Joint Virtual Reality Conference of EGVE - EuroVR (2013)
B. Mohler, B. Raffin, H. Saito, and O. Staadt (Editors)

# Visual Attention to Wayfinding Aids in Virtual Environments

J. Bertrand, E. Ebrahimi, A. Wachter, J. Luo, S. V. Babu,[†] A. T. Duchowski, N. Meehan, A. K. Gramopadhye

Clemson University, USA

**Abstract**

*In an empirical evaluation, we examined participants' visual attention allocation to a dynamic wayfinding map in a complex simulation meant to educate medical practitioners in a hand hygiene protocol. Complex virtual environments (VEs) are novel types of virtual worlds that embody large spaces, interactive virtual humans, static and dynamic virtual entities, and intricate tasks that simulate real-world settings. Previous investigations of wayfinding aids have focused on the evaluation of spatial orientation, knowledge acquisition, and usage. We employed an eye tracker and created visualization tools to quantitatively and qualitatively analyze participants' visual attention to the wayfinding aid in our simulation. Results suggest that the proportion of time of gaze, total gaze count, and gaze transitions between various elements of the VE are altered with the use of the wayfinding aid. Participants also tend to employ innovative visual strategies in order to efficiently plan routes and accomplish tasks in the VE.*

Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Information Interfaces and Presentation]: Multimedia Systems—Artificial, Augmented, and Virtual Realities I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques J.4 [Computer Applications]: Social and Behavioral Sciences—Psychology

## 1. Introduction

As research in virtual environments (VEs) matures, many virtual worlds are becoming increasingly complex with scenarios consisting of large spaces, multiple interactive virtual characters, and users performing difficult cognitive and psychomotor tasks in a simulated setting. Many of these simulations have been developed for the purpose of replicating real-world situations and events for training in specialized cognitive and psychomotor skills [QLF*09, BGC*09], therapy and rehabilitation [HAB*01], simulated social interactions [BLB*02, RJD*06], practicing military exercises [HHG*03], and exploration [Bro99]. Transfer of knowledge gained in interactive virtual environments has been shown to be directly applicable to real world settings [RG08]. In these applications, VEs, rendered in either head-mounted or on large screen displays, provide simulated content in life-size dimensions that users interact with in a natural manner.

Navigation in VEs consists of two components: travel (or locomotion), i.e., moving from one location to another, and wayfinding, which is the cognitive task of spatial knowledge acquisition and usage. Different travel metaphors exist in virtual worlds [ZLB*06]. Researchers have designed, implemented, and evaluated various navigational aids to facilitate spatial perception and cognition, to generally enhance wayfinding in large-scale virtual environments. However, there is little or no prior work examining visual behavior, via eye gaze analysis, to examine quantitative and qualitative aspects of attentional allocation to various elements in complex virtual environments. Eye movements are important to consider when examining the effects of an interaction aid in such environments, as they provide insights related to higher cognition and task performance, such as visual search, attention, coordination of motion, and information usage.

Darken et al. suggest that large-scale environments with wayfinding aids promote enhanced spatial orientation and minimize disorientation [DS96]. Using a map provides a simultaneous geocentric perspective augmenting one's egocentric perspective and enabling the use of environmental cues such as prominent landmarks [TG83]. We thus hypothesize that in VEs that involve large spaces with static and dynamic virtual entities, such as landmarks and autonomous agents, wayfinding aids should facilitate visual search, efficient exploration, spatial orientation, and task performance.

We have previously examined users' performance of a cognitive task in a large complex virtual reality medical sim-

[†] sbabu@clemson.edu



DOI: 10.2312/EGVE.JVRC13.009-016

ulation [GBB*12], and found that, although they were able to learn the simulated task of hand hygiene protocols in the VE, participants were often lost and disoriented, and did not explore the VE as much as we expected. Therefore, we theorize that integrating wayfinding aids in these complex environments should alleviate the cognitive load associated with exploration and other simultaneous tasks. Thus, we expect to observe differences in visual search, attention, and task performance through quantitative and qualitative analysis of visual gaze, with and without a wayfinding aid in a VE.

## 2. Related Work

Navigational aids in virtual worlds have been studied extensively, but few studies have examined users' allocation of visual attention to them. Several metaphors have been proposed for wayfinding in virtual environments. Darken et al. compared the effects of flying above the VE, use of spatial audio markers, visual markers (breadcrumbs), grid navigation, and two metaphors of map views of the world on spatial orientation and wayfinding [DC99]. Results showed that participants preferred a variety of tools for wayfinding, however, the map-based navigational aid was effective in the formation of a cognitive map of the environment.

Ruddle et al. investigated the effects of landmarks on route learning and other spatial cognitive abilities [RPJ97]. They observed that time to task completion improved in environments that contained landmarks. Stoakley et al. used a World-In-Miniature (WIM) to facilitate navigation in large scale virtual environments that are rendered in an head-mounted display (HMD) or CAVE [SCP95]. Users could teleport to a new position in the virtual world by selecting that location in the miniature hand-held world. However, this approach is not suitable as a navigational aid in VEs rendered on a single, large-screen display.

