# The Influence of Shading, Display Size and Individual Differences on Navigation Performance in Virtual Reality in an Applied Industry Setting

**by**

**Barbara Lonnie Hastings**

B.Sc. (Cognitive Systems), University of British Columbia, 2009

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

in the
School of Interactive Arts and Technology
Faculty of Communication, Art and Technology



**SIMON FRASER UNIVERSITY**

**Fall 2013**

# Approval

**Name:** **Barbara Lonnie Hastings**

**Degree:** **Master of Science**

**Title of Thesis:** ***The Influence of Shading, Display Size and Individual Differences on Navigation Performance in Virtual Reality in an Applied Industry Setting***

**Examining Committee:**

**Chair:** Brian Fisher, Graduate Chair

**Dr. Bernhard Riecke**
Senior Supervisor
Assistant Professor

**Dr. John Dill**
Committee Member
Professor Emeritus

**David Kasik**
External Examiner
The Boeing Company

**Date Defended/Approved:**

**Partial Copyright License**

## Abstract

Despite the increasing use of virtual reality within industry and academia, there is a lack of applied usability evaluations within the field. This is problematic for individuals desiring design principles or best practices for incorporating VR into their businesses. The research presented here is a use case study of a virtual reality system used at the Boeing Company for a number of visualization tasks. Twenty eight Boeing employees performed a series of 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 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, training, and effective users for the VR system.

**Keywords**: navigation; spatial cognition; virtual reality; individual differences, fidelity

## Dedication

To my parents, Jeff and Lori Hastings, for their constant support

## Acknowledgements

I would like to thank my supervisor Dr. Bernhard E. Riecke for his intellectual input and guidance in writing this thesis, as well as my committee member Dr. John Dill for his mentorship and encouraging me to come to SIAT in the first place. Thanks also to my fellow i-Spacer Salvar Sigurdarson for his support, insights and empathy along the way.

Lastly, I'd like to thank my Boeing colleagues and advisors: Dr. Adam Richardson for his direction and input while developing this study and Dave Kasik who is a constant source of guidance and the reason I became connected with Boeing to begin with and whose own research was the inspiration for this thesis.

# Table of Contents

## List of Figures

## List of Tables

# 1. Introduction

The roots of Virtual Reality (VR) can be traced back to the1860s when artists began creating 360 degree panoramic murals to produce more immersive experiences for their viewers. Since that time VR has undergone a dramatic transformation and expansion. Not only from an artistic method to a technology based experience, but also from a method of entertainment (panoramic murals to the Sensorama, to video gaming) to a functional tool. Currently, the term Virtual Reality is most often used to describe applications which are computer simulated, 3D, and visually immersive (Demiralp, Jackson, Karelitz, Zhang, & Laidlaw, 2006; Kasik, Troy, & Amorosi, 2002; Sousa Santos et al., 2008; Waller, 2000; Wallet et al., 2011).

In academia, Virtual Reality has proven useful for generating insights into human spatial cognition; encompassing all of the abilities involved in the acquisition and utilization of spatial information from our environment; such as, spatial updating (Riecke, Cunningham, & Bülthoff, 2007; Sigurdarson, Milne, Feuereissen, & Riecke, 2012), navigation and wayfinding (Kasik et al., 2002; Richardson, Montello, & Hegarty, 1999; Satalich, 1995), self-motion illusion (vection) (Schulte-Pelkum, Riecke, von der Heyde, & Bülthoff, 2003), and distance estimation (Loomis & Knapp, 2003; Riecke, Behbahani, & Shaw, 2009; Steinicke et al., 2009; Thompson, Gooch, Creem-regehr, Loomis, & Beall, 2004). There is also significant effort in the research community to understand how aspects of our physical environment contribute to spatial cognition. Spatial cognition studies using real world environments as opposed to VEs, while numerous, are often challenging due to lack of control over and exposure time to the environment. Using VR is a practical alternative for studies of this nature because it provides an ideal combination of a life-like experience while still offering full experimental control.

