Evaluating the Effectiveness of Redirected Walking with Auditory Distractors for Navigation in Virtual Environments

Nicholas Rewkowski* — UNC Chapel Hill

Atul Rungta† — UNC Chapel Hill

Mary Whitton‡ — UNC Chapel Hill

Ming Lin§ — UMD College Park, UNC Chapel Hill

(a) (b) (c) (d) (e)

Figure 1: (a) A user waving to the robot dog distractor; (b) L1: A simple living room scene; (c) L2: Razzaque's firedrill scene; (d) L3: A complex street-crossing VR simulator; (e) OA: A furniture-filled room.

Abstract

Many virtual locomotion interfaces allowing users to move in virtual reality have been built and evaluated, such as redirected walking (RDW), walking-in-place (WIP), and joystick input. RDW has been shown to be among the most natural and immersive as it supports real walking, and many newer methods further adapt RDW to allow for customization and greater immersion. Most of these methods have been demonstrated to work with vision, in this paper we evaluate the ability for a general distractor-based RDW framework to be used with only auditory display. We conducted two studies evaluating the differences between RDW with auditory distractors and other distractor modalities using distraction ratio, virtual and physical path information, immersion, simulator sickness, and other measurements. Our results indicate that auditory RDW has the potential to be used with complex navigational tasks, such as crossing streets and avoiding obstacles. It can be used without designing the system specifically for audio-only users. Additionally, sense of presence and simulator sickness remain reasonable across all user groups.

Index Terms: virtual locomotion—redirected walking—distractors

1 Introduction

Most existing head-mounted display (HMD) systems are tethered to a computer via a cable. This limits walkable physical space. Redirected walking (RDW), a method proposed by Razzaque et al. [31], enables users to really walk in a 3D virtual environment (VE) that is larger than the physical environment (PE). RDW takes advantage of the perceptual inaccuracies evident when humans attempt to walk in a straight line, but unknowingly move in a curve, and the human inability to maintain a precise sense of body orientation.

RDW works by applying an imperceptible amount of "rotational distortion" to the user's rotational movements, e.g., head rotation and its associated look direction, thus rotating the user's look direction so that it is slightly different in the VE than in the real world. The minimum space required for RDW is 5 x 5m, close to the maximum tracking area of the commercial HMD systems we work with [1, 17].

*e-mail: nr@unc.edu

†e-mail: rungta@cs.unc.edu

‡e-mail: whitton@cs.unc.edu

§e-mail: lin@cs.umd.edu

RDW has been studied extensively since its invention and effective implementations include ones that: distort the VE beforehand (resulting in no real-time rotational distortion) [8, 25], customize predictions of user's intended direction based on previous movement [9, 23], rotate based on more factors than head rotation (e.g., eye rotation) [43, 44], and ones that use graph methods to compute possible paths and their likeliness to be taken [49]. Other methods use visual illusions and distortion to improve redirection results [37, 41, 42]. A few methods add *distractors*, objects of interest, e.g., butterflies, to the VE. Distractors attract users' attention and distract them from noticing the rotational distortion that is causing increased physical rotation of their bodies [7, 27, 29].

Most RDW methods using distractors rely on users being able to see the distractors and determine their path forward in the VE based on what they see. We are interested in considering applications for vision-impaired (VI) persons, who cannot benefit from visual input. Some RDW studies consider this type of application and show that lack of vision may not prevent RDW methods from being used, albeit with different physiological thresholds and additional considerations [10, 26]. However, there is little evidence that methods such as distraction that work well for sighted users work well for users with no visual input. Research such as [12, 19, 20] conclude that audio-only navigational performance should generally be worse than vision-guided navigation, but should not be impossible. In Tabitha Peck's work [28], she proposes a distraction method she calls "distractor, audio", in which the user is asked to turn their head to follow the sound of a hummingbird's wings. This method resulted in a high probability of users' noticing the rotational distortion. It was overall considered unnatural by the users but also quite effective at inducing natural rotation compared to methods such as 2:1 turn [48]. However, users could still see the VE as they were not blinded.

In this paper, we expand knowledge of audio-based RDW systems with this report on studies that demonstrate that audio-only RDW methods such as Peck's are effective for users who have no visual cues. Audio distractors can enable "audio-only" RDW users to complete complex navigational tasks such as crossing a street or avoiding obstacles while using the same system that enables sighted users to complete them. Additionally, audio-only RDW can allow for imperceptible rotational distortion and results in no noticeable increase in simulator sickness for most users. To the best of our knowledge, this work is the first research that:

- Demonstrates that distraction is an effective technique allowing audio-only participants to complete navigational tasks;

- Compares performance for audio-only, vision-only, and audio-vision participants under the same system conditions in a RDW-based VR system; and
- Shows that audio-only distraction-based RDW systems can be imperceptible and evoke sense of presence, even with small tracking spaces.

Our results indicate that RDW and the more recent methods that extend its effectiveness can be potentially applicable to audio-based distractors. To summarize, we find:

- Obstacle avoidance, street crossing, and other navigational tasks can be completed using auditory distractors in RDW;
- Simulator sickness, sense of presence, and detection of rotational distortion do not seem to differ between audio and visual RDW;
- Audio-only users have lower walk speeds and distraction ratios (% of completion time being distracted) than users with vision. The data suggest a moderate-strong correlation between these performance variables independent of navigation task design.
- A complex audio environment, such as our Traffic Crossing scene, raises training concerns, e.g. sensory overload, that are not relevant to users with vision.

2 Background

The work reported in this paper builds on recent advances in virtual environments (VEs); evaluation of VEs, including their ability to evoke a sense of presence; virtual locomotion and in particular redirected walking with distractors; and psychoacoustics.

2.1 Immersion, Presence, and Virtual Locomotion

Immersive virtual environments substitute synthetic sensory stimuli for physical stimuli from the real world. Using Slater's definition of immersion, the more senses stimulated and the higher the quality of the synthetic stimuli, the higher the level of immersion and the higher the user's sense of being in the virtual place–often referred to as *sense of presence* [36]. While many senses can be stimulated in VR, vision and hearing provide dominant cues for moving around and we have chosen to investigate them individually and together to begin to develop an understanding of how visually- or hearing-impaired individuals might use VR to learn to how to safely navigate in the physical world.

A detailed survey of ways to measure presence is included in [32]. The Slater-Usoh-Steed Presence questionnaire has been used to measure presence since the mid-1990s [35]. It is easily administered online and, because it is often used, it enables some ability to compare results across studies. We use customized SUS questions for our measure of user's sense of presence.

High-fidelity virtual reality systems–the hardware, software, and the content, evoke a subjective sense of presence in the user: a sense that they are in the virtual place and that what is happening in the VE is really happening [34]. We designed our system with presence-evoking characteristics: users can see virtual representations of part of their bodies (their tracked hands) and they can use their hands to interact with a touch-panel display in the environment. As part of our user study we use a virtual touch panel display to collect user's responses to questions about their experiences with our system.

Of the many virtual locomotion user interface techniques, e.g., joystick/gamepad movement [18], walking-in-place, teleportation [4], and really walking, studies have shown that really walking evokes a higher sense of presence than do other techniques [45] For that reason, and because we hope to train walking people to cross streets, we have chosen real walking as our target locomotion modality. That, however, means we are faced with the problem of the mismatch between the relatively small size of tracked space in popular commercial/commodity VR systems and the comparatively large size of realistic training environments. Redirected Walking locomotion interfaces address this problem.

2.2 Redirected Walking and Distractors

Redirected walking (RDW), first described by Razzaque in [31], takes advantage of the fact that people do not naturally walk in a straight line and almost never have a totally accurate sense of their head or body position at any given moment. RDW adds or subtracts a small amount of rotation to a user's head or body movement so that the moving user is following a slightly different path in the PE than in the virtual environment (VE). To insure the distorted rotation is imperceptible, the amount of positive or negative rotational gain is kept below a predefined threshold. Previous work has reported differing values for these thresholds [6, 38], but there is not yet agreement on what the values should be. In this work we limit the rotational distortion to 15% of head rotation in order to minimize the possibility of simulation sickness while providing enough redirection to traverse the VE comfortably. Additionally, the rotational distortion detection threshold for audio-only distractors may be significantly lower than for vision-based groups [33].

Redirected walking techniques using rotational gains as described above reduce the size of the physical environment required, but the literature variously reports size of the space needed to enable infinite walking in a straight line. Razzaque [30] noted that users are less sensitive to rotational gain when they are already rotating their heads. He exploited this fact by designing a custom fire-drill environment (see our Level 2) that required people to look around (back and forth) after reaching each goal location in order to find the next goal. During multiple cycles of head rotation, most of Razzaque's users physically rotated nearly 180 degrees. The rest of the required 180 degree rotation was made up with subtle rotational distortion added in each frame as the user walked through the virtual and physical environments.