Wu et al. compared three different wayfinding aid metaphors, namely a view-in-view map (inset map in the VE), animated guide, and a human system collaborator for effectiveness, efficiency, and satisfaction [WZZ09]. To help participants find targets in a large-scale virtual world, the view-in-view wayfinding aid was the most helpful. They also noted that the perceived usefulness of the wayfinding aid varied between people with differing spatial abilities.

A wayfinding map is an effective interaction metaphor in facilitating spatial knowledge acquisition and route learning in large scale virtual environments. Prabhu et al. have examined the effect of different configurations of the wayfinding 2D map in an inset view in a navigation task [PSDH96]. They found that a static 2D map with a top-down view of the environment, with a dynamic you-are-here pointer, was most effective in route learning and spatial orientation.

Based on this review, we chose an enhanced wayfinding map with a static view of the environment with a dynamic you-are-here pointer to indicate the position and orientation of the viewer. Our wayfinding aid also includes dynamic annotations of the positions of virtual entities in the VE.

Ruddle et al. have proposed three metrics for evaluating wayfinding aids in virtual environments: (1) users' task performance via time taken, distance traveled, and number of errors committed, (2) physical behavior (locomotion, camera/head rotations, time and error classification), and (3) cognitive effects, such as decision making and questionnaires [RJD*06]. They used after-action review tools to analyze paths taken by participants, and automatically coded the travel strategies via spaghetti plots, with distance traveled and degree of head rotations. In this study, we have extended these visual analytical techniques to include analysis of visual attention allocation to the wayfinding aid in a VE.

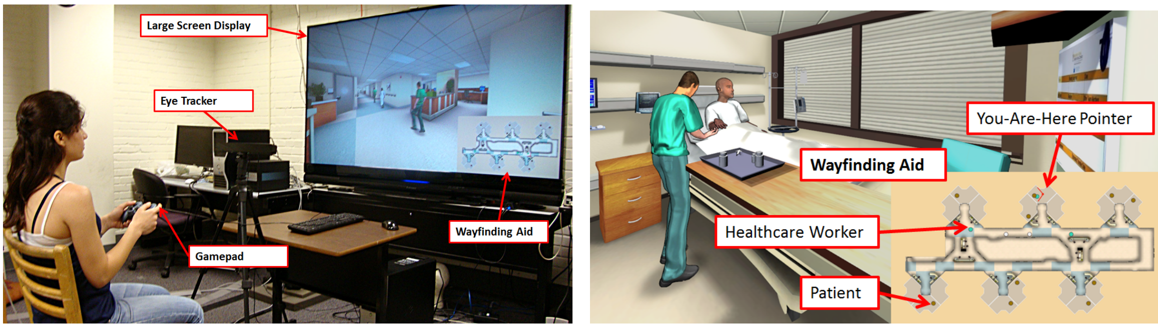
To understand the effects of frequently viewed landmarks on spatial knowledge acquisition, Hamid et al. measured gaze of participants while they learned to navigate a large-scale virtual environment [HSH10]. They found that removal of frequently viewed landmarks (at the end of hallways and intersections) resulted in degradation of performance. This is important as prior research suggests that landmark knowledge enhances route learning in virtual worlds [CWRD07].

Earlier, we examined participants' gaze to a wayfinding map and its benefits to users' navigation in a 3D virtual maze shown in an HMD [VID*04]. We examined different configurations of the wayfinding aid: a 2D map, a 2D map with a directionally ambiguous cue, and a 2D map with a directional you-are-here cue. Eye tracking data was collected and analyzed to examine the correlation between the map configuration and the navigational efficiency of users through the virtual maze. The directional cue was the most effective in analysis of time to complete the task of reaching the center the maze. However, the VE and tasks used lacked complexity: no dynamic entities were present in the VE and the task consisted of finding one's way to the maze center. No comparison was made to a map-less representation of the VE.

We build on previous research by examining visual attention allocation to a wayfinding aid in a complex virtual environment. Environment and task complexity ensue from task demands involving visual search, procedure learning, environment navigation (travel), and interaction with elements populating the VE. The latter includes static and dynamic virtual humans in a realistic simulation of a real-world hospital ward. To the best of our knowledge, this research is among the first to compare gaze allocation to the wayfinding aid in such a dynamically rich VE.

## 3. Experiment Simulation

Our VE is a multi-agent simulation used to teach and train medical practitioners in the US Center for Disease Control (CDC)'s Five Moments of Hand Hygiene [GBB*12]. Using the VE's scenario definition capabilities, we have created a pre-defined experiment scenario with autonomous virtual




*(a) Experimental apparatus.* *(b) VE screenshot.*

**Figure 1:** *Experimental setup: (a) large display with desktop eye tracker; (b) view of the VE with the wayfinding aid.*

healthcare workers walking about and interacting with virtual patients in hospital rooms (see Figure 1(a)). The hospital environment is a large-scale, life-sized virtual ward, a replica of the Roy Carver wing of the University of Iowa Hospital. Participants in the simulation play the role of a healthcare inspector, first learning the hand hygiene protocol [BP08], then traveling through the ward recording as many protocol infringements as witnessed in a given time limit.

