The Influence of Shading, Display Size and Individual Differences on Navigation in Virtual Reality

Lonnie Hastings
Simon Fraser University & The Boeing Company
2810 160th Ave SE
Bellevue, WA, USA
bhasting@sfu.ca

Bernhard E. Riecke
Simon Fraser University
250-13450 102nd Avenue
Surrey, BC, Canada
Ber1@sfu.ca

Abstract

Despite extensive usability research on virtual reality (VR) within academia and the increasing use of VR within industry, there is little research evaluating the usability and benefits of VR in applied settings. This is problematic for individuals desiring design principals or best practices for incorporating VR into their business. Furthermore, the literature that does exist often doesn't account for the characteristics of intended users. This shortage is problematic because individual differences have been shown to have a significant impact on performance in spatial tasks. The research presented here is an evaluation of a VR system in use at The Boeing Company, with 28 employees performing navigation and wayfinding tasks across two shading conditions (flat/smooth) and two display conditions (desktop/immersive). Performance was measured based on speed and accuracy. Individual difference factors were used as covariates. Results showed that women and those with high spatial orientation ability performed faster in smooth shading conditions, while flat shading benefited those with low spatial ability particularly for the navigation task. Unexpectedly, immersive presentation did not improve performance significantly. These results demonstrate the impact of individual differences on spatial performance and help determine appropriate tasks, display parameters, and suitable users for the VR system.

CR Categories: I.3.7 [Three-Dimensional Graphics and Realism]: Virtual Reality; J.4 [Social and Behavioral Sciences]: Psychology; J.6 [Computer-aided Engineering]: Computer-aided manufacturing.

Keywords: Virtual reality, navigational search, shading, individual differences

1 Introduction

It is clear from the exponential growth of virtual reality, that it offers a promising medium for innumerable applications; particularly those that require the acquisition or utilization of spatial knowledge. So, imagine for a moment that you work for a company looking to incorporate virtual reality technology into the business and you have been tasked with choosing and implementing the system. How do you choose the display (large, small, stereo, non-stereo, tracking or no)? Or the interaction device (mouse, wand, keyboard, joystick)? What about the rendering technology? How do you determine which tasks and for which users virtual reality would be a better alternative to a standard desktop approach?

There has been extensive research in this field within academia, however for many questions there are still no unequivocal answers. Take for example the question of whether or not large displays are superior for spatial tasks. Several studies looking at the impact of display size on navigation and distance estimation across multiple display types found no significant effect [Riecke et al. 2009; Kasik et al. 2002; Swindells and Po 2004]. In Kasik [2002], they were interested in determining "whether a larger screen would help a pre-trained user perform a common production task in less time than using a standard 20-inch monitor". The participants, Boeing Commercial Airplane Engineers, performed navigation and wayfinding tasks across three different displays (20" monitor, vision station, and 40" plasma panel). A follow up study by Swindells & Po [2004], used the same data and tasks, but different display conditions (standard desktop monitor, a tiled wall display, and an immersive room (i.e. CAVE-style) environment). Results and discussion from both studies supported the hypothesis that other influences, aside from the display size, may actually have a greater influence on user performance for spatial tasks. In contrast, a series of experiments by Tan et al. [2006] found that physically large displays significantly increased performance on navigation, mental map formation, and spatial memory tasks [Tan et al. 2006]. Furthermore, a study looking at the impact of display size and FOV on gender differences found that larger displays resulted in lower pointing errors, and faster performance for women in particular [Czerwinski et al. 2002]. These ambiguities offer support for the idea that the impact of display type is largely task and user dependent. This provides incentive for evaluations that are application and use case specific. The study presented here is an example of this kind of evaluation, comparing a VR system in use today at the Boeing Company to a standard desktop system.

While previous research has shown that people are capable of acquiring accurate spatial knowledge from virtual environments, the rate of learning and accuracy of performance is almost always inferior to real world performance [Lessels 2005]. It is often assumed that this inferior performance can be attributed to reduced fidelity of VR systems in terms (among others) of visual realism. As a result, there is mounting interest in developing "high fidelity" visualization techniques. The issue of whether or not sophisticated rendering techniques/technologies are worth their computational cost is met with similar ambivalence as display

SAP 2014, August 08 – 09, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-3009-1/14/08 $15.00

size. High fidelity techniques such as ray tracing have been shown to facilitate some spatial tasks; such as, route retracing, scene recognition, and vection [Meijer et al. 2009; Schulte-Pelkum et al. 2003; Slater et al. 2009; Wallet et al. 2011]. Lower fidelity conditions on the other hand, have been shown to be better for distance estimation, and "remembering" [Mania et al. 2006; Waller and Knapp 2001]. As new graphics devices and techniques are coming to market, full geometric complexity and more sophisticated photorealistic rendering techniques are now within reach. It is important to understand what benefit these new techniques provide, particularly when they come at a high computational cost [Dietrich and Wald 2005]. However, there is little consistency within the field in terms of rendering techniques used and tasks performed, which makes it difficult to draw conclusions about the benefit of one technique over another. Additionally, research varies widely in terms of what constitutes "high fidelity". The term is used to describe everything from the addition of landmarks to photorealistic rendering. Furthermore many researchers look at real world performance after a period of virtually learning, making it difficult to evaluate the impact of high fidelity on virtual navigation performance, which is used more often in industry application. These deficiencies suggest the need for further evaluations on how to display VEs on immersive displays within context for their intended use, for virtual navigation applications. To this end, the current study measured the comparative impact of two rendering methods, flat shading and smooth/Gouraud shading, on virtual navigation performance. This shading comparison was chosen as a measure of fidelity because, relative to other measures, it is a simple step in the direction of higher fidelity and because the VR system being tested provides flat and smooth shading as easy export options for users, therefore making it a comparison that represents a real world decision that people make when using the system.