The question then became one of how to induce users to turn their heads without requiring custom environments. Inspired by a fellow student's class project, Peck [28] fully developed and studied the idea of distractors, that is, objects and/or sounds in the scene that move back and forth around the user that the user is told to watch and/or listen to as it moves. Following the object with the eyes and ears induces head rotations and during those rotations the rotational gain can be higher. Peck evaluated several types of visual distractors (ball and hummingbird) as well as audio distractors, e.g., the whir of hummingbird wings. Our work builds on Peck's results endeavoring to increase generalizability of the technique. While generalized redirected walking systems have been built [39], they have typically not been able to effectively redirect a user in the small tracking area of commercial/commodity VR equipment.

2.3 Initiating redirection with distractors

Two goals of redirection are steering the user away from the limits of the PE and, when the user approaches the limit, a potentially unsafe condition, stopping the user and turning them away from the boundary. The second case, known as *resetting*, is the one in which redirection with distractors excels. The question becomes when and where to initiate/trigger the redirection with distractors. Peck [28] and Chen [7] both used a set distance between the user and the boundary of the space as their trigger value; both observed that users could get *stuck* in a corner or tight space and multiple interactions with the distractor were required to get *unstuck*.

These observations led us to consider methods that would address this issue by changing the distractor trigger boundary from the quadrilateral PE boundary to an elliptical curve inscribed within it.

3 Evaluation

We conducted two user studies: one comparing the performance of participants with and without vision in a distractor-based RDW system, the other determining the efficacy of distractors in an audio-only navigation task. We assume that it is impossible for audio-only participants to complete the navigation task without an audio goal, thus they are familiarized with the sound cue of the goal beforehand. Vision participants are likewise told which visual goals to look for (large, green question marks) or goal direction. Familiarizing participants with distractors is handled similarly: audio-only participants are told its distinct audio cue (dog panting sound), vision participants are told what it looks like (red robotic dog). This occurs after the participant wears the HMD. The participants were recruited from the general population of a university campus. None of the participants knew what the PE looked like beforehand. All participants reported normal hearing, vision, and locomotive abilities.

3.1 Base System Overview

In order to compare audio and vision participants in the same context, our system considers failure conditions relevant to both audio and vision participants by providing bimodal (audio/visual) distractor cues, using elliptical bounds to prevent participants from getting "stuck" in corners (causing confusion for audio participants especially), and incorporating a feedback loop allowing all user groups to understand their progress in the environment.

3.1.1 Redirection and Reorientation

In order for the user to physically navigate an open VE that's significantly larger than the PE, we implement both redirection and reorientation by means of distraction. Fig. 2 illustrates this section.

We use the subtle, continuous steer-to-center (S2C) redirection [40] technique used in works such as [7, 29], which applies rotational distortion every frame such that the user ideally faces the center of the PE. Our intended environments are large and open, thus we decide that S2C is sufficient compared to RDW methods such as [23] based on the findings of [13]. Our per-frame rotational distortion coefficient c is $\pm 15\%$ of head rotation, slightly less than the finding in [33] that auditory RDW must remain within a $\pm 20\%$ threshold for the distortion to remain imperceptible.

Reorientation is required for the case in which the user has been predicted to be heading out of the PE bounds for a certain amount of time (in our case, 0.5 seconds). This time buffer is implemented to avoid intervening if the user is only glancing out of the PE bounds.We test a point one meter directly ahead of the user as a conservative estimate of human stride [14]. In order to prevent users from colliding with tracking equipment and to prevent them from getting "stuck" in a corner, we use an elliptical rather than rectangular boundary (See Figure 3). We use a simple trajectory prediction method which projects their position forward a specified distance (in our case, 1 meter) each frame. We implement reorientation using a distractor, which follows a roughly parabolic path around the user at a specific velocity (in our case, 1.5 m/s), during which the user is expected to look in its direction.

In a normal distraction, a distractor follows a path around the user starting from the current goal location and ending near the center of the PE behind them, much like Peck's butterfly [28, 29]. It navigates this path once and returns to its initial position (See Figure 2). This causes slight reorientation under a rotational distortion threshold of 60 degrees over the entire distraction relative to the user's pre-distraction look direction, in addition to the per-frame constraint.

A distraction "fails" if the user is still heading out of bounds *after* the distraction. If distraction fails multiple times in sequence (in our case, 3 times), then a "reset" distractor is activated, which follows a parabolic path starting near the center of the PE and ending on the opposite side of the user, much like Peck's sphere [28]. The distractor begins in front of the user, navigates this path 3 times, and stops in front of the user (causing the user to move their heads back and forth) (See Figure 2). This path rotates the user such that they are looking at the center of the PE by the end of the distraction, guaranteeing a safely walkable path (thus there is no rotational distortion threshold besides the per-frame constraint).

3.1.2 Bimodality

Enabling both audio and vision participants to navigate the VE under the same conditions requires both the goal and distractors to be bimodal, thus they feature spatialized visual and auditory components. Our goals are generally visually represented by a distinct 3D mesh and auditorily represented by a predefined sound cue. The distractor also has a distinct mesh and audio cue. When the user interacts with the goal by physically walking into it, it will play a distinct audio cue and either move itself or start the next phase of the study. Based on findings such as [22], when a distractor is active, the goal audio volume is lowered to mask the rotational distortion.

3.1.3 Feedback Loop

In order for all participants to track their progress, we implement a bimodal feedback loop with input to the control process being whether or not the user is looking at the distractor.

For visual feedback, a message appears in front of the user showing the value of an arbitrary point system that we call "Robot Happiness Points". If the user *is* looking at the distractor, the points will increase and the message is colored green. Otherwise, points decrease and the message is colored red.

For auditory feedback, if the user *is* interacting with the distractor, an unspatialized high-pitched "beeping" sound will play as positive feedback. Otherwise, a low-pitched "booping" sound will play as negative feedback.

3.2 Apparatus

The apparatus consisted of an HTC Vive connected to a desktop workstation (Intel Xeon E5-2630, 64GB RAM, Windows 10, NVIDIA GTX 1080 GPU) while sound was delivered through a pair of Beyerdynamic DT990 PRO headphones. A Leap Motion is mounted to the front of the HMD, and the resulting set of tracked hands appears in the VE to be used for completing questionnaires. The scenes were constructed in Unreal 4. The tracked space was a 5.0x5.5 meter rectangle, with trackers (HTC Vive lighthouses) located inside of the tracking area at two of the corners. The calibrated center of the space matched the physical center. The physical walls were at least at arm's length from the "safe" tracking space to avoid any wall touching. A generic HRTF is used for audio rendering [3].

The physical space was kept silent while conducting the studies to prevent extraneous noise from interfering with the participants' sense of presence. The VR study section was conducted in continuation to prevent the participants from seeing the lab and to make the experience as immersive as possible. This was a necessary measure for all subject groups, as prior knowledge of the environment has a dramatic effect on detection thresholds [26]. In order to offset the amount of time spent in VR, a Kennedy simulator sickness questionnaire (SSQ) was filled out by each participant online while immersed in VR between levels of the study [16]. The total amount of time in the analyzed scenes took around 15-20 minutes, with the entire study taking 40-50 minutes on average, including surveys and SSQs.

3.3 Common Experimental Design Elements

Here we discuss experiment design elements common between the two studies.

3.3.1 Stimuli

The original context for the studies was the training of VI users, so our distractor is a dog. The distractor's visual mesh is always a bright red, metallic robotic dog, and its audio cue is always the sound of a panting dog.

The other audio cues and visual meshes vary between scenes, as does the position of the goal and subject's spawn point. These variables will be identified in sections pertaining to each scene as the following: U_L is the subject's initial location when the scene begins, G_L is the location of the goal in the environment, G_A is

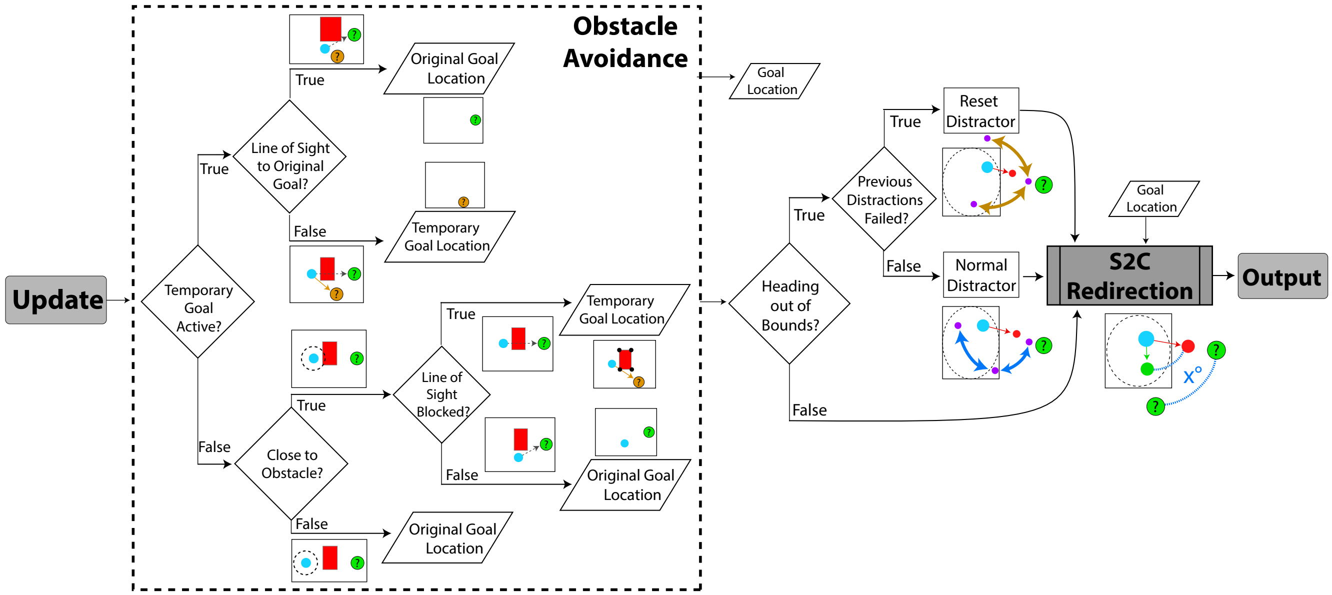

Figure 2: System Framework of the Distractor-based RDW System Used in Our Studies.