While academia has adopted VR as an approach to examining human behavior, industry has also recognized the potential benefits of incorporating VR into their processes.

Indeed a major application and catalyst for further development of virtual reality has been the potential use of such systems for industry training in situations too dangerous or impractical (cost-wise) to carry out in real life; such as, flight and vehicle simulators or military training (Johnson & Stewart, 1999). The first head mounted displays (HMDs), engineered by Philco Corporations and Bell Laboratories, were actually designed with helicopter pilots in mind (Carlson, 2003).

Virtual Reality has also had increasing popularity in the medical community as a therapeutic tool. Virtual Reality Exposure Therapy is used for treatment of eating disorders, addictions (García-Rodríguez, Pericot-Valverde, Gutiérrez-Maldonado, Ferrer-García, & Secades-Villa, 2012; Kuntze et al., 2001), pain reduction (Li, Montaño, Chen, & Gold, 2011), emotional well-being in oncology patients (Espinoza et al., 2012), and post traumatic stress disorder (PTSD) (Difede & Hoffman, 2002; Miyahira et al., 2012; Rizzo & Graap, 2008), to name just a few.

It is clear from the exponential growth of virtual reality that it offers a promising medium for a great variety of 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?

You might consider a literature review of relevant VR research to help make these decisions, however what you will find is that the record is largely equivocal on these topics. 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 performance across multiple display types found no significant effect (Kasik et al., 2002; Riecke et al., 2009; Swindells & 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".

Participants, who were all Boeing Commercial Airplane Engineers, performed navigation and wayfinding tasks across three different displays (20 inch monitor, vision station, and 40 inch 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, discussed in detail in Section 2.2.1.1, support the hypothesis that other influences, aside from the display size, may actually have a greater influence on user performance for 3D visualization tasks. In contrast, a series of experiments by Tan (2006) found that physically large displays significantly increased performance on navigation, mental map formation, and spatial memory tasks (Tan, Gergle, Scupelli, & Pausch, 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, Tan, & Robertson, 2002). Czerwinski et al. (2002) suggest that the large display, wide field of view conditions allow for more VE exploration via head/eye movements, freeing up cognitive resources that might otherwise be engaged in building a cognitive map of the environment. This provides a particular benefit for women who have a propensity for navigating by landmark knowledge.

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 ambivalence similar to that of display 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, Geudeke, & Van Den Broek, 2009; Schulte-Pelkum et al., 2003; Slater, Khanna, & Mortensen, 2009; Wallet et al., 2011). Lower fidelity conditions, on the other hand, have been shown to be better for distance estimation, and "remembering" (Mania, Wooldridge, Coxon, & Robinson, 2006; Waller & Knapp, 2001).

Further complicating the issue is a lack of research studying environmental variables as well as individual differences and their relative impacts on performance in spatial tasks. Those that have, have found that the impact of environment and design variables is largely task dependent, and that individual differences are a major source of variation in both real world and computer related spatial tasks (Bryant, 1982; Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006; Wolbers & Hegarty, 2010). Differences between participants in some cases were large enough to make finding other significant effects difficult (Waller, 2000). Even given these findings that show the high degree of variance that can be attributed to individual differences, the vast majority of VR research continues to address physical aspects of the VR interface, or individual characteristics with little to no research measuring both sources of influence or the interaction between the two.

There is little consistency in the field in terms of VEs, interfaces used, tasks performed, and populations tested. Additionally, there is little work reported concerning usability evaluation for intended users and validation of VR systems for specific tasks as compared to the traditional desktop setups. It would be incredible difficult if not impossible for an individual looking to implement a VR to look at the body of work in the VR community and decide on a clear, concise set of design principles for their system.

It is this void that the research presented here is intended 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 & Po, 2004):

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

The specific goal of this study is to determine how individual differences, visual realism and display size 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.