Another source of ambiguity is the lack of research looking at environmental variables as well as individual differences and their relative impacts on performance in spatial tasks. Studies that have looked exclusively at individual characteristics have identified a number of variables that correlate highly with performance on spatial tasks. For example, large-scale spatial ability [Waller 1999], self-reported sense of direction [Hegarty et al. 2002], gender [Lawton 1994], and familiarity [Thorndyke and Hayes-Roth 1982]. However, these are rarely measured in conjunction with environmental variables and generally measure spatial tasks such as pointing and distance estimation rather than navigation. Those that have compared the two sources of influence have found that the impact of environment and design variables is largely task dependent, and that these kinds of individual differences are a major source of variation in both real world and computer related spatial tasks [Bryant 1982; Hegarty et al. 2006; Wolbers and Hegarty 2010]. Differences between participants in some cases were large enough to make finding other significant effects difficult [Waller 1999]. Despite the high degree of variance that has been attributed to individual differences, the vast majority of VR research continues to focus on physical aspects of the VR interface, or individual characteristics with only little research measuring both sources of influence or the interaction between the two. To this end, the study presented here was a within subjects design measuring spatial ability (using the Guilford-Zimmerman Spatial Orientation Test), gender, familiarity and attitude towards computers. These measures were subsequently used as covariates in the analysis to disambiguate the relative effects of environmental factors (display type and rendering method) and individual differences.

While there is a wealth of previous literature evaluating different design aspects of VR interfaces, there is little consistency in the field in terms of environments tested, interfaces used, or tasks performed. Additionally, there is little work reported concerning usability evaluation for intended users and validation of VR systems for specific tasks within industry settings. Given this, it would be difficult for an individual to determine the relative benefits of VR over a standard desktop approach, or even decide on a specific VR setup for their particular situation.

It is this void that the research presented here is intended to help address. The purpose of this study is to provide an applied use-case study of an immersive VR system currently in use at The Boeing Company. The system is used to visualize aircraft geometry for design and review sessions for at least five major visualization tasks [Kasik et al. 2002; Swindells and Po 2004]:

1. Finding an object
2. Inspecting an object for discrepancies, overlaps, conformity, and interference.
3. Visually scanning scenes
4. Tracing paths, typically through animation, to detect dynamic interference conditions
5. Comparing objects from different design releases to better understand design preference.

The specific goal of this study was to determine how individual differences, visual realism and display type impact performance on two spatial tasks, navigation and wayfinding, that are typical of those performed in review sessions. Both tasks contain aspects of finding an object (1), inspecting an object (2), and visually scanning scenes (3). While the application of the results will be specific to Boeing and to this system, the lessons learned can and should be more generally applicable. For example, information about how individual characteristics impact performance in virtual environments can be used not only to motivate use case studies within companies but also to determine job assignments, targeted training programs, and allocation of resources. This research is also intended to demonstrate the importance of performing use case evaluations for industry and research users of virtual reality by showing the impact of user and system characteristics for one particular system.

Explicitly, the three research questions were:

Q1: What is the effect of display condition (in this case desktop + keyboard/mouse vs. immersive display + wand) on navigation and wayfinding performance in a virtual environment?

Q2: What is the effect of visual fidelity (flat shading vs. smooth/Gouraud shading) on navigation and wayfinding performance in a virtual environment?

Q3: What role do individual differences, specifically prior experience and spatial ability, have on their navigation performance in a complex, geometric virtual environment?

Due to the similarity in tasks and stimuli, we expected the results in regards to display type to be in line with both the Kasik et al. [2002] and Swindells et al. [2004] studies which found that display type itself had no significant effect. However, we did expect to find interactions between display type and individual differences. Specifically, engineers and other expert users who have many years of computer experience with programs like CAD/CATIA, were expected to perform better under familiar conditions; i.e. the desktop + flat shading condition.

In regards to fidelity, since route level knowledge is necessary for successful navigation and wayfinding and high fidelity visuals have been shown to facilitate the acquisition of route level knowledge [Wallet et al. 2011], we expected to find that people without previous experience/bias towards low fidelity (flat shaded) models, would perform better under smooth shaded (high fidelity) conditions.

