stephen B. Gilbert¹, Chengde Wu²

¹VRAC Research Center, Iowa State University, Ames, IA 50011; ²Department of Architecture, Iowa State University, Ames, IA 50011

E-mail: hsabouni@iastate.edu

Abstract. A design challenge in virtual reality (VR) is balancing users' freedom to explore the virtual environment with the constraints of a guidance interface that focuses their attention without breaking the sense of immersion or encroaching on their freedom. In virtual exhibitions in which users may explore and engage with content freely, the design of guidance cues plays a critical role. This research explored the effectiveness of three different attention guidance cues in a scavenger-hunt-style multiple visual search task: an extended field of view through a rearview mirror (passive guidance), audio alerts (active guidance), and haptic alerts (active guidance) as well as a fourth control condition with no guidance. Participants were tasked with visually searching for seven specific paintings in a virtual rendering of the Louvre Museum. Performance was evaluated through qualitative surveys and two quantitative metrics: the frequency with which users checked the task list of seven paintings and the total time to complete the task. The results indicated that haptic and audio cues were significantly more effective at reducing the frequency of checking the task list when compared to the control condition while the rearview mirror was the least effective. Unexpectedly, none of the cues significantly reduced the task-completion time. The insights from this research provide VR designers with guidelines for constructing more responsive virtual exhibitions using seamless attentional guidance systems that enhance user experience and interaction in VR environments.

Keywords: virtual reality, guidance, attention orientation, audio, haptic, user interface, virtual exhibition, informal learning
This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.
[DOI: 10.2352/J.Percept.Imaging.2025.8.000403]

1. INTRODUCTION

The virtual reality (VR) technology market has been rapidly expanding with a projected global growth from \$166.7 billion in 2022 to \$2273.4 billion in 2029 [33]. Virtual reality has emerged as a significant asset across a diverse range of domains, including research [13, 43], healthcare [2, 28, 48], training [24, 32], education [17, 25], and entertainment [22, 23] as it offers a richer, more immersive, and interactive experience. However, a notable challenge in virtual environments (VEs) is user guidance. Where should the user move next? Where should the user look next? Is there an important alert to attend to? What if it is behind the user? Effective guidance mechanisms are essential for user experience in applications

ranging from medical training to interactive educational experiences such as exhibitions. Without effective guidance, users may experience frustration, reduced engagement, or insufficient navigation, which limits the utility of VEs in critical and demanding scenarios.

This research focuses on informal educational environments, particularly virtual exhibition spaces such as museums, which enable people to visit them without travel. The design of these experiences poses an interesting human-computer interaction (HCI) challenge to balance users' freedom with the agenda of the exhibit designer. Users must be able to explore on their own and engage with exhibits that interest them while at the same time, the exhibit designers may have a goal; for example, "I want patrons to learn how important water is to their daily lives and several ways to save water" or "I want patrons to appreciate the dramatic change between Romantic artists and Realist artists and Realism's influence on social class structures." Previous research on guidance systems in real museums has demonstrated that most users value their freedom and resist following a set path [16, 36]. An exhibit will often evoke emotion and even a strong sense of presence or flow [40] and as such, it is important for the user interface to be as simple and unobtrusive as possible so that it does not interrupt that experience [10]. This goal overlaps with some of Weiser's goals of ubiquitous computing [64], which include providing invisible, seamless, or natural interfaces for systems that surround us. It also overlaps with the concept of digital nudging [63]—designing a system that offers users a choice while strongly favoring that they make certain choices over others. The virtual exhibition designer may want to give patrons freedom while nudging them toward certain goals. While visitors in a real exhibition will likely augment their experience with at most their phone and headphones, virtual exhibitions offer many more options for these seamless guidance interfaces. Thus, this research seeks to answer two research questions: What are the potentially suitable types of seamless guidance cues (visual, auditory, and haptic) for virtual exhibitions? How do different types of guidance cues impact users' ability to locate visual targets in a virtual exhibition?

Multiple types of guidance exist in VEs, which include the following: (1) direction of what the user should do next [34, 66], (2) directing the user to a particular locomotion path [14, 60, 62], and (3) directing the user to attend to

Received July 30, 2024; accepted for publication Apr. 2, 2025; published online July 03, 2025. Associate Editor: Rafal Mantiuk.

2575-8144/2025/8/000403/11/\$00.00

specific visual targets [7, 49]. This research focuses on the last category. Unlike more typical visual guidance in the form of alerts (e.g., in a nuclear power plant control room, an airline cockpit, or a hospital surgery), where compliance with the alert guidance is highly consequential, guidance in virtual exhibitions must remain optional and almost invisible. As noted by previous researchers, providing guidance to specific visual targets without constraining the user can pose a challenge in balancing the salience of the guidance cue with the presentation of a freely browsable virtual environment [49].

Previous guidance systems for multiple visual search tasks span a range of sensory modalities, typically visual, audio, and haptic. Although significant research has been conducted to explore the potential of these cues across modalities, each study has its limitations. For example, Felix et al. primarily altered the field of view to explore visual guidance but neglected the rear space [7]. Similarly, although some research has investigated the use of mirrors for rearview guidance, the applications remain limited. Ha et al. employed a mirror to extend the field of view for assembly tasks [20], focusing on a very specific and close space in the front view. Boonsuk et al. used a mirror to detect ten identical red barrels in the rear space [6], where uniform targets with the distinct color help make them easier to locate. Additionally, many studies focused narrowly on one sort of guidance cue. For example, Rebai et al. explored visual guidance cues with a focus on rear space but omitted auditory or haptic comparisons [53]. Some studies focused solely on active cues and overlooked the potential of passive aids [11, 38, 54, 68]. Moreover, several lacked a control condition where no cues were provided, making it difficult to evaluate the inherent challenges of the task or the environment itself [54, 68].

To address these gaps, the present study compares a control condition, a passive visual aid, active audio alerts, and active haptic alerts to evaluate their effectiveness in guiding users to a collection of specific visual targets. This study uses a scavenger hunt task to simulate real-world search scenarios, which offers a unique approach that balances user autonomy with the effectiveness of cue salience.

For this study, participants were tasked with finding seven specific paintings within a virtual environment of the remodeled Louvre Museum. Their performance was measured by (1) how often they referred to the task list of seven to assess how confident they felt with the received cue to finish tasks while navigating and (2) the total time taken to complete the task to analyze the extent to which cues helped participants quickly find target paintings among others. The primary hypothesis of this study is that adding visual, auditory, and haptic cues can help users better locate visual targets in a virtual space compared to solely relying on the front view. In addition, to gain insight into the relationship between the three guidance systems and a quality experience in the VE, other metrics were evaluated, including the level of workload, sense of presence, and cybersickness of users. Moreover, users' satisfaction with the guidance cues' efficiency and effectiveness was evaluated through qualitative open-ended questions.

This study contributes to the field of HCI and VR research, particularly in the design of guidance interfaces for immersive environments. Specifically, the study identifies three potentially suitable types of seamless guidance in VR environments: passive visual aids, active audio alerts, and active haptic alerts. Furthermore, it demonstrates that haptic and audio cues are effective for designing guidance interfaces in immersive VR environments through a task-based experiment. The findings are particularly relevant to virtual museums, educational VR applications, and interactive simulations, offering practical guidance for improving user experience in informal learning environments.