I will first discuss the relevant background literature on each of the variables of interest and how they relate to navigation. Then I will give an overview of other applied industry research before describing in detail the current study and presenting specific results and implications for The Boeing Company as well as the VR community.

# 2. Background and Motivation

## 2.1. Spatial Cognition

In academia, one of the primary uses for virtual reality is to study human spatial cognition by performing spatial tasks such as object rotations, navigation, spatial recall and recognition, distance estimation, map drawing and pointing. Spatial knowledge is a term used to indicate the mental structures or processes that represent the spatial characteristics of a stimulus (Montello, 1993; Waller, 1999a). The term spatial cognition is used to encompass the acquisition, organization and use of that knowledge within an environment (Weatherford, 1982). Just as visual stimuli vary in terms of the kind of spatial information that they afford, spatial tasks also vary in terms of the level of spatial knowledge being tested. This section will begin with an overview of the general principles of spatial knowledge, and then talk more specifically about navigation and wayfinding as those are the tasks used in the present research.

### *2.1.1. General Principals of Spatial Knowledge*

Generally speaking there are two main components of spatial knowledge: small scale knowledge and large (i.e. environmental) scale knowledge (Hegarty et al., 2006). Small scale refers to space that is smaller than the body and can be perceived from a single viewpoint without the need for movement. Small scale abilities encompass the mental operations used to process and manipulate simple objects that can be viewed from one perspective, such as figures, pictures, and maps. Large scale environmental knowledge on the other hand is acquired via movement through a navigable environment that cannot be viewed in its entirety form a single vantage point[1].

#### 2.1.1.1. Small Scale Spatial Knowledge

Tests of small scale spatial ability are numerous and used often; usually involving mental rotations of small objects. A sample of tests of small scale ability from Montello et al. (1999) is shown in Figure 1.

[1] Large Scale space is sometimes subdivided to distinguish between spaces of different scales (spaces that can be mostly viewed from a single vantage vs. spaces that take considerable locomotion vs. spaces that require a map (Montello, 1993)) however, these distinctions are not important for the present research as all virtual environments the same size and the knowledge acquired will be simply referred to as “Large Scale or Environmental”.