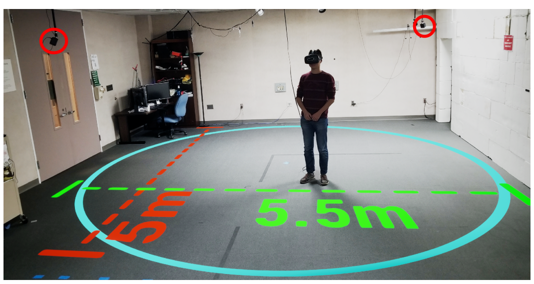

Figure 3: Equipment setup. Using an elliptical boundary prevents collisions with tracking equipment (circled in red).

the goal's audio cue, G_V is the goal's visual mesh, E_A is the set of extra audio sources (e.g., ambience), E_V is the set of extra, *non-environmental/non-Leap Motion hands* visual meshes designed to affect the study (e.g., dynamic objects).

3.3.2 Training

In order to familiarize participants with navigating in VR and interacting with the distractor, all participants complete two training scenes. The study begins when the Vive HMD is placed on the participants before they enter the study room containing the PE, and they first complete a training scene teaching them how to navigate in VR by really walking. In the second training scene, they are taught how to interact with the distractor by physically waving in its perceived direction. All participants have both vision and audio in these scenes, and RDW is not used as the navigable training VEs are smaller than the PE.

3.3.3 Intermediate Scenes

The intermediate scenes require the subject to fill in an SSQ and receive instructions about the next scene, such as what the audio cue and visual mesh or direction of the goal is. If the study coordinator notices any abnormal SSQ responses, then the subject would be asked to leave the study early and be compensated. These scenes include both audio and vision and do not include RDW. They are activated upon the completion of any other scene (including Training). The SSQ is contained in a Google Form shown on a virtual TV screen, which accepts answers through Leap Motion finger selection.

3.4 Study 1: Comparing Sensory Modalities

In this between-group study, we compare and analyze the performance of three subject groups with only the sensory modality condition as an independent variable (audio-only, vision-only, audio-visual) in three virtual environments (Living Room, Firedrill Scenario, Traffic Crossing). A learning effect is assumed between levels of increasing complexity, as RDW is known to require high cognitive load [5], so the VEs are completed in that order. The subject groups will be identified as A, V, and AV, respectively.

3.4.1 Participants

A total of 43 participants took part in the study (20 female, 23 male between the ages of 18 to 58, Mean=23). 14 participants had never tried VR before (4 in A, 4 in V, 6 in VA), 10 had only tried VR once or twice (5 in A, 4 in V, 1 in VA), and 19 had a fair amount of VR experience (6 in A, 6 in V, 7 in VA). They were assigned to one of three conditions (A, V, and AV), after which the study was completed in one session. The A group consisted of 8 females and 7 males, the V group consisted of 5 females and 9 males, and the AV group consisted of 7 females and 7 males. All participants reported normal vision and hearing, were naive as to the purpose of the study, and successfully completed the entire study.

3.4.2 Experimental Design

This study began with two training scenes (discussed in Section 3.3.2), intermediate scenes (discussed in Section 3.3.3), and three distinct scenes of increasing complexity discussed here.

Level 1: Living Room This is a simple scene requiring the subject to virtually walk in a straight line towards a goal, with G_L being the opposite side of the 20x20m living room from U_L, G_V as a large, green question mark, and G_A as a cue of a prerecorded man reading an arbitrary script. A reverberation filter is applied to the audio sources (goal and distractor) for immersion and to prevent front-back confusion [3, 21, 47]. E_A and E_V are empty. See Fig. 4.

Level 2: Firedrill Scenario This scene is a recreation of the virtual environment used in [31]. The subject begins on one side of a 5x21m room and must press buttons along the walls in order to proceed in such a way that they zig-zag between sides of the room. U_L is the same as in the original experiment. G_A and G_V are the same as in Scene 1. G_L is dynamic and moves to the next predefined button location as soon as the subject is close enough to the current button goal. There are a total of 5 buttons that are 8m apart from each other, requiring at least one distraction cycle between buttons. E_A includes a distinct audio cue that plays when the current goal has moved and some industrial background noise that fits the room's aesthetic. E_V is empty. All audio sources are affected by the reverberation filter used in L1. See Fig. 5.

Level 3: Traffic Crossing This section of the study was inspired by [2], in which children were trained to cross a virtual street while gaining real-world experience. In this scene, U_L is one side of a busy, 37m wide street crossing in front of a predefined pedestrian crosswalk and G_L is the opposite sidewalk. Vehicles drive along the street until their light turns red at the crosswalk, after which two distinct crosswalk beeper sounds will play on opposite sidewalks and pedestrians may cross. There is no specific G_V as reaching any position on the opposite sidewalk will trigger a success. G_A is a crosswalk beeper sound distinct from another beeper near U_L, so the participants can differentiate G_A from the starting point. This is especially important for the A group to avoid disorientation in the middle of the street in the absence of a visual reference.

E_V includes virtual, animated human pedestrians that walk along the sidewalk with the subject and randomly decide whether to cross at the crosswalk. When pedestrians may to cross, they will do so and then walk along the opposite sidewalk. E_V also includes various vehicles (eg. police cars, motorcycles) that drive along the street and wait for the subject and virtual agents to cross when appropriate.

E_A includes the distinct crosswalk beeper cue near U_L and pedestrian audio cues meant to help participants in group A correct their paths, such as random audio cues placed at the virtual pedestrians' mouths and footstep sounds located at the pedestrians' feet, based on the findings of [46] that non-speech cues may aid navigation. Additionally, E_A includes vehicle sounds meant to help participants in group A know where *not* to go, such as various engine sounds following the vehicles while they are driving and various honking sounds following the vehicles playing occasionally when they are stopped. Finally, E_A includes city ambience sounds. See Fig. 6.

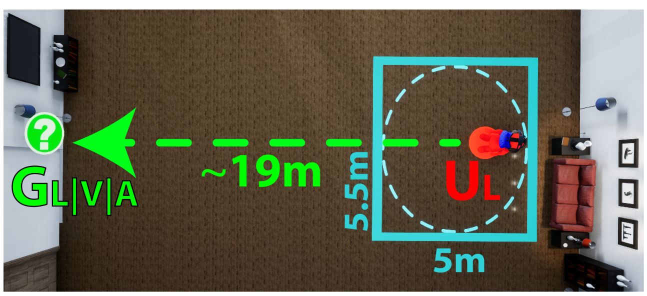

Figure 4: Level 1 (Living Room) layout. The goal is directly ahead of the subject. G_L,G_V, and G_A are at the same location.

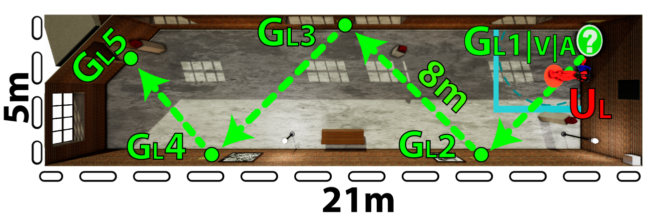

Figure 5: Level 2 (Firedrill Scenario) layout. The subject must touch 5 buttons in the sequence indicated by arrows.

3.5 Study 2: Enabling Complex Navigation

In this single-group study, we evaluate the possibility of using distraction to enable the completion of complex navigational tasks with only audio cues by analyzing the performance of audio-only participants, referred to as Group OA, in a scene requiring them to navigate around obstacles with only goal and distractor sounds.