2. BACKGROUND

Attention orientation is important for virtual tasks as it impacts the ability to process information, navigate environments, and perform tasks efficiently [4]. It involves focusing on specific tasks or objects while disregarding irrelevant information [14], which makes it challenging to measure accurately [58]. Traditional methods, such as self-reports and observations, often capture partial data due to the dynamic nature of VR environments [61]. Previous VR research has explored a variety of methods to guide attention orientation [5, 11, 13, 35, 41, 52]. Typically, attention is measured by the total time users require to complete tasks or the number of right and wrong responses [3, 5, 12, 52, 68]. However, as highlighted in Carlos Lievano et al.'s comprehensive literature review, these methods need to account for factors such as individual differences among users, experimental design and settings, and the characteristics of extended reality (XR) technology in each study [58].

Sensory cues are a key aspect in shaping attention orientation in VEs [37]. These external stimuli, such as visual, auditory, haptic, olfaction, and gustation, can be used to guide users' attention toward specific points of interest. However, they play varying degrees of importance and are particularly critical to navigation and completing tasks due to the challenge of perceiving distance, motion, and direction in VR environments [13, 32]. Among sensory cues, visual, auditory, and haptics appear to have been studied the most [4, 11, 32, 37, 39, 43, 62].

Vision is the primary sensory stimulus humans rely on for interaction with the world [31]. Researchers have investigated the importance of vision in virtual navigation and performance through different perspectives. For instance, a study on the effect of different field of view angles (102°, 52°, and 32°) on a visual search task for training suggested that a higher field of view led to better training effectiveness and virtual performance [45]. A study evaluating gaze-aware visual cues on driving within a VR game showed that the cues improved driver attention and response to unexpected pedestrian crossings [8]. Another study compared the impact of visual cues on finding targets within versus outside the visual field. The results showed that cues significantly improved performance when targets that needed to be located occurred outside the visual field [53]. In addition, some studies introduced the concept of a rearview mirror in

VR. Ha et al. presented a virtual mirror to uncover hidden spatial information, enhancing precision in 3D manipulation. The mirror was automatically controlled and positioned where it could highlight the spatial relationship between the manipulated object and the other objects close to it. The results showed that the mirror was significantly helpful for users in manipulating 3D objects [20]. Another study compared user interfaces with a 90° rearview mirror to a 180° rearview mirror in a scavenger-hunt-style target location task. The results showed that although participants primarily used the front views to acquire targets, they did sometimes use the rearview mirrors [6]. These studies suggested that besides the default front view, a rearview mirror that extends a user's view of the world could be beneficial in a visual search task. Based on this research, the present study chose to include a rearview mirror as one of its guidance systems.

Audio can also play a significant role in attention orientation, particularly in VEs. Researchers have explored its impact in different scenarios to highlight its effectiveness in directing attention [11, 38, 57]. A study comparing participants' attention in both static and dynamic VEs with and without 3D sound cues found that audio cues enhanced participants' attention when scanning for targets [38]. Suzuki et al. introduced a novel method to study sound localization in VEs. The "active listening" method leverages dynamic spatial information generated by participants' movements to improve spatial hearing analysis [57]. The orientation of attention through audio localization has been further investigated using different audio modalities, such as mute, mono, and ambisonics [11]. Research found that salient object sounds had the greatest effect in directing one's attention toward the source [11]. Based on these studies on audio cues, the present study chose to include a simple audio proximity alert as one of its guidance systems.

Haptic cues, particularly vibrotactile feedback in VR, represent another influential cue in orienting attention. Vibration belts with directional cues could help prevent collisions with obstacles in an immersive virtual set [68]. Hong et al. explored the potential of vibrotactile simulations in redirecting visual spatial attention during a visual task. The findings suggested that when haptic cues are valid, they significantly improve response time, indicating the cues' impact on attention enhancement [59]. Another study demonstrated that short, repetitive vibrations could reduce cognitive load if users' gaze were guided toward nearby interactive objects to correct inaccuracies [54]. Localization of haptic cues has also been examined under two conditions with high and low validity. The results indicated that high validity of haptic cues oriented one's attention more naturally and intuitively than low validity [68]. Based on this research, the present study chose to include haptic vibrotactile proximity alert as one of its guidance systems.

A variety of studies have also focused on the combination of such cues [12, 28, 43, 47]. For instance, one study conducted experiments involving multiple simultaneous visual search tasks where visual, audio, and vibrotactile cues were applied. The findings indicated that on-head

vibrotactile cues effectively guided users' attention toward targets without causing any distraction in other simultaneous tasks when compared to audio and visual cues [12]. Another study focused on multisensory integration suggested that multisensory stimuli could be significantly helpful in target detection only if the environment and task required a high perceptual load [37]. Metrics such as task-completion time, identification accuracy, and range of head movement were monitored to assess these effects. The variability in findings suggests that cues may perform differently depending on scenarios and combinations of other cues. To keep the present study simpler and focus on the impact of each cue, the authors did not combine cues. However, the choices of metrics and dependent variables used in previous studies influenced the variables measured in the present study.

Several studies have explored the role of sensory cues, particularly in virtual museums. A mixed-method study of three Indonesian museums examined how virtual multisensory spaces are perceived, focusing on sensory systems as the key factor influencing respondents' experiences. The findings revealed that visual and auditory senses dominated virtual spatial experiences while chemical senses were weaker [15]. Similarly, Guo et al. identified the prominence of visual and auditory cues over haptic and taste cues in creating a holistic digital museum experience. Additionally, it demonstrated that emotional state and sense of presence mediate the impact of multisensory cues [19]. Another study introduced a multisensory virtual museum experience by integrating tactile feedback via 3D-printed artifacts and sensors, enhancing the sense of presence. Experimental results from 32 participants demonstrated that this multisensory approach significantly improved realism, immersion, and user preference compared to traditional audiovisual methods [27]. Although these studies highlighted the potential of sensory cues in creating an immersive virtual museum experience, they neglect to investigate the impacts of individual sensory cues on achieving a seamless experience. Moreover, they focus primarily on immersion and presence, overlooking the critical aspects of how sensory integration affects visitors' sense of freedom and engagement with the virtual environment, which are essential for a comprehensive experience.

The present study seeks to compare passive visual aids like the rearview mirror with active attentional guidance in the form of audio and haptic alerts in the context of a scavenger-hunt-style multiple visual search task (e.g., find seven objects in any sequence). Although extensive human factor research exists in the design of effective alerts (e.g., for medical conditions in a surgical room [39], airline cockpits [44], and nuclear power plant control rooms [69]), these alerts usually serve as alarms. In a scavenger-hunt-style task, where the goal is to allow the user to freely explore while more subtly alerting them to objects of interest, the fields of ubiquitous computing and ambient information systems are perhaps more relevant. These fields focus on methods of alerting users to new information in the background without distracting focus from a primary task [42]. Based on the four dimensions of ambient systems [42], the

authors chose simple alert systems for this initial study. The alerts had low information capacity (“a target is near”), medium notification level (a single sound or vibration), low representational fidelity (the alert did not resemble the target), and low aesthetic emphasis (the alerts were not integrated with the artistic or architectural design of the VE).

3. RESEARCH METHODS

This study uses a mixed-method approach aimed at uncovering the effectiveness of different cues in orienting attention and task completion in a virtual recreation of the Louvre Museum. The investigation combined quantitative assessments of task performance with qualitative analyses of participants' survey and questionnaire responses following the VR task, providing a comprehensive understanding by capturing both measurable outcomes and subjective experiences.

As illustrated in Figure 1, participants were first introduced to the task and the target paintings, followed by a practice session in a simple virtual room to familiarize themselves with the controls. They were then randomly assigned to one of four conditions, each featuring a distinct type of cue. Upon completing the VR experiment, participants filled out a questionnaire to provide feedback on their experience. The following sections present a detailed description of the task design, study design, study variables, and the data analysis approach.