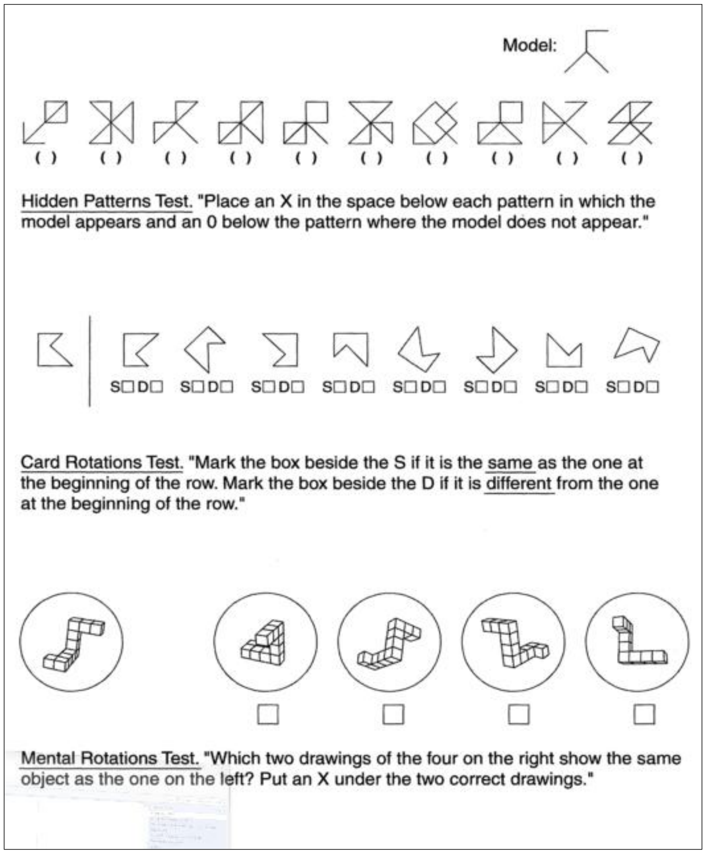
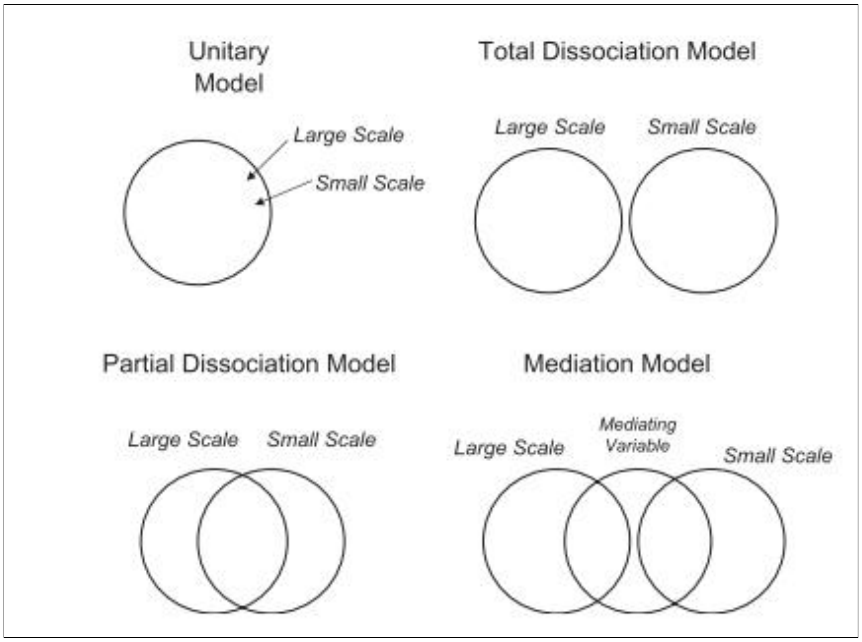



***Figure 1 Example of three popular psychometric tests of small scale spatial ability from Montello (1999): Hidden Patterns & Card Rotations (Ekstrom, French, Harman, & Dermen, 1976), and Mental Rotations (Vandenberg & Kuse, 1978)***

It is not clear exactly how these small scale spatial abilities map to larger scale “environmental” abilities, but there are a number of different models for how they may relate.  Hegarty et.al (2006) outlined and described these potential models with the diagram shown in Figure 2. The Unitary Model describes the two levels of spatial abilities as completely overlapping and depending on the same cognitive processes.

In contrast, the Total Dissociation model proposes that the processes and abilities involved in the respective scales of spatial cognition are wholly separated from each other. The Partial Dissociation model posits some degree of overlap, but with some unique processes involved. Lastly, the Mediation model poses that there are separate processes that are related through a mediating variable. Hegarty et al (2006) also




provide a thorough background on research that may support or disprove each of these

***Figure 2 Hegarty's depiction the different models of interaction between large scale and small scale spatial ability (Hegarty et al., 2006)***

models. She notes that the fact that almost all psychometric tests of spatial ability involve depictions and manipulation of small objects reveals an implicit assumption among researchers that all spatial cognition can be studied with small scale stimuli, as well as an implicit adoption of the Unitary Model.