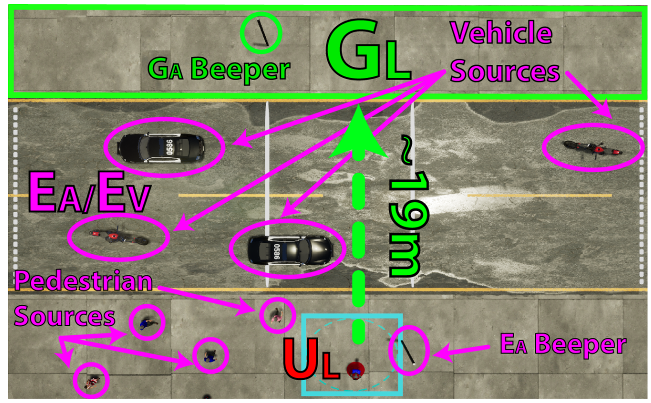

Figure 6: Level 3 (Traffic Crossing) layout. The subject waits for the cars to stop and walks across the street with virtual pedestrians. This scene with multiple moving sound sources is complex and realistic.

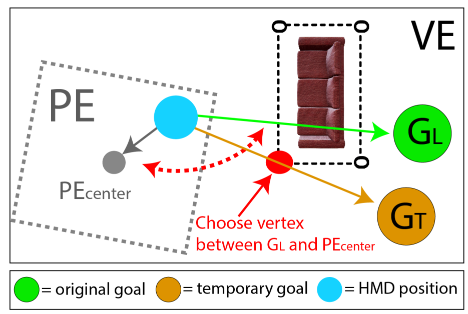

Figure 7: The temporary goal in Level OA is the same distance from the subject as the original goal, albeit in a direction that prevents them from colliding with the obstacle.

3.5.1 Obstacle Avoidance

In this study, we add an obstacle avoidance component to our Study 1 system which assumes that there is sufficient space between obstacles for the participants to navigate around and that there is *always* a navigable path to the goal. We assume that participants with vision do not require such a system as they can see the virtual obstacles.

We generate a rectangular buffer zone of 1 meter around all obstacles. If the subject approaches this zone, then a temporary goal G_T is created in a direction between the original goal G_L and center of the PE PE_{center} that can be reached without colliding with the obstacle (See Figure 7). Additionally, the volume of the original goal's audio cue is muted and the temporary goal is given a distinct cue. When the subject is determined to have line of sight to the original goal (not considering other obstacles), the temporary goal disappears and the original goal's volume is restored.

3.5.2 Participants

A total of 14 (5 females, 9 males) subject took part in this study and were recruited based on flyers and list-servs from the university campus. None had participated in Study I.

3.5.3 Experimental Design

This study began with two training scenes (discussed in Section 3.3.2), intermediate scenes (discussed in Section 3.3.3), and a scene including obstacle avoidance tasks (Scene OA).

Level OA: Obstacles in a Living Room This scene features the same VE as Scene 1 in Section 3.4.2. U_L, G_A and G_V are the same as Scene 1. Obstacles are added between U_L and G_L in the form of living room furniture, including 3 couches, 4 chairs, and 3 large tables. E_V is empty as the participants in Study 2 do not have vision, and E_A only includes a beeping noise that plays at the temporary goal location discussed in Section 3.5.1.

Figure 8: Level OA (Obstacle Avoidance) layout with example paths. It uses the same layout as L1 with the addition of obstacles. The cyan boxes are buffer zones that trigger the OA module. Group OA has no visual cues.

4 Data, Analysis, and Results

Our dependent variables include self-reports of participants' experience in each scene and quantitative measures of participant performance. Unless otherwise noted, data was collected in both studies. We performed statistical analyses using IBM SPSS v25. We report our results as a series of observations based on statistical outcomes.

Please note that descriptive statistics of all variables and figures with prefix "SM" are included in the Supplemental Material.

4.1 Dependent variables and collection methods

We recorded and analyzed data that we divide into *Quality* and *Performance* metrics.

Quality measures, gathered only for Study 1, were generated by participant responses to three questionnaires, two standard and one custom-developed for this study. The questionnaires were:

- **Simulator Sickness**: Kennedy Simulator Sickness Questionnaire (SSQ) [16].
- **Perception of Distortion**: This custom questionnaire, designed to identify whether participants perceive any distortion, is included in the Supplemental Material. Detection measurement methods such as in [38] were not used as we did not assume them to apply to audio or distractor-based RDW, and a specific detection threshold was unnecessary for this work.
- **Sense of Presence**: Slater-Usoh-Steed (SUS) Sense of Presence Questionnaire [35].

After finishing the navigation task in each level, participants completed the SSQ while in the previously described Intermediate Scene. After participants navigated all three levels and completed their third SSQ, they exited the VE and, using an online tool (Qualtrics), filled out the Distortion questionnaire once for each of Levels 1, 2, and 3. Using the same online tool, the participants completed the SUS questionnaire one time, responding based on their entire experience in the VEs. SSQ subscale scores were computed using the method in [16]. Distortion questions were Yes/No responses (with "please explain" follow-up questions to enhance our understanding), *any* response indicating notice of distortion was scored as a {L1,L2,L3} Detection value of 1 (otherwise 0), and were analyzed with logistic regression. SUS questions were on a 6 point Likert scale and, after converting scores to a binary variable ($1,2,3,4 \rightarrow LOW = 0$; $5,6 \rightarrow HIGH = 1$), were analyzed with logistic regression. SSQ comparisons were computed using a Repeated Measures ANOVA with the within-subjects factor as Level and the between-subjects factor as Modality. Equal variances are not assumed, thus our post-hoc analysis is completed with Games-Howell Tests.

Performance measures were gathered and analyzed for each level separately. The metrics computed for both Study 1 and 2 include:

- Count of *normal* and *reset* distraction events
- Time to complete for each Level
- Distraction ratio: percent of time to complete during which a distractor event is active
- Average walk speed during time when distractors are inactive (assuming participants stop walking during distractor events)

Performance metrics collected **only** for Study 2 include:

- Count of number of collisions between participant and nearby obstacles
- Count of number of collisions avoided when the obstacle avoidance method was triggered by participant entering the buffer zone around an obstacle (thus not avoided by chance).

Performance measures were found in or derived from data in the logs collected during each participants VR experience. Data logged on a per-frame basis includes: time stamp, user head pose, and whether a distractor event was active and, if so, whether it was normal or reset. Time to complete a navigation task, amount of time when normal and/or reset distractors were active, and walk speed were computed from time stamps and position data in the logs. Distraction ratio and average walk speed were derived from those data. We determine significance between groups using a Multivariate Two-Way ANOVA with Bonferroni correction, with $p(\alpha < 0.05)$. A repeated measures ANOVA is considered to be inapplicable as the design, content, and primary task of each level is assumed to be too different to reliably assume consistency between scenes.

4.2 Observations

1. As summarized in Figure 9, we find **no significant effect** in Quality values between groups except for in L3Detection, in which logistic regression between Groups A and V results in p=**0.037**.
2. We found that V and AV subjects had average distraction ratios of around 70%, which is higher than in systems such as [7], while A subjects had more typical values of about 50-60%. Refer to Fig. SM6 in Supplemental Material for details.
3. As summarized in Figure 11, we perform a post-hoc multivariate ANOVA with Bonferroni correction resulting in many dependent variables with significant effect between Group A and the others, while we find **no significant effect** between Groups V and AV in *any* dependent variable.
4. In L1: Living Room, we find that Group A encounters **slightly fewer** normal distractions than the other groups (A vs V: p=**0.065**; A vs AV: p=**0.025**) and had **much lower** distraction ratios (A vs V: p=**0**; A vs AV: p=**0**) and walk speeds (A vs V: p=**0.023**; A vs AV: p=**0.004**). Refer to Figure 11 for details.
5. In L2: Firedrill Scenario, we find a significant effect only in distraction ratio, with Group A once again having a **much lower** value (A vs V: p=**0.005**; A vs AV: p=**0.022**). Refer to Figure 11 for details.
6. In L3: Traffic Crossing, we find significant effect between Group A and the others in most values except number of normal distractions. Group A encountered **slightly fewer** reset distractors than AV (p=**0.043**). Group A has **much lower** distraction ratios (A vs V: p=**0**; A vs AV: p=**0**) and average walk

speeds (A vs V: p=**0.028**; A vs AV: p=**0.019**). However, their completion time was **much higher** (A vs V: p=**0.002**; A vs AV: p=**0.026**). Many users in Group A informally noted a jump in complexity in the audio field between L2 and L3. Refer to Figures 11 and Supplemental Figure SM6 for details.

7. As summarized in Figure 12, we find that the effects of Level and Modality on the dependent variables are **independent**, with a trend in distraction ratio ($p = \mathbf{0.062}, observed\,power = \mathbf{0.656}$), thus the correlation between walk speed and distraction ratio may be independent of task.
8. We find a **moderate-strong positive correlation** between walk speed and distraction ratio in general (especially in Group A), with a Pearson correlation of **0.625** over the entire dataset (0=no correlation, 1= total positive correlation). See Figure 13 for more details and pairwise correlations, as well as the Supplemental Material for relevant box plots.
9. After categorizing SSQ questions as in [16], we run Games-Howell Post-hoc tests showing **no significance** between groups in any category. See Fig SM2 in Supplemental Material. However, as shown in Figure 10, we find a significant interaction between Level and Modality on Nausea ($p = \mathbf{0.012}$).
10. We find that our OA solution is successful about 2/3 of the time, as seen in Figures SM8 and SM9 in Supplemental Material.
11. As shown in the MANOVA found in Figure SM11 in Supplemental Material, we find **no significance** in gender or levels of experience.