2 Methods

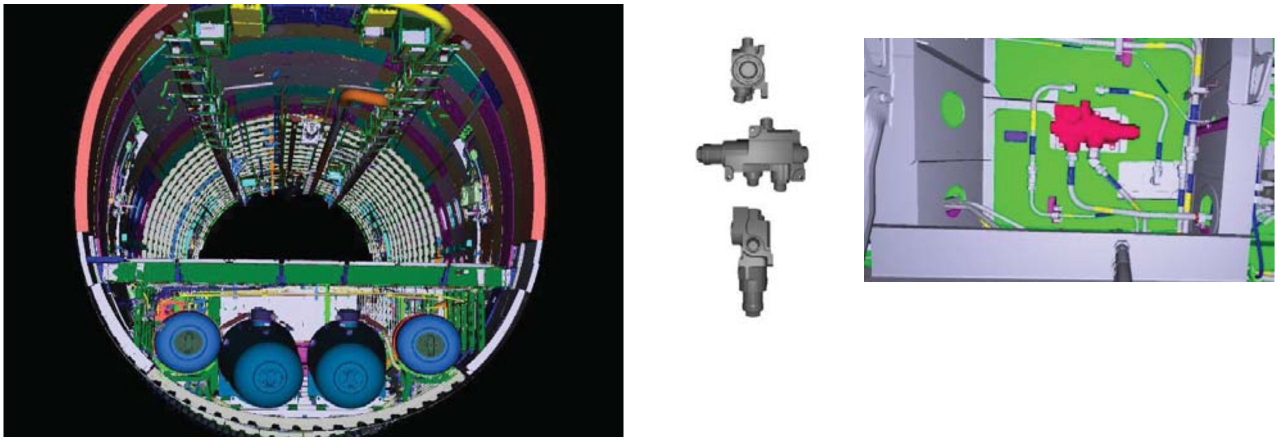

Figure 1 *Example of starting position (left), target reference images (middle), target within context (right)*

2.1 Participants

28 Boeing Employees (17 male, 11 female) were recruited via poster and referrals from other participants. Participants intentionally came from a wide range of backgrounds including: engineering, IT, finance, administration, and manufacturing. No prior experience was required to participate in the study, and $10 Tully's cards were given in exchange for participation.

2.2 Materials

2.2.1 Software

Eight test datasets and two practice datasets were created using internally developed Boeing software called Integrated Visualization Tool (IVT). IVT is a 3D visualization tool that is used to view, manipulate and analyze large quantities of CAD data. In this case it was used to rapidly query against large amounts of engineering data to produce datasets that represented *partially complete* sections of the 787 Dreamliner airplane. This software also allows users to export data in flat shaded or smooth shaded form. This functionality was used to create the two shading conditions for the experiment. All 8 test datasets were subsections of the 787-8 Dreamliner, while the practice datasets were subsections of the 777. This was done as a precaution to ensure that participants were learning the navigation technique but not learning the environmental layout of the test environment during practice. These datasets were then imported into IC:IDO Visual Decision Platform v.9.0.1 which was used to visualize the airplane sections on either a standard desktop display or on a large immersive display described in detail below.

2.2.2 Hardware

The IC:IDO visualization software was run on a Dell Precision Workstation with a NVIDIA Quadro 4000 graphics card, 14GB RAM, and a 2GB Video RAM for both desktop and immersive conditions.

For the desktop conditions, participants viewed datasets on a non-stereo 22' Dell Desktop Monitor with 1680 × 1050 resolution. The Immersive conditions were displayed stereoscopically on the IC:IDO immersive display (approx. 57" × 75") using 2 Christie Digital 1280 × 1024 resolution projectors. Simulated field of view was held constant between display conditions. Desktop navigation was performed using a standard mouse and keyboard setup. The immersive navigation was performed using a wireless, 20 × 10 cm, 6 degree of freedom wand from ART called the FlyStick2. The displays were setup so that the immersive display was just to the left of the desktop and both display conditions were run off of the same workstation.

2.3 Navigation Mode

For both the immersive and desktop conditions, the IC:IDO system was set to the "Fly" navigation mode. In this mode, desktop condition translations (left, right, forward, and backwards) were controlled via the arrow keyboard keys, and the heading or "camera position" was changed by clicking and dragging the mouse in the desired direction. Users could also change their navigation speed using the plus (+) and minus (-) keys. For the immersive condition, users navigated by pressing and holding the left FlyStick2 button with their thumb while moving their hand in the desired direction. For example, to move straight forward with no rotation, the user would hold the button with their thumb and push the wand straight forward on a level plane. Rotation was achieved by rotating the entire wand and hand to the left or right while keeping the center of rotation of the hand on a level plane. A small "cursor" with a speed vector appeared on the screen to show the user where the FlyStick was pointing at any given time, where the center of rotation was, as well as how quickly they were accelerating.

2.4 Procedure

The study used a 2 Display Condition (Immersive/Desktop) × 2 Shading (Flat/Smooth) within-subjects design. Each participant was run separately through one session taking anywhere from 60-80 minutes. Each session began with a consent form and informed consent discussion, followed by a Computer Use Questionnaire [Waller 1999], and an online version of the Guilford-Zimmerman Spatial Orientation (GZ-SO) test [Guilford and Zimmerman 1981; Kyritsis and Gulliver 2009]. The remainder of the study consisted of four conditions (Desktop-Flat, Desktop-Smooth, Immersive-Flat, and Immersive-Smooth).

Display and shading conditions were counterbalanced between participants to control for learning and carry over effects. After completion of the GZ-SO, before beginning the experimental tasks participants were given verbal instructions on how to use the interaction device for whichever display they were starting on.

After these verbal instructions were given, participants had five minutes to practice the technique on a sample dataset before beginning the experimental tasks. After the experimental tasks were completed a short, informal debriefing discussion was held in which the experimenter explained the purpose of the study and asked for feedback.

2.5 Computer Use Questionnaire

A computer use questionnaire, from Waller [1999], was used as a measure of participant's prior experience with computers. Items on the questionnaire assess both a person's attitude toward computers and prior experience. Previous analyses have shown that these scales represent two associated factors of computer use [Waller 1999]. Both factors were then measured as the average score on items from their respective scales. While the target variable from this questionnaire was participant's experience with computers, both scales were used as covariates in the analysis to assess their significance.