Participant's tasks consisted of virtual travel (via a gamepad) through the hospital environment, navigating through narrow and wide corridors, in *movement space*, avoiding other virtual healthcare workers and static objects (i.e., nurse stations and medical equipment), and searching for virtual healthcare workers performing various interactive tasks in patient rooms (i.e., inoculating the patient, taking a tissue sample, or measuring blood pressure etc.). When in patient rooms, participants had to observe the actions of the healthcare worker in this *interaction space*, recording any infringements of the five different types of hand hygiene violations in the CDC's protocol [BP08]. At the end of the simulation, each participant was provided feedback on the accuracy of her hand hygiene inspection task. Technical details of the creation of the large-scale VE, scenario generation of gross motions and fine actions of the virtual agents, the hand hygiene recording interface, and scoring functions are reported elsewhere [GBB*12]. The VE was created using Blender for modeling, and Unity3D Pro 4.0 for rendering.

A wayfinding map was implemented within the VE as an inset (view-in-view metaphor) in the rendered view of the virtual environment. The wayfinding map was located in the bottom right-hand corner and occupied $1/9^{th}$ of the view window. The dynamic map was created by placing a secondary camera in the scene with a top-down static view of the entire environment consisting of hospital rooms, corridors, and objects (e.g., nurse station). The wayfinding map was a 2D abstract decomposition of the 3D world. Colored annotations provided positional information of virtual healthcare workers and patients (green for nurses, white for doctors, and brown for virtual patients). A red dynamic pointer was rendered to illustrate the participant's position in the virtual environment, with two lines emanating from the marker denoting the viewer's orientation (see Figure 1(b)).

## 4. Experimental Procedure & Design

Twenty-seven nursing students at Clemson University participated in the study after providing informed consent. Participants were randomly assigned to one of two conditions in a between-subjects design: VE with wayfinding aid (VE-W), or VE without wayfinding aid (VE-NW). We also characterized differences in aspects of gaze behavior between movement and interaction space as a within-subjects variable.

### 4.1. Purpose

Using an eye tracker, the study aimed to determine how users of a VE interact via gaze with virtual agents, environment, and wayfinding aids, in both movement and interaction space, while performing a challenging cognitive task. We asked the following research questions:

1. What are the quantitative differences in gaze count, total viewing time, and gaze transitions between agents, environment, and the region of screen where the wayfinding aid is present, between conditions VE-W and VE-NW?
2. What are the qualitative differences in participants' visual strategies when using the wayfinding aid in the VE?
3. What are the quantitative and qualitative gaze differences between movement and interaction space in the VE?

### 4.2. Materials and Apparatus

Participants were seated in a chair 80″ away from an 73.2″ (6.1′) × 48.7″ (4′) Mitsubishi large screen non-stereoscopic display with a resolution 1900 × 1080 (1080p; see Figure 1(a)). An S2 Mirametrix eye tracker was set up in front of the participant, on a tripod, at a height that did not obstruct the participants' view of the display. The S2 tracks both eyes over a limited range of head motion (25×11×30 cm volume), sampling at 60 Hz. According to the manufacturer, point-of-gaze accuracy is provided within 0.6° visual angle. Participants were instructed to sit upright in the chair and to attempt to keep their head still throughout the duration of the study. Participants were given a gamepad controller to navigate through the VE and tag hand hygiene infringements. The computational platform that hosted the VE was a Dell Precision workstation with a quad core processor and dual NVIDIA Quadro FX 5600 SLI graphics cards.

### 4.3. Procedure

The procedure was composed of the following steps:

1. *Pre-trial:* Upon arrival, participants completed an informed consent form and three pre-experiment questionnaires: (a) the Guildford-Zimmerman (GZ) Spatial Orientation test; (b) the Visual Memory (MV-2) test; and (c) the Card Rotation (S-1) test of mental rotation. Participants were assigned to a condition (VE-W/VE-NW).
2. *Calibration:* Participants were seated and asked to keep their heads as still as comfortable. The eye tracker was adjusted so that it could image both eyes without obstructing the display. The VE was started and set to the screen resolution. The eye tracker was calibrated, with participants asked to follow a blue dot as it moved between 9 fixed screen coordinates. Calibration was repeated until the eye tracker sampled the participant's eyes at 7 of the 9 points below an error threshold.
3. *Training:* Instruction was given on the CDC's hand hygiene protocol [BP08], how to navigate the VE, and how to tag any hand hygiene infringements made by the virtual healthcare agents. Participants were instructed to follow a designated agent through the hallways into a VE patient room, where the agent failed to wash his hands before and after administering a shot to the patient. Feedback was given on participants' ability to notice this infringement and use of the tagging interface to record it. If participants had no more questions, the experimenter pressed a key to start a timer to begin the trial.
4. *Experiment trial:* Participants had 15 minutes to move freely through the VE to tag any observed infringements.
5. *Post-trial:* At the trial conclusion, the number of correctly tagged observations was displayed out of a total number of possible observations; participants were debriefed.