Experience Quality Metrics (Study 1)

Variable	Logistic Regression Comparison							
	Between Groups		A vs V		A vs AV		V vs AV	
	χ2	p	χ2	p	χ2	p	χ2	p
L1 Detection	0.555	0.456	0.518	0.471	0.518	0.471	0	1
L2 Detection	0.294	0.588	2.025	0.155	0.334	0.563	0.706	0.401
L3 Detection	0.017	0.897	4.363	0.037	0.009	0.924	3.851	0.05
Study 1 Presence	1.681	0.195	0.833	0.362	1.675	0.196	0.144	0.705

Figure 9: Logistic regression analysis comparing user groups' quality metrics from **Study 1**. Significant p-values, i.e. p($\alpha < 0.05$), are highlighted in yellow, trending values are orange. Generally, the experience quality has **no** statistically significant effect.

Test of Level*modality Between-Subjects Effects on SSQ			
SSQ Illness Type	F	p	Power
Nausea	3.457	0.012	0.836
Oculomotor	0.832	0.482	0.219
Disorientation	1.673	0.192	0.377
Total	1.9	0.148	0.433

*all are Greenhouse-Geisser corrected values

Figure 10: A Repeated Measures ANOVA on **Study 1** SSQ data shows a **significant** interaction between level and modality on Nausea ($p = \mathbf{0.012}$, possibly because the audio-only users are not normally blind.)

5 Discussion

Redirection vs Reorientation for Audio-only Participants: Study 1 reveals many differences between audio-only and vision participant groups in the context of RDW. We believe that continuous, per-frame redirection worked best for groups V and AV, as participants moved their heads noticeably more than Group A, often to look at environmental objects, allowing for the S2C redirection to have a greater effect on their motion. Audio-only participants tended to look straight ahead and avoided rotating their heads, causing reorientation by distraction to be more effective than the continuous S2C method. As seen in Observations 4 and 6 (L1 and L3), participants in Group A had many fewer reset distractions than the other groups, which also implies the success of the normal distraction method. We found that those in Group A that required a reset distractor often triggered it by planting their feet and not moving their heads.

Performance Metrics (Study 1)

Level	Variable	MANOVA Comparison		
		A vs V	A vs AV	V vs AV
		p	p	p
Living Room	Normal Distractions (#)	0.065	0.025	1
	Reset Distractions (#)	0.121	0.287	1
	Distraction Ratio (%)	0	0	1
	Average Speed (m/s)	0.023	0.004	1
	Completion Time (s)	1	1	1
Firedrill Scenario	Normal Distractions (#)	1	1	0.837
	Reset Distractions (#)	0.097	0.778	0.894
	Distraction Ratio (%)	0.005	0.022	1
	Average Speed (m/s)	0.314	1	1
	Completion Time (s)	0.991	1	1
Traffic Crossing	Normal Distractions (#)	0.425	0.743	1
	Reset Distractions (#)	0.69	0.043	0.583
	Distraction Ratio (%)	0	0	1
	Average Speed (m/s)	0.028	0.019	1
	Completion Time (s)	0.002	0.026	1

* Note: all values corrected using Bonferroni correction

Figure 11: Multivariate ANOVA analysis comparing performance metrics from **Study 1**. Significant p-values (p($\alpha < 0.05$)) are highlighted in yellow, trending values are orange. Distraction ratio and walk speed are generally **lower** for users in Group A, with these values having a **positive** Pearson correlation of **0.625**.

Test of Level*modality Between-Subjects Effects on Performance			
Dependent Variable	F	p	Power
Number of Normal Distractions (#)	1.679	0.159	0.503
Number of Reset Distractions (#)	0.957	0.434	0.296
Distraction Ratio (% of Completion Time)	2.306	0.062	0.656
Average Walk Speed (m/s)	1.525	0.199	0.46
Completion Time (s)	1.865	0.121	0.552

Figure 12: Test of Level*modality between-subjects effect in Performance metrics from **Study 1**. We find **no significant interaction** between level design and modality on performance, with a trend in the distraction ratio (p=**0.062**, highlighted).

These participants informally noted that they were worried about losing their bearing, as they had trouble localizing the goal to begin with. We also found from informal feedback that the dog distractor was considered more annoying to the vision participants than the audio participants, with many vision participants complaining that it appeared too often, which was also noted by Peck and Chen [7,28].

Correlation between Walking Speed and Distraction Ratio: As noted in Observation 8, our data shows a moderate-strong correlation between walk speed and distraction ratio, which we also noted informally during the study. Participants with vision often walked very quickly, especially in L3 (See Observation 6), in which they tried to match the fast walk speed of the virtual pedestrians (see Supplemental Material, Figure SM10). They also tended to ignore the distractor more often or not stop for it as quickly as Group A, causing it to be triggered multiple times in sequence and triggering resets, especially in L2. Additionally, vision participants tended to walk so quickly that they would leave the PE's safe bounds before realizing the distractor was active, often requiring multiple resets to bring them back to safety. Participants in Group A, however, spent more time standing still, rotating their head and localizing audio sources. Their physical movements were slower and more deliberate, resulting in a more natural curved path and fewer resets (especially in L2; see Observation 5 and descriptive statistics in Supplemental Material) in addition to the known veering behavior of blinded users [15]. This finding agrees with others such as [24], in which slower walk speed aids experience quality and performance.

Effects of Sensory Overload: An unexpected finding is the lack of significance between Quality metrics between groups, as noted in Observation 1. We expected Group A to be significantly less immersed and more likely to notice rotational distortion, but this was

		Bivariate Pearson Correlation Coefficient		
Level	Group\Comparison	Distraction Ratio x Walk Speed	Distraction Ratio x Completion Time	Walk Speed x Completion Time
Living Room	A	**0.698**	-0.188	**-0.577**
	V	**0.877**	**0.558**	0.484
	AV	0.074	0.454	-0.226
Firedrill Scenario	A	**0.562**	0.373	-0.445
	V	**0.69**	**0.633**	0.146
	AV	-0.006	0.14	**-0.68**
Traffic Crossing	A	**0.576**	0.263	0.077
	V	0.23	0.489	-0.286
	AV	0.423	0.013	-0.177
Significant at 0.01 level (99% confidence) (2-tailed)				
Significant at 0.05 level (95% confidence) (2-tailed)				

Figure 13: Pearson correlation calculated between performance values in **Study 1**. We conclude that there is a **moderate-strong** correlation between distraction ratio and walk speed for Group A in all scenes. Other comparisons are **not** consistently significant between levels.

not the case. We suspect that immersive additions to the audio field, such as reverberation and background noise, increased immersion for Group A despite lack of vision. Additionally, we find that in general, based on Observation 3, there was no statistically significant difference between Groups V and AV in terms of both experience quality and performance, implying that audio cues are dominated by visual cues in our case, which was also noted in works such as [11]. Finally, as seen in Observation 4 and 5, we found that completion time was not significant between groups in L1 or L2, which we suspect is because any time saved by Groups V and AV by moving faster is lost after being stopped by the distractor. However, Group A took much longer than the others to complete L3 (See Observation 6), which we suspect is caused by sensory overload in an overly-complex audio field in addition to cognitive load overhead from RDW itself [5]. Thus, we conclude that this complicated navigation task in a virtual environment may require more training and familiarization than the simpler tasks in L1 and L2.

Audio-based Distractors and Simulation Sickness: Another unexpected finding is that the effects of Level and Modality do not appear to be interact in the Performance metrics, as noted in Observation 7. This implies that experience design and task may not affect the noted correlations (e.g., between walk speed and distraction ratio as in Observation 8) or the experience quality of audio-only participants in other RDW scenarios, allowing for a certain degree of flexibility in experience design. We also note in Observation 9 that audio-only RDW and distraction do not generally appear to affect simulation sickness, possibly allowing for sickness-causing factors in VR to be handled the same way between audio-only and vision users. However, we note an interaction between Level and Modality on Nausea, which we believe might be caused by the fact that the audio-only participants did not have real blindness, thus they require a visual horizon that audio-only experiences do not offer to achieve balance, thus causing symptoms of nausea.

Increasing Rotational Distortion for All Groups: We note that the distraction ratio for V and AV participants was higher than in other systems (see Observation 2), which may impede task performance in vision-based applications. Thus, we recommend using a higher rotational distortion ratio closer to the imperceptible vision threshold of 30% [38] in the presence of vision rather than the 15-20%, threshold for audio-only experiences [26], especially since we conclude that experience quality and performance do not appear to differ between groups V and AV (see Observation 3).