3.1 Task Design

3.1.1 Target Introduction

Seven target paintings were introduced to the participants using a printed list, allowing them a moment to get acquainted with the target paintings. During this introduction, all participants were informed that they do not need to memorize the target paintings and they will be provided with a list of all targets in their left hand in the VE.

3.1.2 Practice Task

Following the target introduction, participants began with a three-minute practice session to familiarize themselves with VR headset use, controllers, and the task. This step is recommended to help participants learn the applications and ensure they are comfortable with the virtual environment [1]. The headset lens were adjusted for each participant to ensure an optimal view based on their interpupillary distance. Participants started the practice session in a virtual room called “Simple Box,” as displayed in Figure 2, with only the seven target paintings, received an assigned cue, and learned how to navigate and collect targets. Upon completion of the practice session (about three minutes), participants were asked whether they would like to try it again or whether they were ready to start the task. No data was collected in the practice session.

3.1.3 Task

3.1.3.1. Task Spatial Environment. The design and context of a virtual environment are critical choices in studies focusing on attention and guidance. In educational research, for

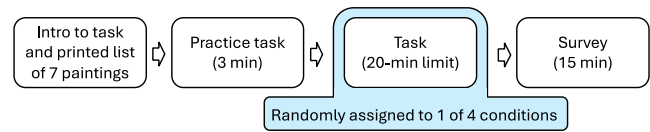


Figure 1. Procedure of VR experiments.



Figure 2. Practice session: the “Simple Box” containing only the seven target paintings.

example, classroom settings are often chosen as conventional environments due to familiarity [58]. Sriworapong et al. created a 3D model of a classroom model that allowed students and educators to engage, present, and discuss materials just as they would in a real-world classroom [56]. In this study, the Louvre Museum in Paris, France, was selected as the virtual exhibition environment to assess the impact of various sensory cues in a scavenger hunt task, mimicking real-world browsing and searching situations. The museum has three wings: Richelieu, Sully, and Denon. A small section of the Denon wing was selected for the virtual environment, rooms 700–702 and 710–712, as shown in Figure 3. These rooms were selected to ensure balance in the layout of the environment. In addition, approximately 30% of the exhibited paintings in the selected rooms were removed to create a focused and engaging experience for participants. These design choices helped limit the complexity of the task and reduce its duration, addressing key operational factors essential for studying attention in VR applications [1].

3.1.3.2. 3D Modeling. The selected rooms were modeled in Rhinoceros 3D. Iconic architectural features such as curved and flat ceilings, towering columns, and walls were modeled according to their current form. Intricate architectural and aesthetic details, however, were omitted in the model to restrict the details and complex artworks from interfering with participants' attention and orientation.

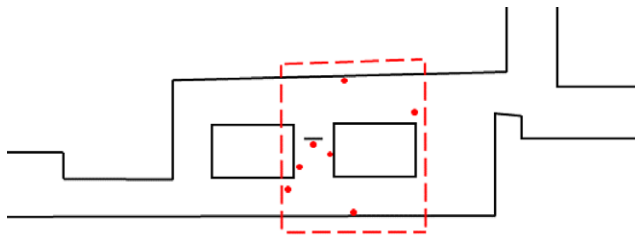


Figure 3. The Louvre Museum layout: the space marked with red dashed lines includes rooms 700–702 and 710–712, and the red dots are the targets. The width of the selected rooms is approximately 20 m.



Figure 4. Rearview mirror condition with the task list at the left of the controller and the rearview mirror at the right.

3.1.3.3. Equipment and Interactions. The partial 3D model of the museum was imported into Unity game engine, where the interactive elements of the task were programmed. Seven paintings were selected as targets from a collection of 48 paintings. They were spread throughout all rooms, each purposefully distanced from other targets (red dots shown in Fig. 3). Participants used a Meta Quest Pro to interact with the environment, utilizing ray interaction through controllers to select and collect the targets. To enhance the virtual experience, a visible counter was incorporated within their field of vision to display the number of paintings they had successfully collected (Figure 4). Participants across all conditions were equally provided with a list of all targets in their left hand, which included graphical images of the paintings (Fig. 4). This feature let participants refresh their memory of the visual targets whenever needed to reduce the effects of memory load.

3.2 Study Design

3.2.1 Participants

The experimental protocol was approved by the university's Institutional Review Board (IRB 23-092).

A recruitment email was sent to all university students. To ensure participant safety, eligibility criteria required individuals to be 18 years or older and to meet the following conditions: no history of photosensitive seizures, motion sickness, nausea, migraines, headaches, balance issues, dizziness, epilepsy, and neurological conditions that could be triggered by visual stimuli; no light sensitivity and uncorrected vision issues (contact lenses were permitted, but eyeglasses were not due to headset fit); no pre-existing

binocular vision abnormalities and psychiatric disorders; no recent medical procedures, including cosmetic treatments; no heart conditions and other serious medical conditions; no use of medical devices, such as cardiac pacemakers, hearing aids, and defibrillators, that could be affected by the VR system's magnets and radio waves; not experiencing fatigue, exhaustion, emotional stress, anxiety, headaches, nausea, dizziness, and lightheadedness; not being ill (including cold, flu, COVID-19, and flu-like symptoms); not being pregnant; and not being under the influence of drugs and alcohol or experiencing a hangover. Participants with or without prior VR experience were eligible to take part.

In total, the sample included 71 participants ranging in age from 18 to 35 with an average age of 26. Of these, 30 were female and 41 were male. The participants had a diverse range of gaming experience levels. Notably, two participants had previously visited the Louvre Museum, but they barely remembered the layout of the building. One explained, “We didn’t visit the entire building due to its large size, and I only remember the overall shape of the building from the outside.” The other participant stated, “That was a long time ago, and I don’t recall any specific details about the building but just a very general impression.”

3.2.2 Physical Experiment Setting

An enclosed physical laboratory space was used to conduct the experiments to eliminate any potential distractions from the surroundings and to ensure participants’ comfort and privacy. Upon participants’ arrival at the dedicated lab, they were guided to take a seat in front of a desk and screen. They were instructed to remain seated throughout the entire session. All navigation and interactions within the virtual environment were designed to be done using controllers. Participants would translate through the environment using one of the controller joysticks and could rotate the forward view in 15-degree increments with the other. Furthermore, while remaining seated, they were able to look around and rotate freely. As participants recognized the targets, they collected them by pressing the right trigger.

3.2.3 Conditions

Most prior studies focused on attention and engagement in XR learning used a between-subject design to limit the variables participants were exposed to, reduce study duration, and minimize cybersickness [58]. Following this approach, this study also adopted a between-subject design. A control condition with no additional sensory cues was designed to compare task completion with three other conditions equipped with sensory cues: an extended field of view through a rearview mirror, audio, and haptics. The goal was to identify the most effective cue for orienting attention in the virtual environment. All four conditions were designed to maintain consistent mechanics and objectives with the only difference being the unique sensory cue provided. Each participant was randomly assigned to only one condition. Each condition is described below.

3.2.3.1. Control Condition. Under this condition, a 90° front horizontal field of view was available without additional cues being provided.

3.2.3.2. Mirror Condition. A rearview mirror was available throughout the session and extended participants' vision. Using this mirror, participants could see paintings behind. The mirror was 0.5 × 0.5 m and located 0.9 m away from the participant. It had a field of view of 60°, rotated 30° along the yaw axis, and was located on the middle right side of their front view. It was also transparent to prevent blocking the portion of the front view underneath the rear view. Fig. 4 shows a player's view of this condition.