## 5. Quantitative and Qualitative Measures

When the simulation starts, the eye tracker sends data packets to the VE. At each rendering cycle of the VE, the left and right gaze information is retrieved from the most recently received tracker packet. The tags for retrieved information are: left $x$, left $y$, right $x$, right $y$ and an error metric. The simulation writes data to three output files at each rendering frame: (a) a data file with a timestamp, and the participant's position and orientation in the VE, (b) an agent file with the same information at each frame, but for each virtual healthcare worker found in the scene, and (c) a gaze file with a timestamp, and left and right gaze point in normalized screen coordinates. Each participant's performance data (e.g., number of tagged infringements) are also stored in the data file.

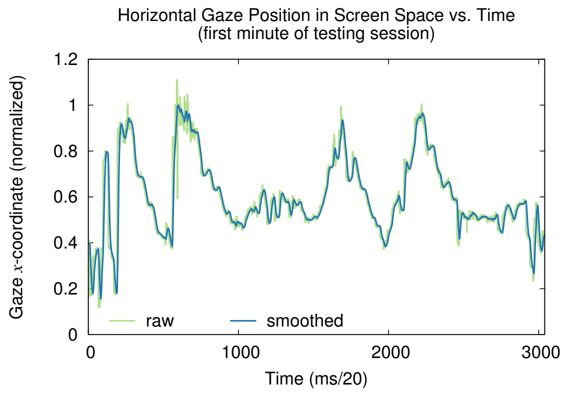



**Figure 2:** *Differences between raw (green) and smoothed (blue) gaze position over a short time interval.*

The eye tracker outputs a stream of screen gaze points $(x_i, y_i)$. Typically, this data is noisy and requires smoothing. Treating $x_i$ or $y_i$ independently, smoothing or differentiating (to order $s$) is achieved via convolution with a $2m+1$ point filter $h_i^{t,s}$ at midpoint $i$:

$$f_n^s(t) = \frac{1}{\Delta t^s} \sum_{i=-m}^{m} h_i^{t,s} x_i, \quad \text{or} \quad f_n^s(t) = \frac{1}{\Delta t^s} \sum_{i=-m}^{m} h_i^{t,s} y_i, \qquad (1)$$

where $n$ and $s$ denote the polynomial fit to the data and its derivative order, respectively [Gor90, OD12]. Based on prior work, we chose a $2^{nd}$ order Butterworth filter to smooth the raw gaze data with sampling and cutoff frequencies of 60 and 1.65 Hz, respectively [DPHW11] (see Figure 2).

## 6. Visual Analysis and Automatic Coding Tool

We created a visual analysis and automatic coding tool (VAACT) to extract the gaze dependent and independent quantitative variables from the smoothed gaze data. Additionally, for qualitative evaluation, experimenters can play back a participant's raw and smoothed gaze sequence through the entire duration of the experiment. The gaze data is displayed as a series of colored circles whose degree of transparency is based on the positive and negative amount of time from the current frame. The radius of the circles is set to 1.75° visual angle. The time range enables analysis of where participants looked over a 1 second timespan. A GUI allows the experimenter to toggle between raw, smoothed, and fixation-classified gaze. The playback can be rendered at 1×, 2×, or 4× the framerate of the VE simulation. Using

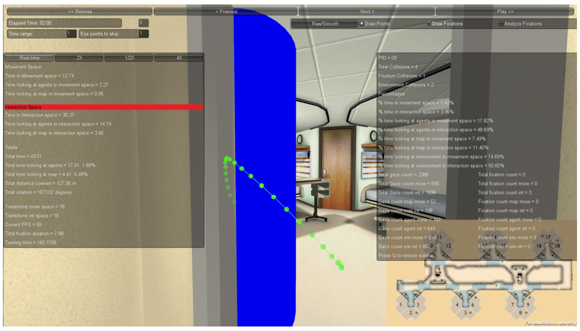


**Figure 3:** *VAACT screenshot showing frame where smoothed gaze data intersects virtual agent's bounding volume cylinder (highlighted in blue).*

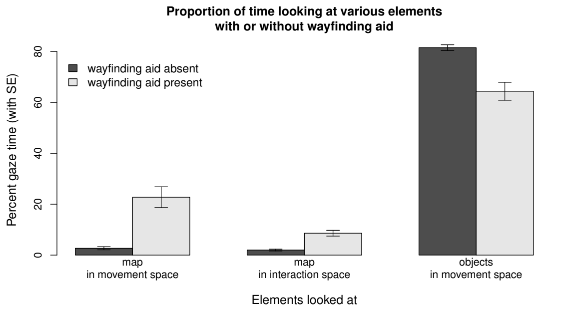



**Figure 4:** *Mean proportion of gaze time over map or environment in movement or interaction space, with or without wayfinding aid.*

VAACT, experimenters (who were blind to the condition experienced by the participant) reviewed the participant's data file to examine and code gaze behavior of participants with and without the wayfinding aid (a wayfinding aid was rendered in both conditions in VAACT), and the subject's gaze in interaction and movement spaces. The VAACT was implemented in Unity3D Pro (see Figure 3).