Effects of Audio Cues in Navigation around VEs: Study 2 provides evidence that RDW with distractors may be sufficient for navigational tasks such as obstacle avoidance. Our OA module resulted in approximately 2/3 of obstacles being successfully avoided (see Observation 10), and we suspect that this success rate could have been increased by relying on a learning effect between levels as in Study 1. As shown in Figure SM9 in Supplemental Material, the participants were supported through multiple paths to the goal, implying a versatility to our OA solution.

Generalizability: While our study implies the possibility of using RDW for VI users, such user populations would require a separate evaluation and cannot be assumed to behave the same as sighted users with simulated blindness, as in our Group A. The results of our study imply the possibility of VI applications after a number of future steps. Additionally, the use of distractors may limit possible applications of any RDW system, and while [29] implies that distraction should not significantly affect user perception of the VE, tasks requiring great concentration would not benefit from this system. However, the small size of commercial tracking areas may still require distraction.

6 Conclusion

In this paper, we report on two studies that evaluated the ability of audio-only RDW systems with audio distractors to enable audio-only users to complete complex navigational tasks in VR.

Our first study compared audio-only, vision-only, and audiovisual user groups' performance and the quality of their experiences in three increasingly complex scenarios. We found that performance of vision-only and audiovisual groups was similar to each other, while the audio-only group had lower distraction ratios and average walk speeds. We found those two varibles to be moderately to strongly correlated. Differences in sense of presence, detection of rotational distortion, and susceptibility to simulator sickness were not significantly different between conditions.

Our exploratory second study examined the performance of an audio-only group in an obstacle avoidance scenario. This study was performed using the same base audio-only RDW system as Study 1 except that we added an obstacle avoidance module. We found that with our obstacle avoidance framework, participants could successfully navigate through the scene avoiding 2/3 of the obstacles they encountered. This leads us to believe that VR systems using RDW and audio distractors have potential for training VI users.

The results from our two studies suggest that auditory distraction can be an effective technique for use in audio-only VR navigation tasks, and that distractors can be designed in a general way to handle both audio-only and vision-having user groups without modification.

Future Work: Although this early and exploratory work shows the potential of distraction as a method of extending audio-only VR experiences, there are still many avenues of future work. We did not design our questionnaire to evaluate user preference for redirection or reorientation with distractors between groups, so it would be worth evaluating if user preference agrees with our speculation that reorientation is more appropriate for audio-only users. Additionally, our rotational distortion threshold of 15% can be increased to the 20% found in [33], as we are now confident that 15% is already highly immersive but provides too high of a distraction ratio for vision users. Our obstacle avoidance framework can be optimized to handle more complicated obstacles and to precompute paths around the obstacles, which can lead to powerful training applications for audio-only user groups. Finally, we suspect that other locomotion methods that cause visual distortion such as [8] may work well for audio-only groups that would not notice visual imperfections, assuming that distortion in the audio field can be avoided. A future extension of this study on VI users will further verify its general applicability beyond the sighted population.

Acknowledgments

This project is supported in part by National Science Foundation and Elizabeth Stevinson Iribe Chair Professorship.

References

[1] M. Azmandian, T. Grechkin, M. Bolas, and E. Suma. Physical space requirements for redirected walking: how size and shape affect performance. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, pages 93–100. Eurographics Association, 2015.

[2] O. Bart, N. Katz, P. L. Weiss, and N. Josman. Street crossing by typically developed children in real and virtual environments. *OTJR: Occupation, Participation and Health*, 28(2):89–96, 2008.

[3] D. R. Begault, E. M. Wenzel, and M. R. Anderson. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society*, 49(10):904–916, 2001.

[4] B. Bolte, F. Steinicke, and G. Bruder. The jumper metaphor: an effective navigation technique for immersive display setups. In *Proceedings of Virtual Reality International Conference*, 2011.

[5] G. Bruder, P. Lubas, and F. Steinicke. Cognitive resource demands of redirected walking. *IEEE transactions on visualization and computer graphics*, 21(4):539–544, 2015.

[6] G. Bruder, F. Steinicke, K. H. Hinrichs, and M. Lappe. Reorientation during body turns. In *EGVE/ICAT/EuroVR*, pages 145–152, 2009.

[7] H. Chen and H. Fuchs. Towards imperceptible redirected walking: integrating a distractor into the immersive experience. In *Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, page 22. ACM, 2017.

[8] Z.-C. Dong, X.-M. Fu, C. Zhang, K. Wu, and L. Liu. Smooth assembled mappings for large-scale real walking. *ACM Transactions on Graphics (TOG)*, 36(6):211, 2017.

[9] D. Engel, C. Curio, L. Tcheang, B. Mohler, and H. H. Bülthoff. A psychophysically calibrated controller for navigating through large environments in a limited free-walking space. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 157–164. ACM, 2008.

[10] T. Feigl, E. Kõre, C. Mutschler, and M. Philippsen. Acoustical manipulation for redirected walking. 2017.

[11] M. Gonzalez-Franco, A. Maselli, D. Florencio, N. Smolyanskiy, and Z. Zhang. Concurrent talking in immersive virtual reality: on the dominance of visual speech cues. *Scientific Reports*, 7(1):3817, 2017.

[12] M. Gröhn, T. Lokki, and T. Takala. Comparison of auditory, visual, and audiovisual navigation in a 3d space. *ACM Transactions on Applied Perception (TAP)*, 2(4):564–570, 2005.

[13] E. Hodgson and E. Bachmann. Comparing four approaches to generalized redirected walking: Simulation and live user data. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):634–643, 2013.

[14] V. T. Inman, H. J. Ralston, and F. Todd. *Human walking*. Williams & Wilkins, 1981.

[15] C. S. Kallie, P. R. Schrater, and G. E. Legge. Variability in stepping direction explains the veering behavior of blind walkers. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1):183, 2007.

[16] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[17] E. Langbehn, P. Lubos, G. Bruder, and F. Steinicke. Application of redirected walking in room-scale vr. In *Virtual Reality (VR), 2017 IEEE*, pages 449–450. IEEE, 2017.

[18] J. J. LaViola Jr. Bringing vr and spatial 3d interaction to the masses through video games. *IEEE Computer Graphics and Applications*, 28(5), 2008.

[19] T. Lokki and M. Grohn. Navigation with auditory cues in a virtual environment. *IEEE MultiMedia*, 12(2):80–86, 2005.

[20] J. R. Marston, J. M. Loomis, R. L. Klatzky, and R. G. Golledge. Nonvisual route following with guidance from a simple haptic or auditory display. *Journal of Visual Impairment & Blindness*, 101(4):203–211, 2007.

[21] J. C. Middlebrooks and D. M. Green. Sound localization by human listeners. *Annual review of psychology*, 42(1):135–159, 1991.

[22] G. A. Miller. The masking of speech. *Psychological bulletin*, 44:105–129, 1947.

[23] T. Nescher, Y.-Y. Huang, and A. Kunz. Planning redirection techniques for optimal free walking experience using model predictive control. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 111–118. IEEE, 2014.

[24] C. T. Neth, J. L. Souman, D. Engel, U. Kloos, H. H. Bulthoff, and B. J. Mohler. Velocity-dependent dynamic curvature gain for redirected walking. *IEEE transactions on visualization and computer graphics*, 18(7):1041–1052, 2012.

[25] N. Nitzsche, U. Hanebeck, and G. Schmidt. Extending telepresent walking by motion compression. *Tagungsbeiträge*, 1:83–90, 2002.

[26] M. Nogalski and W. Fohl. Acoustic redirected walking with auditory cues by means of wave field synthesis. In *Virtual Reality (VR), 2016 IEEE*, pages 245–246. IEEE, 2016.

[27] T. C. Peck. *Redirected free exploration with distractors: A large-scale real-walking locomotion interface*. PhD thesis, The University of North Carolina at Chapel Hill, 2010.

[28] T. C. Peck, H. Fuchs, and M. C. Whitton. Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):383–394, 2009.

[29] T. C. Peck, H. Fuchs, and M. C. Whitton. Improved redirection with distractors: A large-scale-real-walking locomotion interface and its effect on navigation in virtual environments. In *Virtual Reality Conference (VR), 2010 IEEE*, pages 35–38. IEEE, 2010.

[30] S. Razzaque. *Redirected walking*. University of North Carolina at Chapel Hill, 2005.

[31] S. Razzaque, Z. Kohn, and M. C. Whitton. Redirected walking. In *Proceedings of EUROGRAPHICS*, volume 9, pages 105–106. Citeseer, 2001.

[32] J. Richard Skarbez Frederick P. Brooks and M. C. Whitton. A survey of presence and related concepts. *ACM Comput. Surv.*, 50(6), 2017.

[33] S. Serafin, N. C. Nilsson, E. Sikstrom, A. De Goetzen, and R. Nordahl. Estimation of detection thresholds for acoustic based redirected walking techniques. In *Virtual Reality (VR), 2013 IEEE*, pages 161–162. IEEE, 2013.

[34] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535):3549–3557, 2009.

[35] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 3(2):130–144, 1994.

[36] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and virtual environments*, 6(6):603–616, 1997.

[37] F. Steinicke and G. Bruder. Using perceptual illusions for redirected walking. *IEEE computer graphics and applications*, 33(1):6–11, 2013.