3.2.3.3. Audio Condition. The headset provided a beep when the participant moved within a five-meter radius from a target painting. The beep played for half a second and did not repeat unless the participant moved outside the five-meter radius and re-entered it. This distance was chosen based on preliminary testing, which indicated that 5 m provides sufficient time and space to react to the cue. Within this radius, users have multiple paintings that allow them to identify and pick one rather than being limited to a single painting that indicates the correct target.

3.2.3.4. Haptic Condition. The controllers vibrated when the participant moved within a five-meter radius from a target painting. The controllers vibrated for half a second. Just like audio, the vibration did not repeat unless the participant re-entered the five-meter radius. This distance was chosen based on the same reasons applied to the audio condition.

3.3 Study Variables and Data Analysis Approach

The investigation combined quantitative assessments of task performance with qualitative analyses of participants' survey and questionnaire responses following the VR task. Because individual performance measures can indicate the speed of task completion but cannot explain why one guidance method might yield faster results than another, qualitative measures were invoked to aid in the triangulation of why the quantitative measures emerged as they did. This mixed-method approach provides a comprehensive understanding by capturing both measurable outcomes and subjective experiences.

3.3.1 Performance Variables

The task was designed to require both visual search (for each painting) and memory (of the list) even though the list was constantly available. Therefore, performance variables included the following: (1) a count of how many times the user checked the task list (measured by when the user raised their left-hand controller to show the list regardless of duration) and (2) the total task-completion time. After the task began, a custom script in Unity captured participants' performance data. The session had a duration limit of 20 minutes and all participants finished the task within the time limit. The participants' performance data included their IDs

and timestamp, event, and painting variables. Timestamp indicated when a cue or action occurred in the task session. Event showed when a cue was received (the player was within a distance of 5 m from a target) and removed (the player was located farther away than 5 m from targets) and whether a target painting was collected. The painting variable recorded the name of the target that was visually seen in the control or rearview mirror condition or the target that provided feedback in audio or haptic conditions and if it was collected. All of these raw data could be used to calculate the final performance variables.

3.3.2 Survey

In addition to participants' performance, collecting their feedback provides deeper insights into their experience, allowing for a more holistic evaluation. Understanding user engagement and the role of guidance cues in virtual environments requires examining factors beyond objective task completion. Existing research on attention in experimental settings largely depends on post-test evaluations [58], underscoring the importance of capturing subjective user responses. To assess these experiential aspects, participants completed a post-task survey presented on a nearby computer screen. The survey included eight questions exploring their enjoyment, immersion, perceived performance, cue effectiveness, and willingness to engage with similar tasks in different virtual environments. Control condition participants skipped the question regarding the effectiveness of the given cue. The complete list of questions is as follows:

- (1) Did you feel you were trying your best to finish the game?
- (2) Did you feel the urge to see what was happening around you (in the VE)?
- (3) To what extent did you find the game challenging?
- (4) To what extent did cues help you if you received any? Otherwise write "No cue received."
- (5) To what extent did you lose track of time?
- (6) To what extent did you enjoy playing the game rather than something you were just doing?
- (7) How well do you think you performed in the game?
- (8) Would you like to visit somewhere else, for example, the College of Design, this way?

Thematic analysis [46] was conducted to identify recurring themes in the free-response survey questions using Taguette [29]. Afterward, the NASA-TLX [21], Presence Questionnaire (PQ, Version 3.0) [67], and Simulator Sickness Questionnaire (SSQ) [2] were also completed, requiring approximately 15 minutes for the whole survey. The NASA-TLX and PQ were used as measures that might correlate with the seamlessness of the interface. The SSQ was used to note whether cybersickness occurred. Some XR studies conduct a pre-SSQ baseline and a post-SSQ and examine the change [9, 30]; other studies carry out only the post-SSQ [50]. Because a study at the local research lab recently conducted pre- and post-SSQ but found the pre-SSQ data to be irrelevant in a sensitivity analysis, the current study used only the post-SSQ data [51].

3.3.3 Data Analysis

Quantitative data were analyzed using SPSS, with analysis of variance (ANOVA) employed to assess statistical differences between conditions. Qualitative data were examined using Taguette, applying thematic analysis to identify recurring patterns, themes, and key insights from participant responses.

3.3.4 Predictions

Based on previous research, the multiroom scavenger-hunt-style multiple visual search task, and the design of the cues, it was predicted that the rearview mirror's expanded field of view would aid the user's search, decreasing the completion time. It was also hypothesized that the active alerting systems (audio and haptic) would decrease the number of times users would need to reference the task list since the alerts would indicate that a target painting was nearby. It was predicted that all three of the guidance systems would lead to faster task-completion times than the control condition. Finally, it was predicted that the audio and haptic conditions would lead to higher ratings of presence than the mirror since they did not interrupt the visual experience.

4. RESULTS

The purpose of this study was to compare the effect of three sensory cues on attentional orientation in VR. Of the 71 participants, three chose to stop the task before completion due to cybersickness. Their performance data were excluded from our analysis, resulting in $n = 17$ for each of the four groups.

4.1 Frequency of Checking Task List

After ensuring that the assumptions of an ANOVA were met, a one-way ANOVA was run to ascertain whether variation existed in the number of times participants of different conditions checked the task list attached. The ANOVA was chosen as it allows for comparing means across multiple independent groups to determine whether statistically significant differences exist. There was one extreme outlier in the control condition and two in the haptic condition falling above or below 1.5 times the interquartile range, which were removed. The updated data are shown as boxplots in Figure 5. Although the assumption of homogeneity of variance was not met as assessed by Levene's test for equality of variance ($p < .001$), the analysis was continued because of the robustness of ANOVA to such violations, especially with the nearly equal sample size groups across the four conditions. The number of times the participants checked the painting task list was significantly different across conditions ($F(3, 61) = 11.170$, $p < .001$, $\omega^2 = 0.31$). Using the Tukey post hoc pairwise analysis, the number of times participants checked the painting list was significantly lower in the haptic condition ($M = 0.07$, $SD = 0.25$) when compared to the control condition ($M = 4.56$, $SD = 6.50$; 95% CI [.08, 8.91], $p = .04$) and the rearview mirror condition ($M = 8.12$, $SD = 6.50$; 95% CI [3.70, 12.40], $p < .001$). In addition, the frequency in audio condition ($M = 0.35$, $SD = 0.60$),

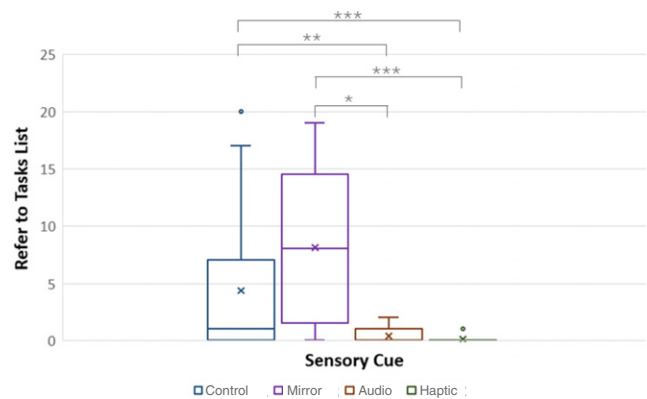


Figure 5. Frequency of referring to the task list. Means are indicated by the X in each boxplot.

was significantly lower than in both rearview mirror (95% CI [3.55, 11.98], $p < .001$) and in control conditions (95% CI [.08, 8.91], $p = .05$). The difference between the control and mirror conditions was not significant.