The VAACT reads a participant's log files which include their position and orientation, that of agents, as well as raw, and smoothed gaze data. Dependent variables are shown on a translucent layer on top of the visualization, continuously updating in real-time in order to enable visual verification of the accuracy of the aggregated values. These include number of gaze points, proportion of gaze, gaze transitions, and total time looking at elements of the VE (agents vs. objects vs. wayfinding aid). These variables are grouped by whichever space the participant occupied in the VE (movement vs. interaction). Autonomous agents are rendered abstractly as cylinders in the VAACT. Gaze intersections with agents, objects, or wayfinding aid are computed via ray casting, and highlighted in the VACCT to allow experimenters to code the event appropriately.

Other events such as frontal collisions between the viewer and environmental objects such as desks, chairs, and autonomous agents are also automatically coded using the underlying Unity3D physics engine. The viewers' total distance covered and view frustum rotations are also automatically coded by the VAACT and stored for statistical analysis.

## 7. Results and Analysis

Of the 27 participants, 23 were considered for data analysis. Four participants were excluded due to loss of eye tracking data during the experiment session. This resulted in 13 participants in the VE-NW condition and 10 in VE-W. Overall no significant differences were found between Spatial Orientation (GZ), Visual Memory (MV-2), or Mental Rotation scores between participants in the VE-NW and VE-W conditions. We did not find any significant differences in gaze independent performance or spatial cognition including collisions, total view rotations, total distance traveled, hand hygiene observation scores or the number of observations relative to the presence of the wayfinding aid in the VE.

We also conducted a correlation analysis of variance-covariance between participants' spatial ability, mental rotation and visual memory scores and variables related to visual attention and performance. We did not find any significant correlations of interest. In the following sections, we highlight the significant effects we found in the pairwise comparisons of the variables related to visual attention.

### 7.1. Between-Subjects VE-NW and VE-W

An independent samples *t*-test revealed that participants in the VE-W condition ($M = 174.62$ s, $SD = 84.97$) spent a significantly longer amount of time gazing at the map than participants in the VE-NW condition ($M = 19.97$ s, $SD = 12.60$), $t(21) = 6.51, p < 0.001$. This suggests that when participants are presented with a wayfinding aid, they are likely to visually attend to it during the simulation.

**Proportion of Time Analysis.** An independent samples *t*-test was used to test the differences in percentage of time spent gazing at the agents, objects, and the wayfinding map, with and without the wayfinding aid. In the absence of the aid (the VE-NW condition), we compared participants' gaze towards the screen region where the wayfinding aid would be present (the VE-W condition). This allowed us to compare intentional goal-oriented visual attention to the map, versus natural gaze drawn towards that portion of the screen.

Participants in the VE-W condition ($M = 24.65\%$, $SD = 14.83$) spent a larger proportion of time looking at the map in movement space than participants in the VE-NW condition ($M = 2.6\%$, $SD = 2.36$), $t(21) = 5.31, p < 0.001$. Participants

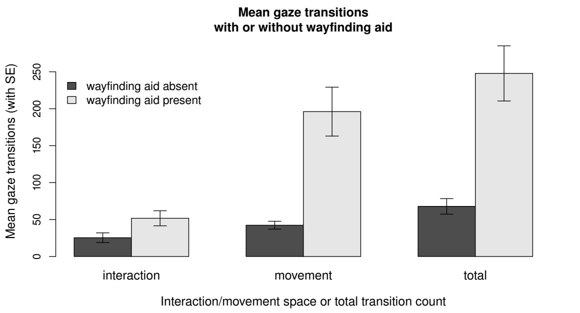



**Figure 5:** *Mean number of (smoothed) gaze transitions in movement and interaction space.*

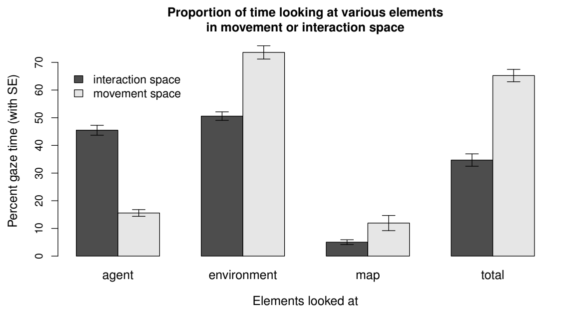



**Figure 6:** *Mean proportion of gaze time over elements or environment in movement or interaction space.*

in the VE-W condition also ($M = 9.06\%$, $SD = 4.14$) spent a larger proportion of time looking at the map in interaction space than participants in the VE-NW condition ($M = 1.83\%$, $SD = 1.34$), $t(21) = 5.93, p < 0.001$. Participants appeared to spend a larger proportion of time gazing at the wayfinding aid in movement space, where they usually perform navigation oriented tasks, than in interaction space where they performed non-spatial cognitive tasks (see Figure 4).