[38] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics*, 16(1):17–27, 2010.

[39] F. Steinicke, G. Bruder, T. Ropinski, and K. Hinrichs. Moving towards generally applicable redirected walking. In *Proceedings of the Virtual Reality International Conference (VRIC)*, pages 15–24. IEEE Press, 2008.

[40] E. A. Suma, G. Bruder, F. Steinicke, D. M. Krum, and M. Bolas. A taxonomy for deploying redirection techniques in immersive virtual environments. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pages 43–46. IEEE, 2012.

[41] E. A. Suma, S. Clark, D. Krum, S. Finkelstein, M. Bolas, and Z. Warte. Leveraging change blindness for redirection in virtual environments. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 159–166. IEEE, 2011.

[42] E. A. Suma, S. L. Finkelstein, S. Clark, and Z. Wartell. An approach to redirect walking by modifying virtual world geometry. In *Workshop*

on Perceptual Illusions in Virtual Environments, pages 16–18. Citeseer, 2009.

[43] Q. Sun, A. Patney, L.-Y. Wei, O. Shapira, J. Lu, P. Asente, S. Zhu, M. McGuire, D. Luebke, and A. Kaufman. Towards virtual reality infinite walking: dynamic saccadic redirection. *ACM Transactions on Graphics (TOG)*, 37(4):67, 2018.

[44] Q. Sun, L.-Y. Wei, and A. Kaufman. Mapping virtual and physical reality. *ACM Transactions on Graphics (TOG)*, 35(4):64, 2016.

[45] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking¿ walking-in-place¿ flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364. ACM Press/Addison-Wesley Publishing Co., 1999.

[46] B. N. Walker and J. Lindsay. Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice. *Human factors*, 48(2):265–278, 2006.

[47] F. L. Wightman and D. J. Kistler. Resolution of front–back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, 105(5):2841–2853, 1999.

[48] B. Williams, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr, J. Rieser, and B. Bodenheimer. Exploring large virtual environments with an hmd when physical space is limited. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 41–48. ACM, 2007.

[49] M. Zank and A. Kunz. Optimized graph extraction and locomotion prediction for redirected walking. In *3D User Interfaces (3DUI), 2017 IEEE Symposium on*, pages 120–129. IEEE, 2017.

A Supplemental Material

SSQ Level	Illness Category	User Group A $\bar{x}$	A σ	V $\bar{x}$	V σ	AV $\bar{x}$	AV σ
Living Room	Nausea	10.9	14.9	8.2	14.4	11.9	16.8
	Oculomotor	16.2	18.5	13.5	16.1	16.4	17.4
	Disorientation	30.8	39.1	15.9	17.1	24.4	26.6
	Total	20.6	23.1	14.2	16.0	19.3	20.3
Firedrill Scenario	Nausea	14.3	20.7	12.3	16.5	9.5	11.5
	Oculomotor	16.8	16.9	11.4	15.1	17.1	16.5
	Disorientation	28.8	33.4	13.9	14.4	24.4	26.6
	Total	21.6	24.2	14.2	15.6	18.7	17.0
Traffic Crossing	Nausea	25.2	31.0	8.9	14.2	12.7	15.9
	Oculomotor	26.0	25.0	11.6	15.0	20.2	22.0
	Disorientation	45.7	52.5	15.9	14.3	22.0	32.2
	Total	35.0	36.8	15.0	15.3	20.9	24.4

Figure SM1: Descriptive values for SSQ Data in **Study 1.**

Games-Howell Post-hoc Tests for SSQ (Repeated Measures)

Symptom	Comparison	p
Nausea	A vs V	0.557
	A vs AV	0.718
	V vs AV	0.952
Oculomotor	A vs V	0.547
	A vs AV	0.963
	V vs AV	0.739
Disorientation	A vs V	0.193
	A vs AV	0.641
	V vs AV	0.595
Total	A vs V	0.347
	A vs AV	0.771
	V vs AV	0.728

Figure SM2: Games-Howell (equal variance not assumed) post-hoc tests performed on Repeated Measures ANOVA for SSQ data in Study 1. We find **no significance** between groups in this comparison.

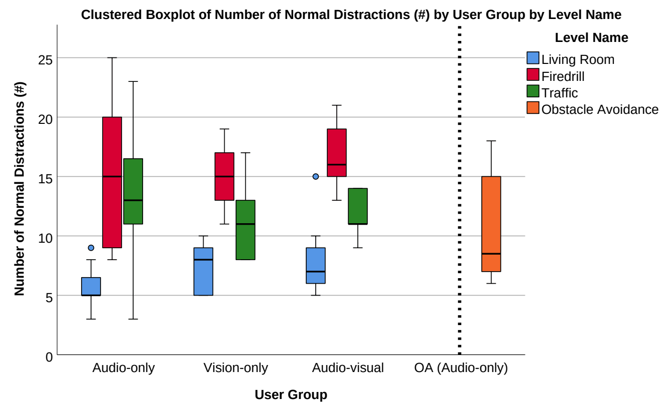

Figure SM3: Performance data for number of normal (#) distractions (distractor moving from goal to PE center and back one time). We only find a statistically significant effect between Groups A and V in L1, p($\alpha < 0.05$)=**0.065**.

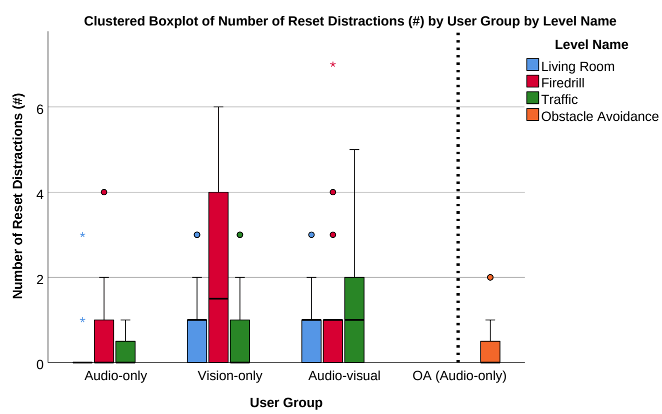

Figure SM4: Performance data for number of reset (#) distractions (distractor moving from center of PE to opposite side of userthreetimes). There is generally no statistically significant effect between groups except in L3 between A and AV (p($\alpha < 0.05$)=**0.043**).

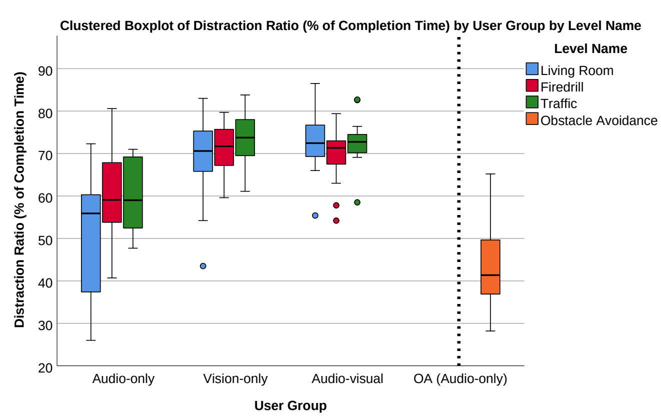

Figure SM5: Performance data for distraction ratio (% of completion time for which any distractor is active). Values are less consistent but generally much lower for Group A.

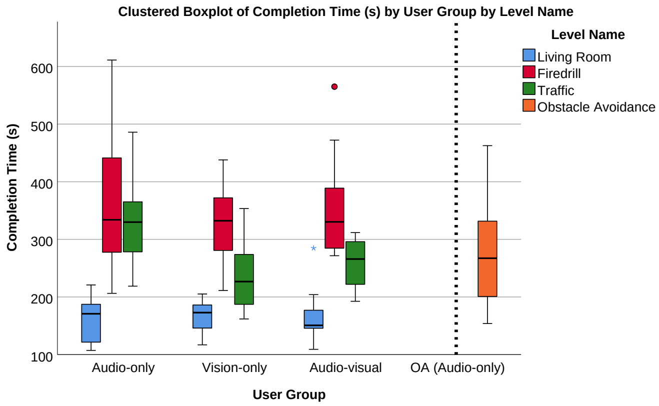

Figure SM6: Performance data for completion time (s). We only find a statistically significant effect in L3 between Group A and the others, with A taking much longer to finish (A vs V: p($\alpha < 0.05$)=**0.002**; A vs AV: p($\alpha < 0.05$)=**0.026**).