2.6 Guilford Zimmerman-Spatial Orientation

The Guilford Zimmerman Spatial Orientation test was designed to measure what the authors described as "an ability to appreciate spatial relations with reference to the body of the observer" [Guilford and Zimmerman 1981]. This includes the awareness of whether one object is to the right or left, nearer or farther, etc. Unlike other psychometric tests of spatial ability, the GZ-SO has been shown to predict performance in large scale spaces [Infield 1991].The test requires the examinee to look at two pictures and imagine that he/she is riding in a boat whose prow is visible in front of them in the first scene, along with other reference objects that give information as to the current position. In the second picture the boat has changed its position, and the goal of the participant is to determine how the boat has moved. The online version has an answering system slightly different from the original paper based test however, developers conducted a study validating that the pattern of Spatial Orientation scores would be the same regardless of the medium [Kyritsis and Gulliver 2009].

2.7 Where's Waldo Navigational Search Task

Participants began each test condition with a navigational search task called Where's Waldo. In this task, participants were given a clipboard with grayscale images of a target object from three different orientations (front, top, left)[1]. They were then positioned at either the front or back of the airplane, and instructed that they had five minutes to find the target part.

The target was positioned in its normal location, orientation, and context and the experiment guaranteed that the target would appear in the scene. There were no restrictions placed on the method or path of navigation. Timing began when the participant indicated that they were ready. If the participant failed to find the target within the allotted time, their final completion time was listed as five minutes. If the participant identified an incorrect part as the target, an error was recorded and they were allowed to continue to look for the correct part for whatever time remained. The task ended when the subject had identified the correct part or the five minutes had expired.

2.8 Hansel and Gretel Wayfinding Task

The second task was a wayfinding task called "Hansel and Gretel." The task began with the participants viewing a short clip (60-90 seconds), of the experimenter navigating indirectly to a target part. At the end of the clip, the participant was then virtually placed back at the start location (front or back of the airplane section) and the task from there was for the participant to find their way back to the target they had just been shown. Participants were instructed that they did not need to follow the same path as shown in the clip. The same "Fly" navigation mode was used for both tasks. The task ended when the subject had identified the correct part or the five minutes had expired.

3 Results

The two tasks were analyzed separately in two 2 (display) × 2 (shading) repeated measures ANCOVAs in SPSS v.17.0. Computer Experience, Attitude towards Computers, gender, and Spatial Ability Score (GZ-SO) were used as covariates. Participants were measured for both completion time and accuracy, however there were very few "misses" and therefore very little accuracy data to analyze. A trial was considered a "miss" if the participant misidentified a target part during a task. Out of 224 trials only 5 misses were recorded, in other words less than 2% of the trials were misses. Due to the lack of data, only results for completion time will be reported.

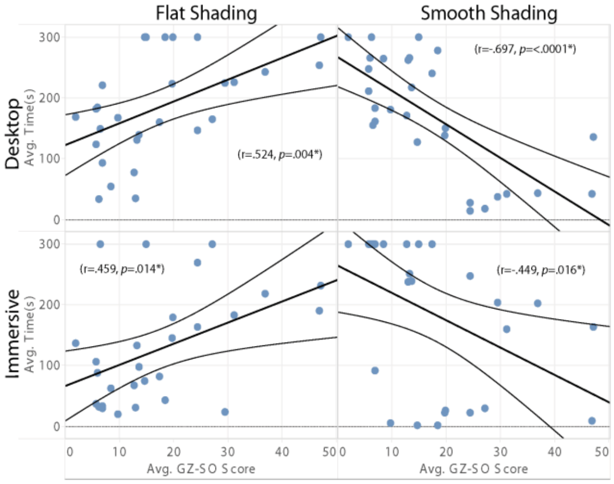

Figure 2 *Correlation between completion time and spatial ability (GZ-SO score) (with confidence intervals) for all four display × shading conditions for the Where's Waldo task*

Q1: What is the effect of display size on navigation and wayfinding performance in a virtual environment? The results, as expected showed that there was no significant main effect of display type on navigation or wayfinding performance. However, interestingly, results from debriefing discussions showed that while the immersive display did not facilitate performance, 27 out of the 28 participants still preferred it over the desktop.

Q2: What is the effect of visual fidelity on navigation and wayfinding performance in a virtual environment?

The ANCOVA for the Where's Waldo navigational search task revealed a significant main effect of shading ($F(1, 23) = 4.15$, $p = .05$, $\eta_p^2 = .153$). Participants were significantly faster in the flat shading condition ($M = 156.5$, $SE = 17.4$) than in the smooth shading condition ($M = 178.2$, $SE = 20.7$).

Q3: What role do individual differences, specifically prior experience and spatial ability, have on their navigation performance in a complex, geometric virtual environment?