4.2 Task-Completion Time

A one-way ANOVA was completed to determine whether the total time spent in VR to complete the task was different for individuals in different conditions. There was one extreme outlier in control, four in haptic, and one in audio condition, falling above or below 1.5 times the interquartile range, which were removed. Figure 6 shows the boxplot of the updated data with extreme outliers removed. Despite the violation of homogeneity of variance assumptions as assessed by Levene's test for equality of variance ($p < .001$), the analysis was continued because ANOVA is robust to these violations when sample sizes are almost equal across all conditions. The total time spent was significantly different across conditions ($F(3, 58) = 17.88$, $p < .001$, $\omega^2 = 0.44$). The total completion time was statistically significantly higher in the rearview mirror condition ($M = 6.11$, $SD = 0.35$) than in the control condition ($M = 4.21$, $SD = 0.90$; 95% CI [0.87, 3.11], $p < .001$), the audio condition ($M = 3.34$, $SD = 1.09$; 95% CI [1.74, 3.98], $p < .001$), and the haptic condition ($M = 3.65$, $SD = 1.61$; 95% CI [1.45, 3.66], $p < .001$). Other pairwise comparisons were not significantly different.

4.3 Survey

Participants were asked to answer eight open-ended questions that were provided after VR task completion to gather qualitative feedback regarding their experience in this study. The in-depth questions offered a deeper understanding of the engagement level and the usefulness of the cues and the task list. Following the data collection, the responses were imported into Taguette, an open-source text coding tool and coded [29]; this approach helps systematically organize and interpret data for an easier identification of patterns and insights. The recurring responses were then categorized into five groups that emerged: beneficial, difficult, fun/enjoyment, time tracking, and retrying the

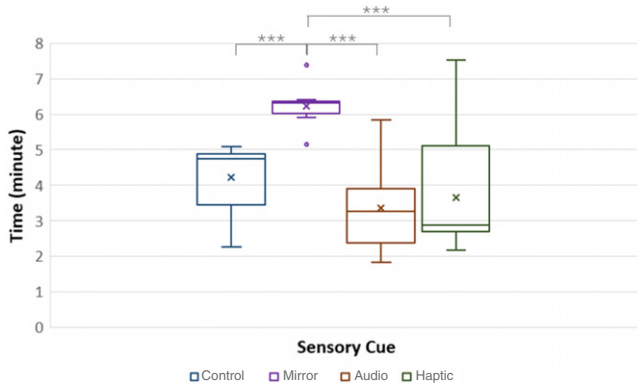


Figure 6. Total time spent to finish the task. Means are indicated by the X in each boxplot (right).

game. These were then fit into two main themes: effectiveness and engagement.

4.3.1 Effectiveness

The effectiveness of cues was emphasized across conditions. In both audio and haptic conditions, 100% of participants found the cues beneficial. One participant highlighted the received cue's efficiency this way: "At the end I missed one picture, and I was just walking around waiting for the cue to help finding the missing picture." Conversely, in the mirror condition, 47% of participants said they did not use the rearview mirror, 23% saw it as enhancing environment scanning, 11% used it slightly, and 6% used it frequently. Comments from participants highlighted the importance of cues in aiding task completion. One in the audio condition expressed their experience thus: "There were a few times where I definitely would have missed a painting or two if I did not have the cue." In the mirror condition, feedback varied with about 7% finding it immersive but not particularly effective for task accomplishment. Nearly 6% of participants in this condition also reported verbally that the mirror partially blocked their view. Additionally, 85% of participants in the control condition and 65% in the mirror condition found the task challenging. Meanwhile, 85% and 100% of participants in the audio and haptic conditions, respectively, found the task only slightly challenging, citing similar reasons for their experiences such as some similarities in the style of the paintings and uncertainty about the total number of rooms in the game.

4.3.2 Engagement

Participants reported that cues kept them engaged and focused throughout the session. Regarding enjoyment, 75% in the control condition, 83% in the mirror condition, 100% in the audio condition, and 86% in the haptic condition found the session more enjoyable compared to their typical daily activities. Participants of audio and haptic conditions highlighted cues as helpful reminders in task focus and completion. For instance, an audio condition participant mentioned that cues kept them on mission despite the urge

Table 1. Means (standard deviations) for results from NASA-TLX, PQ, and SSQ.

	Control	Mirror	Audio	Haptic
TLX: Performance	14.82 (5.44)	19.29 (1.45)	18.00 (2.30)	18.29 (3.39)
TLX: Mental	7.00 (4.66)	7.45 (4.54)	6.35 (5.34)	7.41 (3.69)
TLX: Physical	2.88 (2.85)	5.29 (5.21)	2.82 (3.13)	2.18 (2.60)
TLX: Temporal	6.64 (5.32)	6.11 (6.33)	4.53 (4.14)	2.23 (5.11)
TLX: Effort	8.35 (4.12)	7.88 (5.70)	6.53 (5.22)	7.06 (4.78)
TLX: Frustration	4.00 (3.00)	4.88 (6.78)	2.12 (2.98)	2.76 (3.73)
PQ	4.43 (2.09)	4.31 (2.08)	4.83 (2.06)	4.95 (1.98)
SSQ	1.24 (1.03)	1.35 (0.99)	1.41 (2.09)	1.38 (0.97)

to explore: "The audio was a nice reminder to make sure to thoroughly check the area I was exploring." Participants using the rearview mirror reported satisfaction with their performance; however, the mirror's role in engagement and joy was not specifically mentioned. Time perception varied with 70% in the control condition, 41% in the mirror condition, 76% in the audio condition, and 35% in the haptic condition experiencing a loss of time. Additionally, 85% of participants in both control and mirror conditions and 100% in both audio and haptic conditions expressed willingness to try this method again.

4.4 NASA-TLX, PQ, and SSQ

Three questionnaires were conducted to assess the impact of sensory cues on mental workload in VR (NASA-TLX), sense of presence (PQ), and cybersickness (SSQ). A one-way ANOVA was run to find whether participant differences were statistically significant across conditions. No significant differences were found. The means and standard deviations of the data are presented in Table 1.

Performance Consistency: Across all conditions, the Performance scores remained high, with mean values ranging from approximately 14.8 to 19.3.

Mental Demand: The mirror ($M = 7.47$, $SD = 4.54$) and haptic ($M = 7.41$, $SD = 3.69$) conditions had slightly higher Mental Demand compared to the control ($M = 7.0$, $SD = 4.66$) and audio ($M = 6.35$, $SD = 5.34$) conditions, but these differences were not statistically significant.

Physical Demand: The mirror condition ($M = 5.29$, $SD = 5.21$) had higher Physical Demand than the other conditions, all of which were below 3. The haptic condition showed the lowest Physical Demand ($M = 2.18$, $SD = 2.60$). However, these differences were not statistically significant.

Temporal Demand: The control condition ($M = 6.64$, $SD = 5.32$) exhibited the highest Temporal Demand. The audio condition ($M = 4.53$, $SD = 4.14$) showed the lowest Temporal Demand. However, due to the high variances, these differences were not statistically significant.

Effort: Effort was highest in the mirror condition ($M = 7.88$, $SD = 5.70$). In the control condition ($M = 8.35$, $SD = 4.12$), it was slightly higher than in the audio ($M = 6.53$,

$SD = 5.22$) and haptic ($M = 7.06$, $SD = 4.78$) conditions. These differences were not statistically significant.