Participants also spent a significant proportion of time gazing at the environment in the absence of the wayfinding aid ($M = 81.57$, $SD = 4.48$), than those with the aid present ($M = 61.10$, $SD = 10.29$), $t(21) = -6.45, p < 0.001$.

**Gaze Count and Gaze Transition Analysis.** Gaze count denotes the number of times a participant's smoothed gaze point fell on either the map, agents or objects, at every frame of the VE. In movement space, an independent samples *t*-test showed that participants in the VE-W condition ($M = 8624$, $SD = 5030$) issued significantly more gazes toward the wayfinding aid than those in the VE-NW condition ($M = 776$, $SD = 569$), $t(21) = 5.62, p < 0.001$. In interaction space, participants in the VE-W condition ($M = 1738$, $SD = 1211$) also issued significantly more gazes to the wayfinding aid than those in the VE-NW condition ($M = 387$, $SD = 376$), $t(21) = 3.81, p = 0.001$. In interaction space, participants in the VE-NW condition ($M = 8550$, $SD = 2011$) produced a significantly larger gaze count towards virtual agents than participants in the VE-W condition ($M = 7006$, $SD = 1036$), $t(21) = -2.20, p = 0.039$. With no wayfinding aid, participants tend to pay more attention towards virtual agents when they are performing the goal-oriented task of observing healthcare workers for violation of the hand hygiene protocol. In movement space, participants in the VE-NW condition ($M = 27380$, $SD = 6065$) produced a significantly larger gaze count towards environment objects than those in the VE-W condition ($M = 21917$, $SD = 5460$), $t(21) = -2.23, p = 0.036$.

Gaze transitions are the number of switching movements of eye gaze from the virtual environment or agents to the wayfinding aid and vice versa (see Figure 5). There was a significantly larger number of gaze transitions between the map region in the VE-W condition ($M = 278.2$, $SD = 118.27$) than in the VE-NW condition ($M = 67$, $SD = 41$), $t(21) = 6.022, p < 001$. Gaze transition is more frequent to the wayfinding aid when it is present than to the same screen region when it is absent. Participants in the condition VE-W ($M = 219.5$, $SD = 110.8$) showed significantly higher number of transitions to the map region in movement space, than those in the VE-NW condition ($M = 41.76$, $SD = 20.66$), $t(21) = 5.69, p < 0.001$. Participants in VE-W ($M = 58.70$, $SD = 33.86$) also showed significantly more transitions in interaction space to the map, than those in the VE-NW condition ($M = 25.23$, $SD = 25.65$), $t(21) = 2.70, p = 0.013$. More gaze transitions occur to the map region when the wayfinding aid is present in both movement and interaction spaces. However, the number of transitions in interaction space appears less than the number of transitions in movement space.

### 7.2. Within-Subjects Movement and Interaction

In this section we compare within-subjects quantitative gaze differences between movement and interaction space.

**Proportion of Time Analysis.** Participants in interaction space ($M = 46.29\%$, $SD = 7.61$) spent a significantly larger proportion of time looking at virtual agents than they did in movement space ($M = 16.23\%$, $SD = 6.06$), $t(21) = -16.28, p =< 0.001$ (see Figure 6). Participants in movement space ($M = 12.19\%$, $SD = 14.76$) spent a significantly larger proportion of time looking at the map than in interaction space ($M = 4.97\%$, $SD = 4.62$), $t(21) = 2.93, p = 0.007$. Participants also spent a significantly larger proportion of time looking at environment objects in movement space ($M = 72.67\%$, $SD = 95.93$) than in interaction space ($M = 49.93\%$, $SD = 6.48A$), $t(21) = 8.83, p < 0.001$.

**Total Gaze and Transition Count Analysis.** Participants in interaction space ($M = 7879$, $SD = 1805$) issued a significantly larger number of gazes toward virtual agents than participants did in movement space ($M = 5240$, $SD = 2340$), $t(21) = -3.84, p < 0.001$. Participants in movement space ($M = 4188$, $SD = 5133$) produced a significantly larger number of gazes directed at the map than participants in interaction space ($M = 974$, $SD = 1071$), $t(21) = 3.32, p = 0.003$. Participants also showed a significantly larger number of gazes directed at the environment in movement space ($M = 25005$, $SD = 6319$) than in interaction space ($M = 9364$, $SD = 3701$), $t(21) = 8.17, p < 0.001$. Overall, we found that the total number of gaze transitions between the wayfinding aid and the rest of the environment was significantly larger in movement space ($M = 119$, $SD = 115.6$) than in interaction space ($M = 39.78$, $SD = 33.4$), $t(21) = 3.64, p = 0.001$.

## 8. Qualitative Analysis

Using the VAACT, researchers who were blind to the participant's condition, walked through the participants' data, and examined the visual attentional and behavioral strategies. The following lists some key findings of the qualitative analysis of participant's behaviors categorized by condition.