	Variable	User Group							
		A		V		AV		OA	
		x̄	σ	x̄	σ	x̄	σ	x̄	σ
Quality (Study 1 only)	L1 Detection*	0.40	0.51	0.50	0.52	0.36	0.50		
	L2 Detection*	0.67	0.49	0.36	0.50	0.64	0.50		
	L3 Detection*	0.80	0.41	0.36	0.50	0.64	0.50		
	Study 1 Presence^	0.67	0.49	0.79	0.43	0.71	0.45		
Performance — L1: Living Room	Normal Distractions (#)	5.67	1.59	7.43	1.74	7.71	2.53		
	Reset Distractions (#)	0.27	0.80	1.00	1.04	0.86	0.95		
	Distraction Ratio (%)	50	14	69	10	73	8		
	Average Speed (m/s)	0.37	0.18	0.55	0.17	0.59	0.17		
	Completion Time (s)	163	39	165	29	164	42		
Performance — L2: Firedrill Scenario	Normal Distractions (#)	15.33	6.20	14.93	2.46	16.64	2.27		
	Reset Distractions (#)	0.67	1.11	2.14	2.14	1.43	1.99		
	Distraction Ratio (%)	61	12	71	6	70	7		
	Average Speed (m/s)	0.47	0.14	0.54	0.12	0.50	0.10		
	Completion Time (s)	363	119	329	65.2	354	84		
Performance — L3: Traffic Crossing	Normal Distractions (#)	13.40	4.93	11.43	3.03	11.86	1.79		
	Reset Distractions (#)	0.27	0.46	0.79	1.12	1.36	1.60		
	Distraction Ratio (%)	60	9	73	6	73	6		
	Average Speed (m/s)	0.43	0.15	0.58	0.15	0.59	0.16		
	Completion Time (s)	327	82	238	63	261	42		
Performance — Obstacle Avoidance (S2)	Normal Distractions (#)							10.75	4.33
	Reset Distractions (#)							0.42	0.79
	Distraction Ratio (%)							44	11
	Average Speed (m/s)							0.31	0.10
	Completion Time (s)							276	89
	Obstacles Avoided (#)							1.9	0.9
	Obstacles Hit (#)							1.0	0.9

* Values near 1 indicate high chance of noticing rotational distortion

^ Values near 1 indicate high sense of presence based on mini-SUS

Figure SM8: Quality and performance values recorded for all groups. The Obstacle Avoidance level refers to the group in **Study 2**, which only completed that one level. Quality metrics were not recorded for **Study 2.**

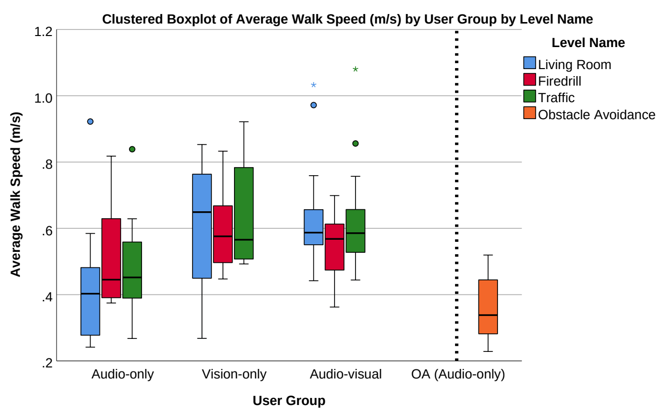

Figure SM7: Performance data for average walk speed (m/s). Group A generally walked more slowly, which we find correlates to lower distraction ratio.

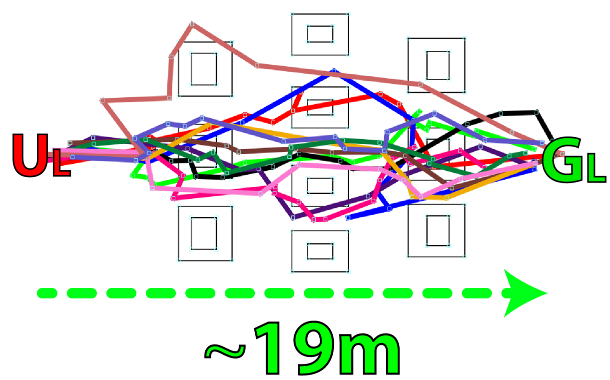

Figure SM9: Smoothed paths taken by the Participants in Study 2 (Group OA). The inner boxes are the furniture meshes; the outer boxes are buffer zones. Participants move from left to right.

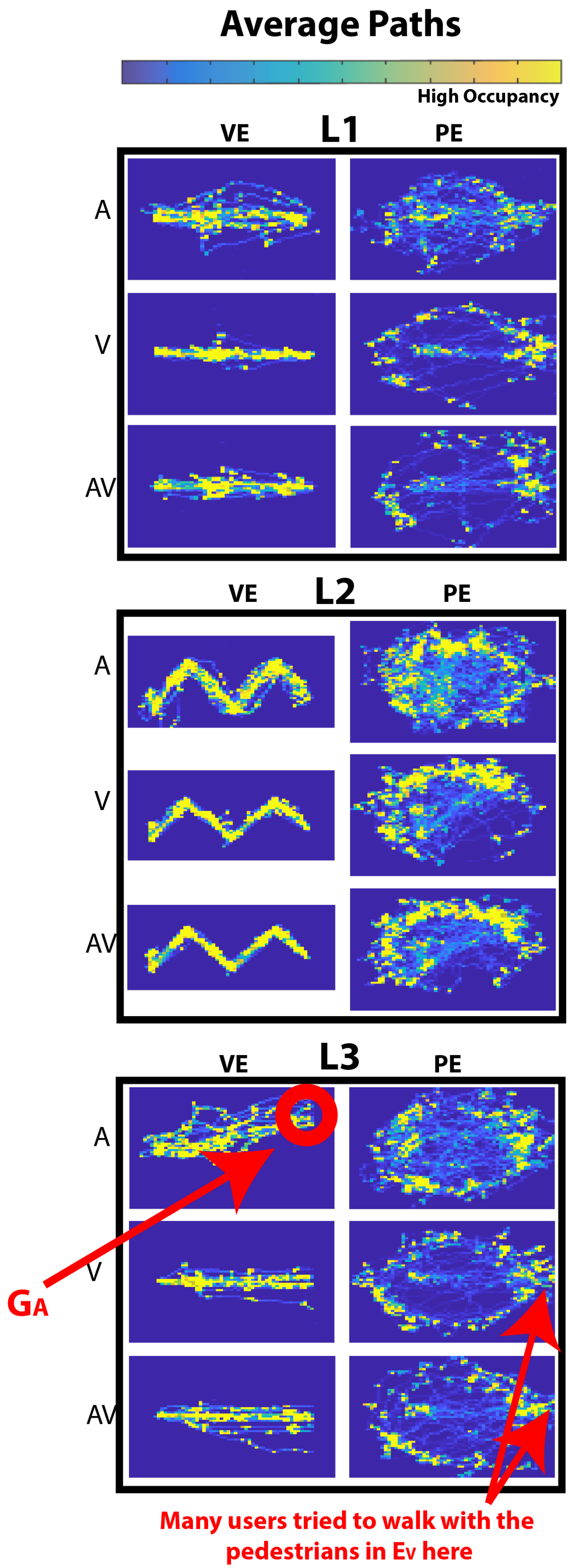

Figure SM10: Heatmap of typical paths taken by Participants.

Demographic	Tested Variable	Significance	
		F	p
Gender	Normal Distractions	0.034	0.855
	Reset Distractions	0.134	0.715
	Distraction Ratio	0.017	0.895
	Average Speed	0.113	0.737
	Completion Time	2.244	0.137
VR Experienc	Normal Distractions	1.111	0.294
	Reset Distractions	0.320	0.573
	Distraction Ratio	0.189	0.665
	Average Speed	2.443	0.121
	Completion Time	1.134	0.289

Figure SM11: MANOVA testing for statistical significance in gender and amount of VR experience on performance metrics in **Study 1**. We find no significance in either variable.

	Stimulus\Level	Living Room	Firedrill	Traffic	OA
A	G_A	man talking	man talking	chirping beeper	
	G_V				
	G_L	opposite side of room	each button in sequence	opposite sidewalk	
	Distractor Audio	dog panting	dog panting	dog panting	
	Distractor Video				
	E_A		machinery ambience, touched button cue	people talking, footsteps, car engines, honking, fast beeper	
	E_V				
V	G_A				
	G_V	green question mark	green question mark		
	G_L	opposite side of room	each button in sequence	opposite sidewalk	
	Distractor Audio				
	Distractor Video	red robot dog	red robot dog	red robot dog	
	E_A				
	E_V	environment, Leap hands	environment, Leap hands	walking people, moving/stopped vehicles, environment, Leap hands	
AV	G_A	man talking	man talking	chirping beeper	
	G_V	green question mark	green question mark		
	G_L	opposite side of room	each button in sequence	opposite sidewalk	
	Distractor Audio	dog panting	dog panting	dog panting	
	Distractor Video	red robot dog	red robot dog	red robot dog	
	E_A	machinery ambience, touched button cue	people talking, footsteps, car engines, honking, fast beeper	temporary goal noise	
	E_V	environment, Leap hands	environment, Leap hands	walking people, moving/stopped vehicles, environment, Leap hands	
OA	G_A				man talking
	G_V				
	G_L				opposite side of room
	Distractor Audio				dog panting
	Distractor Video				
	E_A				temporary goal noise
	E_V				

Figure SM12: Table of audio and visual stimuli presented to each user group.