---

[1] Gray scaling prevented subjects from inferring location from color, a convention used to encode different airplane systems

For the Where's Waldo navigational search task, significant interaction effects were found between Shading × Spatial Ability (F(1, 23) = 11.62, p = .002, η_p^2 = .336), Shading × Gender (F(1, 23) = 8.25, p = .009, η_p^2 = .264), and (Display × Shading × Spatial Ability F(1, 23) = 14.90, p = .001, η_p^2 = .393).). Post hoc analysis of each of these interactions revealed that participants with high spatial ability performed better in the smooth shading conditions and that those with low-to-medium spatial ability performed better in the flat shading conditions, this trend is shown in Figure 2. This pattern was true for both display types. Contrary to our hypothesis, computer experience was not a significant covariate with shading. Looking further into the interaction between shading and gender showed that women performed significantly better in smooth shading conditions while the opposite pattern was true for men.

For the Hansel and Gretel wayfinding task, there were three significant interactions: Shading × Spatial Ability (F(1, 23) = 7.73, p = .011, η_p^2 = .252), Display × Shading × Spatial Ability (F(1, 23) = 4.74, p = .040, η_p^2 = .171), Display × Shading (F(1, 23) = 4.63, p =.042, η_p^2 = .168). Post Hoc analysis revealed that participants performed significantly faster in the flat shading condition on the immersive display, shown in Figure 3, and slower in the flat shading condition for the desktop. The three way interaction between Display × Shading × Spatial Ability revealed that this pattern was particularly true for those with high spatial ability.

Interestingly, the three way interaction between Display × Shading × Spatial Ability was significant for both the Where's Waldo and the Hansel and Gretel Task. However, there is an interesting inverted pattern between the two tasks. For the Hansel and Gretel task, high spatial ability correlated with faster completion times across all display and shading conditions. In the Where's Waldo task however, high spatial ability had a positive correlation with completion time. In other words those with high spatial ability actually took longer under these conditions.

There was no significant correlation between spatial ability and completion time on either display for the smooth shading conditions. There was no difference between males and females in terms of self-reported attitude or experience towards computers. Males overall had a slightly higher average on the GZ-SO (M = 18.0, SE = 3.0) than women (M = 16.8, SE = 3.4), however this trend did not reach significance. For women, there was a significant overall negative correlation between GZ-SO score and completion time (r(9) = .66, p = .025).

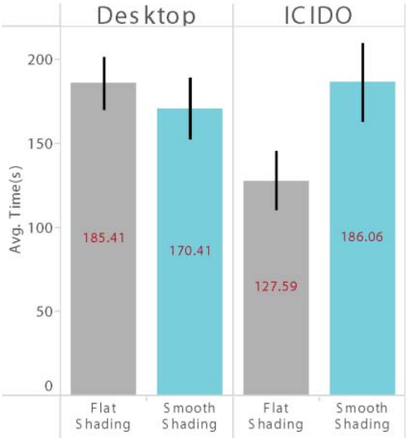

Figure 3 *Significant interaction between Display × Shading F(1,23)=4.63, p=.042, ηp2=.168) for the Hansel and Gretel task, with error bars displaying +- 1 standard error of the mean*

4 Discussion

The two tasks performed in this study were developed by Boeing subject matter experts to represent standard tasks performed in design review sessions and were meant to evaluate a person's ability to find an object by relying on visual reference and visual memory respectively. There were several hypotheses this research hoped to address. The first was that engineers and other "expert" users would perform best under familiar conditions; i.e. in the desktop display and flat shaded condition. This hypothesis was only partially supported in the results. While computer experience was not a significant covariate, participants with high spatial ability did have faster completion times on the desktop for the wayfinding task; however, there was very little difference in performance between shading conditions for these users. Unexpectedly, those with high spatial ability performed worse on the desktop display with flat shading for the navigational search task. There are a number of possible explanations for this. Many participants who were experienced CAD/CATIA users reported that the desktop navigation technique was frustrating because it was similar to what they were familiar with but did not have all of the functionality that they were used to. For example, CATIA allows users to graphically select an object and then use that object as the center of rotation. This study did not support that capability, and several users commented that it was hard to tell where the center of rotation was in the desktop conditions. That functionality would not have impacted their performance in the Hansel and Gretel wayfinding condition because after watching the clip and seeing where the target object was located they could choose a path that limited the amount of rotation required, which would explain the better performance. Interestingly, this did not seem to affect desktop performance for the smooth shading conditions. It is possible that the higher fidelity visuals compensated for the lack of functionality in the software, but further research would need to be done to confirm that hypothesis.

A second hypothesis was that participants who did not have previous experience with or a bias towards flat shaded models, would perform better under smooth shaded (high fidelity) conditions. This hypothesis was based on the fact that high fidelity visuals have been shown to facilitate the acquisition of route level knowledge [Wallet et al. 2011]. This hypothesis was also only partially supported in the results. In the desktop navigation condition, people with lower spatial ability performed better than those with high spatial ability. This was particularly true of men. Of note, while the interaction effect between shading and spatial ability was present for both the Hansel and Gretel wayfinding and the Where's Waldo navigational search tasks, the correlation between spatial ability and completion time for the flat shaded condition was inverted between the two tasks for both displays. In the Hansel and Gretel wayfinding task, people with high spatial ability performed better under all conditions; however for the Where's Waldo navigational search task the flat shading seems to have been a disadvantage for people with high spatial ability. This is an unexpected and counterintuitive finding; further research needs to be done to determine the persistence of this correlation. Again this could have been due to the discrepancies between the navigation modes in CAD/CATIA and IC:IDO. For example, participants with high spatial ability most often had engineering roles or other occupations in which they would be using 3D software regularly. IC:IDO has similar but not identical navigation methods to other standard engineering applications, and these small differences could have caused frustration or confusion resulting in the performance disadvantage.