*VE-NW—No Wayfinding Aid:*

- Most participants followed virtual agents they met in the corridors outside the patient rooms and "shadowed" them for some time. Even if they were walking in circles, participants tended to follow them until they realized that it was not worthwhile to follow the same agent further.
- Some participants tended to hover in the alcoves outside the patient rooms waiting for another virtual agent to visit the patient rooms nearby, reluctant to walk around and visit patients without spatial awareness.
- Other participants walked about the virtual ward and took a peek into patient rooms to observe a virtual healthcare worker performing a procedure on a virtual patient. In this manner of trial-and-error the participants noticed a virtual healthcare worker in a patient room and then recorded potential infringements of the hand hygiene procedure.

*VE-W—With Wayfinding Aid:*

- Some participants tended to stop in the corridors, gaze at the map, plan where the nearest rooms were located with virtual agents visiting patients, and then went there to perform the hand hygiene observation task. Those participants preferred to stop travel, acquire spatial knowledge from the wayfinding aid, then travel to a destination.
- Some participants showed evidence of the "Cell Phone Effect"—they walked through the VE mostly watching the wayfinding aid rather than paying attention to objects in the environment and avoiding oncoming virtual agents. For this subset of participants, collision with other agents and objects in the environments was high. They seemed distracted by the presence of the wayfinding aid, and this behavior seemed to hinder travel performance.
- Still others attended to the VE and oncoming agents in movement space, and tended to look at the wayfinding aid mostly outside alcoves to check if healthcare workers were present in nearby rooms. This resulted in an efficient use of the aid as a quick reference, facilitating attention to landmarks and dynamic virtual entities. Some checked the wayfinding aid just as they were about to leave a patient room (in interaction space), and then traveled directly to an area of interest while paying attention to environment cues rather than the wayfinding aid during travel.

## 9. Discussion

Gaze counts and the proportion of time looking at the wayfinding map in interaction and movement spaces were significantly higher when the map was present than when it was absent, i.e., not surprisingly, in both instances, the map is used when available. Overall, the proportion of time gazing at the environment is significantly higher when the map is unavailable. The number of gaze transitions to the region where the map is located is higher when the map is there than when it is not, meaning that attention to the map is goal-directed and not random. Our visual analysis tools showed that users use the map by quickly glancing back and forth between it and the environment to gain spatial awareness, to enable more efficient path planning.

Interestingly, the proportion of time gazing at virtual agents in movement space was significantly less than in interaction space. Qualitatively, when participants travel between rooms, they tend to pay more attention to the wayfinding aid (if available), or are distracted by other artifacts in the environment, and pay less attention to oncoming virtual humans. However, they tended to look more at virtual agents when they are in interaction space performing tasks.

Qualitative evaluation using the visual analysis tool revealed interesting users' visual attentional strategies. The wayfinding aid allowed users to check for virtual healthcare agents in nearby patient room alcoves and then to plan navigational routes to those rooms. When on their way to those locations, some used the wayfinding aid to quickly glance back and forth between the map and the environment to enable them to travel more efficiently. Without the wayfinding aid, navigational strategies were often reduced to shadowing virtual agents, sometimes walking in circles.

When navigating (in movement space), overuse of the wayfinding aid (as measured by a smaller proportion of gaze allocated to virtual agents than the map) sometimes led to collisions between users and virtual agents (the "Cell Phone Effect"). When observing agents (in interaction space), although some gaze was allocated to the wayfinding aid, it was largely ignored, since it is not particularly relevant to the given task. Concomitantly, we can suggest the following design guideline: wayfinding aids can potentially be made adaptive, such that they are made largely transparent when not needed, e.g., when the system can detect the user engaged in a task other than travel (locomotion), or when no gaze is directed towards the screen region where the map is typically displayed, i.e., a gaze-contingent wayfinding aid.

## 10. Conclusions and Future Work

We conducted a comparative evaluation to examine the differences in visual attention to a wayfinding aid in a large scale complex Virtual Environment. We also created a visual analysis tool (VAACT) to automatically code the smoothed visual gaze data of participants with (VE-W) and without a wayfinding aid (VE-NW). Our framework provides an empirical testbed for thorough quantitative and qualitative examination of visual attentional within the environment under differing interactive conditions. Our framework enables scientific investigation of the interplay between cognitive task performance, navigation, visual search, and visual attention that occurs in complex virtual environments.

Future work will focus on two key directions: (1) analysis of visual attention to the wayfinding aid via classification of gaze into fixations, saccades, and smooth pursuits; and (2) empirical evaluation of the differences in visual attention to various types of wayfinding aids in VEs such as "breadcrumbs", dynamic landmarks, and worlds-in-miniature.

## 11. Acknowledgments

This work was supported in part by the US National Science Foundation under Award #1104181.