Frustration: Frustration was highest in the mirror condition ($M = 4.88$, $SD = 6.78$). The audio condition had the lowest Frustration ($M = 2.12$, $SD = 2.98$), followed by the haptic condition ($M = 2.76$, $SD = 3.73$).

4.4.1 Presence Questionnaire

The PQ results showed no significant differences in the sense of presence across conditions. The overall mean presence score across all conditions was 4.41 ($SD = 2.09$). The mirror condition had a mean score of 4.31 ($SD = 2.08$), the audio condition had a mean of 4.83 ($SD = 2.06$), and the haptic condition had the highest mean score of 4.95 ($SD = 1.98$).

4.4.2 Simulator Sickness Questionnaire

The SSQ assessed symptoms across three subscales (Nausea, Oculomotor, and Disorientation). The total scores ranged from 1.24 in the control condition to 1.41 in the audio condition. The standard deviations ranged from 0.96 to 1.03 across conditions. The audio condition showed the highest mean symptom severity ($M = 1.41$, $SD = 0.96$), followed by the haptic condition ($M = 1.38$, $SD = 0.97$) and the mirror condition ($M = 1.35$, $SD = 0.99$). The control condition demonstrated the lowest severity ($M = 1.24$, $SD = 1.03$). Three participants chose to stop the game due to dizziness: one stopped in the middle of the task and two near the end.

5. DISCUSSION

In line with the previous literature [12, 32, 45, 55], the findings of this study indicated that the guidance cues impacted users' attention orientation. The audio and haptic cues (active cues) significantly reduced the frequency of referring to the task list, which supported our predictions. However, the rearview mirror condition (passive cue) had unexpectedly high task list use and time to complete compared to the active alert systems. A reason for this could be that the interface design or mirror placement made users less certain about the location of the painting targets or the mirror partially blocked their view, and they had to double-check their task list more regularly than with the active alert conditions. In terms of time for task completion, the rearview mirror required significantly greater time than even the control condition as well as the other two conditions. It may have added extra cognitive or visual demands, slowing down participants' performance. On the other hand, the audio and haptic conditions did not significantly differ from the control condition, indicating that these sensory cues did not drastically impact the task-completion time while adding some value in terms of guidance. This result suggests that the audio and haptic interfaces are closer to the desired seamless interaction that offers the user additional guidance without breaking the sense of presence. This result also contrasts with previous studies [59], suggesting that user differences, task characteristics, and environment may play a crucial role [1]. Further investigation is needed to explore these influences in more detail.

The survey responses resulted in two core themes: engagement and effectiveness. These two themes were crucial for assessing the effectiveness of cues because they showed that guidance cues impacted the participants' experience interacting with the VE and completing the task. Almost all users in the four conditions enjoyed the task they were assigned to, as they all mentioned in the survey, unaware of the existence of other conditions. They were engaged sufficiently to finish the task successfully regardless of the number of times they referred to the task list or the speed at which they finished the task. Survey responses indicated that haptics and audio were considered highly effective, especially for locating the more challenging paintings, which were those with intricate details and multiple objects, as opposed to simpler paintings featuring a single figure with few elements. Moreover, the paintings that remained elusive after longer exploration were primarily found by relying on the receiving haptic or audio cues; five participants mentioned that they were waiting for the cue to assist them to locate the last painting. Notably, this type of experience was noted in only a small number of participants in the rearview mirror condition. It is also notable that some users explained that they forgot to check the mirror and just relied on their front view. This result might stem from the habit that people use a rearview mirror primarily for safety reasons in vehicles rather than as a tool to locate targets.

The three questionnaires did not result in statistically significant differences between different conditions but offered important insights into the investigated cues on the designed task in VR. The results of NASA-TLX indicated that none of the cues resulted in a significantly increased workload even in the rearview mirror condition, where participants referred to the task list the most and spent a longer time to complete the task. Some variables had means that seemed notably different, but the high standard deviations removed the possibility of statistically significant differences. However, this pattern suggests that with a greater sample size, the mean differences might become significant; these areas are thus worthy of further research. Examples include the NASA-TLX Physical Demand and Frustration (with mirror having larger means than other conditions) and Temporal Demand (with control having a larger mean than other conditions). The mean presence score was lowest with the mirror condition and highest with the haptic condition (though not significantly different). The higher presence with the haptic condition is consistent with previous findings [18]. The audio condition also performed well, highlighting its potential to enhance immersion [26]. However, when confronted with a different scenario featuring a more challenging task or graphically different VE, these cues might cause a different sense of presence, necessitating further investigation.

6. CONCLUSION

Systems for attention orientation are crucial in VR tasks that require the guidance of users, and a balance of offering seamless, invisible guidance while still offering users freedom to explore is critical in virtual exhibition environments. The

findings of this study suggested that the active alert systems (audio and haptic cues) reduced the users' need to refer to the task list; they confidently counted on receiving help throughout the entire process with minimal interruptions. On the other hand, participants of the passive guidance system (rearview mirror condition) most frequently checked the task list to complete the task. Furthermore, the total time spent in the virtual Louvre Museum to locate all targets and complete the task was significantly longer in the rearview mirror condition compared to the other three conditions, possibly due to the mirror's limited effectiveness in guiding users and partially obstructing the front view.

The survey results highlighted that the haptic and audio cues were the most favorable cues in finding elusive targets after extended exploration, with some participants even waiting for cues to guide them to the target's location. These cues seemed to successfully achieve the desired balance of seamlessness while not inhibiting user exploration of the environment. Since this study focused on a virtual museum setting, caution is needed when generalizing the results to other scenarios, as the novelty of different VEs and task requirements may influence outcomes. However, the findings of this research emphasize the potential of well-designed guidance cues and offer VR designers guidelines for designing VEs that require a balance of user freedom and guidance.

ACKNOWLEDGMENT

The authors thank Dr. Jonathan Kelly, Dr. Kimberly E. Zarecor, and Professor Pete Evans for early guidance in the design of this study and feedback on the analysis.