The immersive display had faster completion times only for the navigational search task; the desktop display produced faster wayfinding performance. Another way to look at this would be to say that the immersive display was better for "searching" for a part, but not as good for travelling along a known path. This is most likely due to the larger screen size, which participants reported to perceive as displaying more of the airplane at any given time (although the simulated FOV was constant), this resulted in a reduction in the amount of unnecessary movement performed and facilitating searching. This finding is in line with Tan et al [2006] who found that larger displays bias participants towards an egocentric viewpoint which results in participants moving shorter distances while navigating, regardless of a constant simulated FOV. The desktop was better when the participant knew exactly where they needed to go, probably due to higher interface proficiency, although interface proficiency was not measured.

Another interesting finding was that, women saw the most navigation performance benefit from the immersive display condition. This finding is in line with a previous study which found that an increased FOV and wide display provided a specific benefit to women [Czerwinski et al. 2002]. Men had a higher average score for the GZ-SO, however, contrary to previous studies, this finding was not significant. This could have been affected by a specialized sample population, the majority of female participants had an engineering background, and this may explain the incongruous results. The discrepancy could also be due to a lack of statistical power due to a smaller female sample size as compared to males (11 vs. 17).

While not measured systematically, considerable differences in travel technique were also observed between participants, particularly for the desktop condition. Many participants use a "fly then look around" method in which they would first position the camera in the direction they wished to travel, and then translate forward to the desired location. Once there, they used the mouse to scan the immediate area for the target. Another technique was to "fly" around graphically selecting parts, which would highlight the part in bright yellow and allowed the user to get a better look at the exact shape while ignoring the surrounding objects. More experienced participants often looked at the target on the clipboard and determined what subsystem of the airplane they believed it belonged to and then only search that subsystem. For example, one user noted that the target below had many "inputs and outputs" so they concluded that it was most likely some kind of hydraulic component and searched only in areas where they saw a lot of electrical wiring and/or cords and hoses.

Unfortunately these methods of navigation were informally observed, and not quantitatively measured, but it would be an interesting topic for future research to measure the frequency of the different methods as well as how they correlate to overall performance on the task at hand. This would be particularly interesting to consider with respect to models of spatial knowledge. Participant's navigation behaviors are most likely symptoms of varying degrees of spatial knowledge acquisition. The realm of environmental spatial knowledge is typically discussed in terms of three levels outlined in Siegel and White's [1975] sequential hierarchical model. Landmark knowledge is a relatively simple and declarative form of environmental knowledge which employs specific strategic focal points. Route knowledge consists of landmark-action pairings and is predominantly sequential, allowing an individual to travel only along known paths. Survey knowledge is referred to as a "cognitive map." This level of enables global knowledge of all landmarks and routes in an environment and is necessary for creating novel paths or shortcuts. While this model has been widely adopted for its descriptive properties of the different levels of knowledge, the sequential nature has been largely discredited [Ishikawa and Montello 2006; Klatzky et al. 1990; Montello 1998]. It is now generally accepted that there are different levels of spatial knowledge that a user may acquire during exploration of a new environment but that these levels are independent and develop concurrently rather than sequentially. Spatial cognition and navigation is also often characterized in terms of timeless allocentric spatial representations vs. dynamic updating of spatial representations [Wang and Spelke 2002; Mou et al. 2004]. While participants were observed taking novel paths and short cuts, and from the literature we know that these behaviors exhibit some degree of survey level knowledge and an allocentric representation, a more structured approach to measuring the different navigation methods used by participants could offer intriguing evidence for theories on the acquisition and use of varying levels of spatial knowledge.

5 Limitations

There were a number of limitations to this study, which may have limited the depth of insights. Based on previous literature, it would have been beneficial to measure familiarity, for example, familiarity with the geometry or structure of an airplane, or with the interface or mode of interaction. The nearest measure that was taken was the self-reported measure of "previous experience", however this measure was quite broad and addressed general computer and gaming experience rather than experience in this particular software. In hindsight it also would have been interesting to give a pretest that measured interface proficiency to see if this was a significant predictor of performance. Previous research has shown that together with spatial ability, interface proficiency can account for approximately 20% of the variance in measures of spatial knowledge in VEs [Waller 1999]. Finally, with regard to individual differences, it would have been interesting to include a measure of self-reported spatial ability which has been shown to correlate with tasks measuring knowledge acquired by navigation in environmental spaces as mentioned in the introduction [Hegarty et al. 2002].

Another potential confound was the difference in resolution between the display conditions. Studies comparing VE performance across display types have shown that screen resolution can be a significant factor, specifically low resolution displays result in poorer performance, regardless of display size [Loomis and Knapp 2003; Riecke et al. 2009]. While the immersive screen did have lower resolution in this study, the higher resolution desktop display did not result in better performance overall, and in fact the immersive screen had a lower average completion time for the Where's Waldo navigational search task. This study also would have been more comprehensive if a desktop/stereo display condition was included. With the current design it is not possible to distinguish whether significant results can be attributed to the display size, the stereo image or both. Ideally, this study would have included a desktop/stereo and immersive/nonstereo conditions. While it would have been possible to turn off one of the projectors in the immersive display conditions to produce the immersive/nonstereo conditions, creating stereo images on the desktop display was not an option. Swindells et al. [2004] did make this full comparison (with the addition of head tracking) and still found that display condition appeared to have little influence on task performance [Swindells and Po 2004]. This study, however, was a between subjects design and suffered from high variance between participants which would have made it difficult to observe the effect if it was present.