However there is considerable evidence provided by neuroscience literature that this may not be a well-founded assumption (Hegarty et al., 2006). Several studies have shown that spatial abilities are primarily right hemisphere controlled, however at different scales of space, difference structures and mechanisms are involved (Kimura, 1992; Springer & Deutsch, 1985). Small scale spatial tasks are associated with activation of the parietal lobes (Kosslyn & Thompson, 2003) whereas large scale spatial tasks are associated with activation in the hippocampus (Wolbers & Hegarty, 2010). Further supporting the dissociation theory is evidence from brain lesion patients who, depending on the location of damage, often experience impairment in large scale navigation tasks while experiencing no change in small scale abilities and vice versa (Philbeck, Klatzky, Behrman, Loomis, & Goodridge, 2001). These patterns seem to offer strong support in favor of a total dissociation relationship model. The model of interaction between these two levels of spatial knowledge matters greatly to VR researchers who hope to spatial ability as a controlling measure. As mentioned previously, the prevailing method for evaluating spatial ability pre-study is to use one of many psychometric tests of spatial ability, which are a reflection of small scale ability only and have largely been found to not predict large scale spatial ability (Hegarty et al., 2006).

#### 2.1.1.2. Large Scale/Environmental Spatial Knowledge

The realm of environmental spatial knowledge is typically discussed in terms of three levels outlined in Siegel and White's (1975) sequential hierarchical model.

1. Landmark knowledge is a relatively simple and declarative form of environmental knowledge. Landmarks themselves are specific strategic focal points in an environment from which people travel and orient.
2. Route knowledge consists of landmark-action pairings. In other words, it is a mental representation of how to get to and from known landmarks. Route knowledge is predominantly sequential, allowing an individual to travel only along known paths.
3. Configurational or Survey knowledge is sometimes conceived of as a "cognitive map." This level of spatial knowledge enables global simultaneous knowledge of all landmarks and routes in an environment. Survey knowledge is necessary for creating novel paths or shortcuts or for navigating effectively in a changing environment.

While this model has been widely adopted for its descriptive properties of the different levels of knowledge, the sequential nature which Siegel and White emphasize has been largely discredited (Ishikawa & Montello, 2006; Klatzky et al., 1990; Montello, 1998). Research has shown that Route Level knowledge is not a requirement for acquiring Survey Level knowledge (Thorndyke & Hayes-Roth, 1982). Additionally, while Siegel and White suggested that Survey knowledge developed gradually over an extended period of time, subsequent research has shown that it can be formed quickly (Richardson et al., 1999). 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.

An acquired spatial representation can depend on many variables, including the structure of the environment, the strategy of exploration, the task or goal of exploration. This is true of both real and virtual worlds, but is particularly important to keep in mind when deciding on a spatial task in a virtual environment since the task and environment itself will affect the spatial knowledge acquisition that is subsequently being tested. For example, while designing the present study which consists of two navigation tasks that test route and survey level knowledge, the exposure time, level of detail in the environment, as well as the verbal instructions were all important factors to consider.

The following section will discuss different types and strategies of exploration and the kind of spatial knowledge that they afford.

### *2.1.2. Navigation and Wayfinding*

Navigation and wayfinding are arguably the most common spatial tasks that people perform within virtual reality research. Both terms refer to purposefully moving though a large-scale environment while maintaining orientation (Montello & Sas, 2006; Waller, 1999). Typically navigation tasks are performed after a “learning period” in which participants are given time to learn or observe the environment before being tested. Performance in navigation tasks is most often used as a measure of spatial knowledge acquisition, although it has also been used as a measure of interface proficiency (Satalich, 1995; Waller, 1999). Spatial cognition researchers typically subdivide navigation into two components: a motor component, often referred to as travel or simply navigation, and a cognitive component, referred to as wayfinding (Satalich, 1995). Another important distinction is that wayfinding requires that there be a goal destination or target of navigation (Montello & Sas, 2006).

#### 2.1.2.1. Passive vs. Active Navigation

Previous literature also draws a clear distinction between active and passive navigation. This distinction is important because they are thought to afford different levels of spatial knowledge. Active navigation, for example, is thought to play an essential role in the acquisition of survey knowledge. Passive navigation usually involves the presentation of visual information about a path or environment to a stationary observer (Chrastil & Warren, 2012). Active navigation can encompass any exploration that involves physical activity ***or cognitive activity***. Observing navigation passively obviously also involves cognitive activity, but in order for navigation to be considered “active,” the cognitive component necessarily involves allocation of attention, decision making, or mental manipulation. Chrastil & Warren identify five components of active navigation:

1. Efferent motor commands to determine the path of locomotion
2. Proprioceptive and vestibular information for self motion
3. Allocation of attention to navigation related features of the environment
4. Cognitive decisions about direction of travel
5. Mental manipulation of spatial information.