REFERENCES

- S. Baldoni, M. S. Hadj Sassi, M. Carli, and F. Battisti, "Definition of guidelines for virtual reality application design based on visual attention," *Multimed. Tools Appl.* **83**, 49615–49640 (2024).
- S. A. Balk, M. A. Bertola, and V. W. Inman, "Simulator sickness questionnaire: twenty years later," *Proc. 7th Int'l. Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design* (University of Iowa, Bolton Landing, NY, USA, 2013), pp. 257–263.
- Y. Baram and A. Miller, "Virtual reality cues for improvement of gait in patients with multiple sclerosis," *Neurology* **66**, 178–181 (2006).
- F. Biocca, C. Owen, A. Tang, and C. Bohil, "Attention issues in spatial information systems: directing mobile users' visual attention using augmented reality," *J. Manage. Inf. Syst.* **23**, 163–184 (2007).
- R. S. Bolia, W. R. D'Angelo, and R. L. McKinley, "Aurally aided visual search in three-dimensional space," *Hum. Factors* **41**, 664–669 (1999).
- W. Boonsuk, S. Gilbert, and J. Kelly, "The impact of three interfaces for 360° video on spatial cognition," *Proc. SIGCHI Conf. on Human Factors in Computing Systems* (ACM, Austin, Texas, USA, 2012), pp. 2579–2588.
- F. Bork, C. Schnelzer, U. Eck, and N. Navab, "Towards efficient visual guidance in limited field-of-view head-mounted displays," *IEEE Trans. Visual. Comput. Graphics* **24**, 2983–2992 (2018).
- E. Bozkir, D. Geisler, and E. Kasneci, "Assessment of driver attention during a safety critical situation in VR to generate VR-based training," *ACM Symposium on Applied Perception 2019* (ACM, Barcelona, Spain, 2019), pp. 1–5.
- P. Brown and W. Powell, "Pre-exposure cybersickness assessment within a chronic pain population in virtual reality," *Front. Virtual Reality* **2**, 672245 (2021).
- M. Chalmers and I. MacColl, "Seamful and seamless design in ubiquitous computing," *Workshop at the Crossroads: The Interaction of HCI and Systems Issues in UbiComp* (2003), Vol. 8.
- F.-Y. Chao, C. Ozcinar, C. Wang, E. Zerman, L. Zhang, W. Hamidouche, O. Deforges, and A. Smolic, "Audio-visual perception of omnidirectional video for virtual reality applications," *2020 IEEE Int'l. Conf. on Multimedia & Expo Workshops (ICMEW)* (IEEE, Piscataway, NJ, 2020), pp. 1–6.
- T. Chen, Y.-S. Wu, and K. Zhu, "Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality," *Proc. 24th ACM Symposium on Virtual Reality Software and Technology* (ACM, Tokyo, Japan, 2018), pp. 1–5.
- P. Cipresso, I. A. Chicchi Giglioli, M. Alcañiz Raya, and G. Riva, "The past, present, and future of virtual and augmented reality research: a network and cluster analysis of the literature," *Front. Psychol.* **9**, 2086 (2018).
- S. Cosgrove and J. J. LaViola, "Visual guidance methods in immersive and interactive VR environments with connected 360° videos," *2020 IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (IEEE, Piscataway, NJ, 2020), pp. 652–653.
- R. Damayanti, B. P. Redyantanu, and F. Kossak, "A study of multi-sensory senses in museum virtual-visits," *Conf. Ser.: Earth Environ. Sci.* **907**, 012020 (2021).
- N. Dezelon, "Designing an Inclusive Audio Guide Part 4: Content Development: Telling the Warhol Story," (4 August 2016). [Online]. Available: <https://www.warhol.org/designing-an-inclusive-audio-guide-part-4-content-development-telling-the-warhol-story/>.
- C. Englund, A. D. Olofsson, and L. Price, "Teaching with technology in higher education: understanding conceptual change and development in practice," *Higher Educ. Res. Dev.* **36**, 73–87 (2017).
- J. K. Gibbs, M. Gillies, and X. Pan, "A comparison of the effects of haptic and visual feedback on presence in virtual reality," *Int. J. Hum.-Comput. Stud.* **157**, 102717 (2022).
- K. Guo, A. Fan, X. Lehto, and J. Day, "Immersive digital tourism: the role of multisensory cues in digital museum experiences," *J. Hospitality Tourism Res.* **47**, 1017–1039 (2023).
- W. Ha, M. G. Choi, and K. H. Lee, "Automatic control of virtual mirrors for precise 3D manipulation in VR," *IEEE Access* **8**, 156274–156284 (2020).
- S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): results of empirical and theoretical research," *Adv. Psychol.* **52**, 139–183 (1988).
- T. Hartmann and J. Fox, "Entertainment in virtual reality and beyond: the influence of embodiment, co-location, and cognitive distancing on users' entertainment experience," in *The Oxford Handbook of Entertainment Theory*, edited by P. Vorderer and C. Klimmt (Oxford University Press, Oxford, UK, 2021), pp. 717–732.
- P. Hock, S. Benedikter, J. Gugenheimer, and E. Rukzio, "CarVR: enabling in-car virtual reality entertainment," *Proc. 2017 CHI Conf. on Human Factors in Computing Systems* (ACM, Denver Colorado USA, 2017), pp. 4034–4044.
- H. Huygheleir, B. Schraepen, C. Lafosse, N. Vaes, F. Schillebeeckx, K. Michiels, E. Note, V. Vanden Abeele, R. Van Ee, and C. R. Gillebert, "An immersive virtual reality game to train spatial attention orientation after stroke: a feasibility study," *Appl. Neuropsychol.: Adult* **29**, 915–935 (2022).
- S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, "A systematic review of virtual reality in education," *Themes Sci. Technol. Educ.* **85**–119 (2017).
- A. C. Kern and W. Ellermeier, "Audio in VR: effects of a soundscape and movement-triggered step sounds on presence," *Front. Robotics AI* **7**, 494828 (2020).
- K. Kim, O. Kwon, and J. Yu, "Evaluation of an HMD-based multisensory virtual museum experience for enhancing sense of presence," *IEEE Access* **11**, 100295–100308 (2023).
- S. E. J. Knobel, B. C. Kaufmann, N. Geiser, S. M. Gerber, R. M. Müri, T. Nef, T. Nyffeler, and D. Cazzoli, "Effects of virtual reality-based multimodal audio-tactile cueing in patients with spatial attention deficits: pilot usability study," *JMIR Serious Games* **10**, e34884 (2022).
- Knowledge Bank, "The CAQDAS. Software for Qualitative Analysis," mvorganizing.org (2018).
- P. Kourtesis, J. Linnell, R. Amir, F. Argelaguet, and S. E. MacPherson, "Cybersickness in Virtual Reality Questionnaire (CSQ-VR): a validation and comparison against SSQ and VRSQ," *Virtual Worlds* **2**, 16–35 (2023).