## References

[BGC*09] Babu S. V., Grechkin T., Chihak B., Ziemer C., Kearney J., Cremer J., Plumert J.: A virtual peer for investigating social influences on children's bicycling. In *IEEE VR* (2009), pp. 91–98. 1

[BLB*02] Blascovich J., Loomis J., Beall A. C., Swinth K. R., Hoyt C. L., Bailenson J. N.: Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry 13* (2002), 103–124. 1

[BP08] Boyce J., Pittet D.: Guidelines for hand hygiene in health-care settings: Recommendations of the healthcare infection control practices advisory committee. *Infection Control 23* (2008), 3–41. 3, 4

[Bro99] Brooks F. P.: What is real about virtual reality? *IEEE Computer Graphics & Applications 19* (1999), 16–27. 1

[CWRD07] Cliburn D., Winlock T., Rilea S., Donsel M. V.: Dynamic landmark placement as a navigational aid in virtual worlds. In *ACM VRST* (2007), pp. 211–214. 2

[DC99] Darken R., Cevik H.: Map usage in virtual environments: Orientation issues. In *IEEE VR* (1999), pp. 133–140. 2

[DPHW11] Duchowski A. T., Pelfrey B., House D. H., Wang R.: Measuring Gaze Depth with an Eye Tracker During Stereoscopic Display. In *Applied Perception in Graphics & Visualization (APGV)* (Toulouse, France, August 27-28 2011), ACM. 4

[DS96] Darken R., Sibert J. L.: Wayfinding strategies and behaviors in large scale virtual worlds. In *ACM SIGCHI* (1996), pp. 142–149. 1

[GBB*12] Gupta M., Bertrand J. W., Babu S. V., Polgreen P., Segre A. M.: An evolving multi-agent scenario generation framework for simulations in preventive medicine education. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (New York, NY, 2012), IHI '12, ACM, pp. 237–246. 2, 3

[Gor90] Gorry P. A.: General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method. *Analytical Chemistry 62*, 6 (1990), 570–573. 4

[HAB*01] Hodges L., Anderson P., Burdea G., Hoffman H., Rothbaum B.: Treating Psychological and Physical Disorders with VR. *IEEE Computer Graphics & Applications 21*, 6 (2001), 25–33. 1

[HHG*03] Hill R. W., Hill A. W., Gratch J., Marsella S., Rickel J., Swartout W., Traum D.: Virtual humans in mission rehearsal exercises. In *Workshop on AI Embodied Conversational Agents* (2003), vol. 17, pp. 32–38. 1

[HSH10] Hamid S. N., Stankiewicz B., Hayhoe M.: Gaze patterns in navigation: Encoding information in large-scale environments. *Vision Journal 10*, 12 (2010), 1–11. 2

[OD12] Ouzts A. D., Duchowski A. T.: Comparison of Eye Movement Metrics Recorded at Different Sampling Rates. In *Eye Tracking Research & Applications (ETRA)* (Santa Barbara, CA, March 28-30 2012), ACM. 4

[PSDH96] Prabhu G. V., Shalin V. L., Drury C. G., Helander M.: Task-Map Coherence for the Design of in-Vehicle Navigation Displays. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (1996), vol. 40, pp. 882–886. 2

[QLF*09] Quarles J., Lampotang S., Fischler I., Fishwick P., Lok B.: Technical section: Scaffolded learning with mixed reality. *Computer Graphics 33*, 2 (2009), 34–46. 1

[RG08] Reger G., Gahm G.: Virtual reality exposure therapy for active duty soldiers. *Journal of Clinical Psychology 64*, 8 (2008), 940–946. 1

[RJD*06] Raij A., Johnson K., Dickerson R., Lok B., Cohen M., Bernard T., Oxendine C., Wagner P., Lind D. S.: Interpersonal scenarios: Visual approximate real? In *IEEE VR* (2006), pp. 59–66. 1, 2

[RPJ97] Ruddle R., Payne S., Jones D.: Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experiences. *Journal of Experimental Psychology 3*, 2 (1997), 143–159. 2

[SCP95] Stoakley R., Conway M. J., Pausch R.: Virtual reality on a wim: interactive worlds in miniature. In *SIGCHI* (1995), pp. 265–272. 2

[TG83] Thorndyke P. W., Goldin S. E.: Spatial Learning and Reasoning Skill. In *Spatial Orientation*, Pick Jr., Herbert L. and Acredolo, Linda P., (Ed.). Plenum Press, New York, NY, 1983, pp. 195–217. 1

[VID*04] Vembar D., Iyengar N., Duchowski A. T., Clark K., Hewitt J., Pauls K.: Effect of Visual Cues on Human Performance in Navigating Through a Virtual Maze. In *EuroGraphics Symposium on Virtual Environments (EGVE)* (Grenoble, France, June 8-9 2004), EuroGraphics. 2

[WZZ09] Wu A., Zhang W., Zhang X.: Evaluation of wayfinding aids in virtual environments. *IJHCI 25*, 1 (2009), 1–25. 2

[ZLB*06] Zanbaka C., Lok B., Babu S., Ulinski A., Hodges L. F.: Comparison of Path Visualizations and Cognitive Measures relative to Travel Technique in a Virtual Environment. *IEEE Transactions on Visualization and Computer Graphics 11*, 6 (2006), 694–705. 1