6 Conclusion

The goal of this research was to provide a use case evaluation of an immersive VR system in use at the Boeing Company. While the results are specific to the IC:IDO system at the Boeing company, the implications extend to other industry users and the VR community as a whole.

To begin, however, it is clear that the following considerations should be made within Boeing when evaluating the benefit of using the IC:IDO system over a standard desktop setup. Most importantly, the task for which the system is intended to be used needs to be considered. The tasks presented here were examples of standard tasks which the IC:IDO system is currently used for. Results showed that the use of the immersive IC:IDO system produces equivalent task performance for wayfinding and faster completion times for the navigation or "search" task. This is most likely due to the larger screen size, which participants reported to perceive as displaying more of the airplane at any given time (although the simulated FOV was constant), this resulted in a reduction in the amount of unnecessary movement performed and facilitating searching. Tan et al [2006] observed similar behavior and attributed it to the larger screen biasing participants towards an egocentric viewpoint. The desktop had a slight, although not significant, advantage when the participant knew exactly where they needed to go (i.e., in the Hansel and Gretel Task), probably due to higher interface proficiency, although, again, this was unfortunately not measured formally.

It is also important to consider what benefits the system is meant to achieve and how those benefits should be prioritized. The immersive display provided some benefits for completion time as discussed above, but the differences between the two display conditions were negligible in terms of accuracy. It is also worth noting that although there was no statistically significant performance differences between display types, the overwhelming majority (27 out of 28) participants reported that they preferred the immersive conditions to the desktop conditions. A few comments of note were that it felt more like "they were in the plane" and that it was just "more fun" and "reminded them of playing Nintendo Wii." Considering the immersive display was not a detriment to performance, perhaps this overwhelming preference should be taken into account when comparing the two options.

Lastly, and probably most importantly, the intended users of the system need to be considered. It has been shown in many previous studies that there is wide variation between individuals in the ability to acquire and use spatial information in a virtual environment. For this study spatial ability, prior experience, attitude toward computers, and gender were used as measures of individual differences. In line with previous research, the results showed that spatial ability and gender were significant covariates for both the navigation and wayfinding tasks [Lawton 1994; Montello et al. 1999; Waller 1999]. These results can be used to create guidelines for industry users of virtual reality. The simple interpretations could be that the shading technique used for tasks of this nature should vary with the spatial ability of the intended users. However, the results could also be taken as a training opportunity. Further research into the persistence of the impact of spatial ability is needed.

Independent of spatial ability, the high fidelity shading technique had lower mean completion times across all conditions. This is a notable result, considering the variety and complexity of "high fidelity" rendering techniques, the technique used in this study was actually quite simple as compared to the possibilities that exist. If high fidelity continues to provide performance benefits, there are many more visually realistic rendering techniques that could be explored. For example, lighting, shadows, ray tracing, and hidden line removal. Potential and current industry users of VR will have to consider that these more sophisticated techniques come at additional computational costs. Some research to consider would be perceptually efficient rendering, or relating the importance of display visually parameters to the accuracy of perception of the virtual environment [Rodger and Browne 2000; Yang 2005].

This research is by no means meant as a comprehensive study applying virtual reality in industry settings, or even of this particular industry application. It is, however meant as a step in the direction of use-case studies for industry applications. As evidenced by the results in this study, the effectiveness or benefit derived from the use of virtual reality depends on the task, the fidelity of the visual environment, as well individual characteristics of the user population. This list is not absolute, but reflects what was reasonable to measure within the confines of a single study. There are many potential areas for future research. For example, evaluating the influence of the interaction method or additional rendering techniques including: lighting, shadows, and hidden line removal. Hubona et al [2004] showed that shadows can improve accuracy on object positioning tasks, but that they increase completion times [Hubona et al. 2004]. It would be interesting to see the effect for these navigation and wayfinding tasks.

A matrix of design and participant variables which may affect performance in these kinds of task would ideally be built and populated with references to help guide a potential user to relevant literature; however, at this point the populated matrix would look quite sparse. As more research is performed we will get a better picture of under what conditions and for whom the system is beneficial. While the variables were listed with the IC:IDO system in mind, they are variables that could be applied to most VR applications so the list could be used to guide other researches in completing cumulative and comprehensive use case evaluations for other systems.

References

BRYANT, K.J., 1982. Personality correlates of sense of direction and geographical orientation. *Journal of personality and social psychology*, 43(6), pp.1318–24.

CZERWINSKI, M., TAN, D.S. AND ROBERTSON, G.G., 2002. Women take a wider view. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, New York, USA: ACM Press, pp. 195–202.

DIETRICH, A. AND WALD, I., 2005. Large-scale CAD model visualization on a scalable shared-memory architecture. In *Proceedings of the 10th International Fall Workshop-Vision, Modeling, and Visualization*. pp. 303–310.

GUILFORD, J. AND ZIMMERMAN, W., 1981. The Guilford-Zimmerman Aptitude Survey manual of instruction and interpretation. *CA: Consulting Psychologists Press*.

HEGARTY, M., MONTELLO, D.R., RICHARDSON, A.E., ISHIKAWA, T. AND LOVELACE, K.L., 2006. Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34(2), pp.151–176.