- 31 C. K. Lankford, J. G. Laird, S. M. Inamdar, and S. A. Baker, "A comparison of the primary sensory neurons used in olfaction and vision," *Front. Cell. Neurosci.* **14**, 595523 (2020).
- 32 N. Li, N. Sun, C. Cao, S. Hou, and Y. Gong, "Review on visualization technology in simulation training system for major natural disasters," *Nat. Hazards* **112**, 1851–1882 (2022).
- 33 X. Lin, C. Hwang, and Y. Xu, "Classification of virtual reality fashion shows: from the perspective of user experience," *Int'l. Textile and Apparel Association (ITAA) Annual Conf. Proc.* (Iowa State University Digital Press, 2024), Vol. 80.
- 34 A. MacAllister, S. Gilbert, J. Holub, E. Winer, and P. Davies, "Comparison of navigation methods in augmented reality guided assembly," *Proc. 2017 Interservice/Industry Training, Simulation, and Education Conf. (IITSEC)* (National Training and Simulation Association (NTSA), 2016), Vol. 2017, p. 17208.
- 35 E. Magosso, A. Serino, G. Di Pellegrino, and M. Ursino, "Crossmodal links between vision and touch in spatial attention: a computational modelling study," *Comput. Intell. Neurosci.* **2010**, 1–13 (2010).
- 36 S. Mannion, A. Sabiescu, and W. Robinson, "An audio state of mind: understanding behaviour around audio guides and visitor media," *MW2015: Museums and the Web 2015* (Museums and the Web LLC, Chicago, 2015).
- 37 M. Marucci, G. Di Flumeri, G. Borghini, N. Sciaraffa, M. Scandola, E. F. Pavone, F. Babiloni, V. Betti, and P. Aricò, "The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence," *Sci. Rep.* **11**, 4831 (2021).
- 38 J. P. McIntire, P. R. Havig, S. N. J. Watamaniuk, and R. H. Gilkey, "Visual search performance with 3-D auditory cues: effects of motion, target location, and practice," *Hum. Factors* **52**, 41–53 (2010).
- 39 K. Momtahan, R. Héту, and B. Tansley, "Audibility and identification of auditory alarms in the operating room and intensive care unit," *Ergonomics* **36**, 1159–1176 (1993).
- 40 J. Nakamura and M. Csikszentmihalyi, "The concept of flow," *Handbook of Positive Psychology* **89**, 105 (2002).
- 41 E. Nonino, J. Gislis, V. Holzwarth, C. Hirt, and A. Kunz, "Subtle attention guidance for real walking in virtual environments," *2021 IEEE Int'l. Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (IEEE, Piscataway, NJ, 2021), pp. 310–315.
- 42 Z. Pousman and J. Stasko, "A taxonomy of ambient information systems: four patterns of design," *Proc. Working Conf. on Advanced Visual Interfaces* (Association for Computing Machinery (ACM), 2006), pp. 67–74.
- 43 W. Powell, T. Garner, S. Shapiro, and B. Paul, "Virtual reality in entertainment: the state of the industry," *British Academy of Film and Television Arts* (British Academy of Film and Television Arts (BAFTA), 2017).
- 44 A. R. Pritchett, "Reviewing the role of cockpit alerting systems: implications for alerting system design and pilot training," *SAE Technical Paper* 1–6 (2001).
- 45 E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo, and R. P. McMahan, "Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task," *IEEE Trans. Vis. Comput. Graph.* **21**, 794–807 (2015).
- 46 S. Riger and R. Sigurvinsdottir, "Thematic analysis," *Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative, and Mixed Methods* (Oxford University Press, Oxford, UK, 2016), pp. 33–41.
- 47 A. A. Rizzo, T. Bowerly, J. Galen Buckwalter, D. Klimchuk, R. Mitura, and T. D. Parsons, "A virtual reality scenario for all seasons: the virtual classroom," *CNS Spectr.* **11**, 35–44 (2009).
- 48 D. Romero-Ayuso, A. Toledano-González, M. D. C. Rodríguez-Martínez, P. Arroyo-Castillo, J. M. Triviño-Juárez, P. González, P. Ariza-Vega, A. Del Pino González, and A. Segura-Fragoso, "Effectiveness of virtual reality-based interventions for children and adolescents with ADHD: a systematic review and meta-analysis," *Children* **8**, 70 (2021).
- 49 S. Rothe, D. Buschek, and H. Hußmann, "Guidance in cinematic virtual reality-taxonomy, research status and challenges," *MTI* **3**, 19 (2019).
- 50 M. Sanaei, S. B. Gilbert, N. Javdpour, H. Sabouni, M. C. Dorneich, and J. W. Kelly, "The correlations of scene complexity, workload, presence, and cybersickness in a task-based VR game," *Int'l. Conf. on Human-Computer Interaction* (Springer, Nature Switzerland, 2024), pp. 277–289.
- 51 M. Sanaei, S. B. Gilbert, A. J. Perron, M. C. Dorneich, and J. W. Kelly, "An examination of scene complexity's role in cybersickness," *Ergonomics* **67**, 496–520 (2024).
- 52 R. Soret, P. Charras, C. Hurter, and V. Peysakhovich, "Attentional orienting in virtual reality using endogenous and exogenous cues in auditory and visual modalities," *Proc. 11th ACM Symposium on Eye Tracking Research & Applications* (ACM, Denver, Colorado, 2019), pp. 1–8.
- 53 R. Soret, P. Charras, I. Khazar, C. Hurter, and V. Peysakhovich, "Eye-tracking and virtual reality in 360°: exploring two ways to assess attentional orienting in rear space," *ACM Symposium on Eye Tracking Research and Applications* (ACM, Stuttgart, Germany, 2020), pp. 1–7.
- 54 O. Spakov, J. Rantala, and P. Isokoski, "Sequential and simultaneous tactile stimulation with multiple actuators on head, neck and back for gaze cuing," *2015 IEEE World Haptics Conf. (WHC)* (IEEE, Piscataway, NJ, 2015), pp. 333–338.
- 55 C. Spence and A. Gallace, "Recent developments in the study of tactile attention," *Canadian J. Exp. Psychol.* **61**, 196–207 (2007).
- 56 S. Sriworapong, A. Pyae, A. Thirasawasd, and W. Keereewan, "Investigating students' engagement, enjoyment, and sociability in virtual reality-based systems: A comparative usability study of Spatial.io, Gather.town, and Zoom," *WIS 2022, CCIS* (Springer, Cham, 2022), Vol. 1626, pp. 140–157.
- 57 Y. Suzuki, A. Honda, Y. Iwaya, M. Ohuchi, and S. Sakamoto, "Toward cognitive usage of binaural displays," in *The Technology of Binaural Understanding*, edited by J. Blauert and J. Braasch, Modern Acoustics and Signal Processing (Springer, Cham, 2020).
- 58 C. L. Taborda, H. Nguyen, and P. Bourdot, "Engagement and attention in XR for learning: literature review," *Virtual Reality and Mixed Reality, EuroXR 2024, Lecture Notes in Computer Science* (Springer, Cham, 2025), Vol. 15445.
- 59 H. Z. Tan, R. Gray, C. Spence, C. M. Jones, and R. Mohd Rosli, "The haptic cuing of visual spatial attention: evidence of a spotlight effect," *Proc. SPIE* **7240**, 724001 (2009).
- 60 R. Tanaka, T. Narumi, T. Tanikawa, and M. Hirose, "Guidance field: Potential field to guide users to target locations in virtual environments," *2016 IEEE Symposium on 3D User Interfaces (3DUI)* (IEEE, Piscataway, NJ, 2016), pp. 39–48.
- 61 W. Tarnig, I. C. Pan, and K. L. Ou, "Effectiveness of virtual reality on attention training for elementary school students," *Systems* **10**, 104 (2022).
- 62 H. R. Tsai, Y. C. Chang, T. Y. Wei, C. A. Tsao, X. C. Y. Koo, H. C. Wang, and B. Y. Chen, "GuideBand: intuitive 3D multilevel force guidance on a wristband in virtual reality," *Proc. 2021 CHI Conf. on Human Factors in Computing Systems* (ACM, Yokohama Japan, 2021), pp. 1–13.
- 63 M. Weinmann, C. Schneider, and J. v. Brocke, "Digital nudging," *Bus. Inf. Syst. Eng.* **58**, 433–436 (2016).
- 64 M. Weiser, "The world is not a desktop," *Interactions* **1**, 7–8 (1994).
- 65 B. K. Wiederhold and S. Bouchard, "Sickness in virtual reality," *Advances in Virtual Reality and Anxiety Disorders* (Springer, US, Boston, MA, 2014), pp. 35–62.
- 66 S. Wijewickrema, Y. Zhou, J. Bailey, G. Kennedy, and S. O'Leary, "Provision of automated step-by-step procedural guidance in virtual reality surgery simulation," *Proc. 22nd ACM Conf. on Virtual Reality Software and Technology* (ACM, Munich Germany, 2016), pp. 69–72.
- 67 B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: a presence questionnaire," *Presence: Teleoperators Virtual Environ.* **7**, 225–240 (1998).
- 68 B. Wöldecke, T. Vierjahn, M. Flasko, J. Herder, and C. Geiger, "Steering actors through a virtual set employing vibro-tactile feedback," *Proc. 3rd Int'l. Conf. on Tangible and Embedded Interaction* (ACM, Cambridge, UK, 2009), pp. 169–174.
- 69 X. Wu and Z. Li, "A review of alarm system design for advanced control rooms of nuclear power plants," *Int. J. Hum.-Comput. Interact.* **34**, 477–490 (2018).