HEGARTY, M., RICHARDSON, A.E., MONTELLO, D.R., LOVELACE, K. AND SUBBIAH, I., 2002. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5), pp.425–447.

HUBONA, G.S., SHIRAH, G.W. AND JENNINGS, D.K., 2004. The Effects of Cast Shadows and Stereopsis on Performing Computer-Generated Spatial Tasks. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 34(4), pp.483–493.

INFIELD, S., 1991. *An investigation into the relationship between navigation skill and spatial abilities*. University of Washington.

ISHIKAWA, T. AND MONTELLO, D.R., 2006. Spatial knowledge acquisition from direct experience in the environment: individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive psychology*, 52(2), pp.93–129.

KASIK, D., TROY, J. AND AMOROSI, S., 2002. Evaluating graphics displays for complex 3d models. *Computer Graphics and Applications, IEEE*, 22(3), pp.56–64.

KLATZKY, R.L., LOOMIS, J.M., GOLLEDGE, R.G., CINCINELLI, J.G., DOHERTY, S. AND PELLEGRINO, J.W., 1990. Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, 22, pp.19–43.

KYRITSIS, M. AND GULLIVER, S.R., 2009. Gilford Zimmerman orientation survey: A validation. In *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on*. IEEE, pp. 1–4.

LAWTON, C., 1994. Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex roles*, 30(11-12), pp.765–779.

LESSELS, S., 2005. *The effects of fidelity on navigation in virtual environments*. The University of Leeds, School of Computing.

LOOMIS, J.M. AND KNAPP, J.M., 2003. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger & M. W. Haas, eds. *Virtual and adaptive environments*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 21–46.

MANIA, K., WOOLDRIDGE, D., COXON, M. AND ROBINSON, A., 2006. The effect of visual and interaction fidelity on spatial cognition in immersive virtual environments. *IEEE transactions on visualization and computer graphics*, 12(3), pp.396–404.

MEIJER, F., GEUDEKE, B.L. AND VAN DEN BROEK, E.L., 2009. Navigating through virtual environments: visual realism improves spatial cognition. *Cyberpsychology and Behavior*, 12(5), pp.517–521.

MONTELLO, D.R., 1998. A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In *Spatial and temporal reasoning in geographic information systems*. New York: Oxford University Press, pp. 143–154.

MONTELLO, D.R., LOVELACE, K.L., GOLLEDGE, R.G. AND SELF, C.M., 1999. Sex Related Differences and Similarities in Geographic and Environmental Spatial Abilities. *Annals of the Association of American Geographers*, 89(3), pp.515–534.

MOU, W., MCNAMARA, T.P., VALIQUETTE, C.M. AND RUMP, B., 2004. Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), pp.142–157.

RIECKE, B.E., BEHBAHANI, P. AND SHAW, C., 2009. Display size does not affect egocentric distance perception of naturalistic stimuli. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*. Chania, Crete, Greece: ACM, pp. 15–18.

RODGER, J. AND BROWNE, R., 2000. Choosing rendering parameters for effective communication of 3d shape. *Computer Graphics and Applications, IEEE*, 20(2), pp.20–28.

SCHULTE-PELKUM, J., RIECKE, B.E., VON DER HEYDE, M. AND BÜLTHOFF, H.H., 2003. Circular vection is facilitated by a consistent photorealistic scene. In *Proceedings of Presence 2003*.

SIEGEL, A. AND WHITE, S.H., 1975. The development of spatial representations of large-scale environments. *Advances in child development and behavior*, 10, pp.9–55.

SLATER, M., KHANNA, P. AND MORTENSEN, J., 2009. Visual realism enhances realistic response in an immersive virtual environment. *Computer Graphics and Applications, IEEE*, 29(3), pp.76–94.

SWINDELLS, C. AND PO, B., 2004. Comparing CAVE, wall, and desktop displays for navigation and wayfinding in complex 3D models. In *Computer Graphics International, 2004. Proceedings*. IEEE, pp. 420–427.

TAN, D.S., GERGLE, D., SCUPELLI, P. AND PAUSCH, R., 2006. Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interaction*, 13(1), pp.71–99.

THORNDYKE, P. AND HAYES-ROTH, B., 1982. Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, 14(4), pp.560–589.

WALLER, D., 1999. *An assessment of individual differences in spatial knowledge of real and virtual environments*. University of Washington.

WALLER, D. AND KNAPP, D., 2001. Spatial representations of virtual mazes: The role of visual fidelity and individual differences. *Human Factors: The Journal of the*, 43(1), pp.147–158.

WALLET, G., SAUZÉON, H., PALA, P.A., LARRUE, F., ZHENG, X. AND N'KAOUA, B., 2011. Virtual/real transfer of spatial knowledge: benefit from visual fidelity provided in a virtual environment and impact of active navigation. *Cyberpsychology, behavior and social networking*, 14(7-8), pp.417–23.

WANG, R.F. AND SPELKE, E.S., 2002. Human spatial representation: Insights from animals. *Trends in Cognitive Sciences*, 6(9), pp.376–382.

WOLBERS, T. AND HEGARTY, M., 2010. What determines our navigational abilities? *Trends in cognitive sciences*, 14(3), pp.138–46.

YANG, X., 2005. Perceptual level of detail for efficient raytracing of complex scenes. *UK Theory and Practice of Computer Graphics*, pp.1–6.