These components also include updating knowledge of one's spatial position, i.e. spatial updating, and updating or forming a mental representation of one's current position within an environment, or path integration.

The distinctions between these two forms of navigation and what level of spatial knowledge they afford are important for designing and evaluating VR systems in the context of our intended tasks. However, in practice the distinctions are not clear. For example, several studies have shown that survey level knowledge can be formed from passively studying maps, which doesn't involve ideothetic information (Goldin & Thorndyke, 1982; Presson & Montello, 1994; Satalich, 1995). These results lead to a further distinction between "primary" and "secondary" survey level knowledge, the latter being built through map or picture study alone. This discrepancy is also not quite clear, Satalich (1995) compared a group who had direct virtual navigation experience of an environment to those who had just been given a map, and found that by all measures map users performed either equivalently or better than those who had active navigation experience. Conversely, Thorndyke and Hayes-Roth (1984) performed a similar experiment and found the opposite; that people with direct navigation experience within a building were more accurate on subsequent pointing, route and distance estimation tasks than those who only studied a map of the same building. Furthermore, the participants with the longest exposure time to the building prior to experimentation performed the best. The authors concluded that primary survey knowledge (from direct navigation) was superior to secondary survey knowledge (from passive navigation).

While the combination of results from these studies may seem indecisive in terms of what kind of navigation is necessary or sufficient to acquire different "levels" of spatial knowledge a few conclusions can be drawn. For example:

- The sufficiency of passive navigation for acquiring route and survey level knowledge is dependent on the task at hand.
- Familiarity or exposure time to the environment prior to the experimental task has a significant effect on primary survey knowledge but probably not on secondary.

These methods of navigation and their potential impact on the spatial information acquired are important considerations when evaluating a virtual reality system for effectiveness and the subsequent spatial knowledge that it supports. It is crucial for potential users and practitioners of VR to evaluate systems in terms of intended tasks, and experience level of the intended users as these variables determine the quality of spatial information necessary to make the system effective. The current study is an example of this context specific approach, evaluating an in-use Boeing VR system in terms of the intended users (Boeing employees across multiple fields of expertise) and two spatial tasks which represent core tasks the system is used for (**Finding** an Object , **Inspecting** and object, and visually **scanning** scenes) .

## 2.2. Environmental and Individual Factors Influencing Spatial Performance

### *2.2.1. Environmental Factors*

#### 2.2.1.1. Display Type

There have been numerous studies on the impact of display type on spatial performance in virtual environments; however the findings are often equivocal and hard to generalize. Many studies have shown performance benefits for immersive display variables, for example size, wide FOVs, and stereoscopy, while others found no significant effects. This is unsurprising considering the diversity in terms of displays tested (HMDs, CAVE environments, fish tank VR, wall projected, etc.) and tasks measured (navigation, data analysis, memory recall, perceived spatial relations, subjective physical response). This section will discuss some of the relevant findings within the previous literature and further motivate the need for task and user specific evaluations.

Several studies have compared immersive virtual reality displays to standard desktops for information visualization and perception tasks (Arns, Cook, & Cruz-Neira, 1999; Bowman & McMahan, 2007; Schuchardt & Bowman, 2007). Riecke et al (2009) looked at distance estimation performance across 3 different display sizes (HMD, 24", and 50"). The primary goal of this study was to determine if the observed phenomenon of distance compression in HMDs would hold on larger size displays and whether or not this effect could be mediated by using realistic visuals. They found that distance estimations were unexpectedly accurate and showed no signs of distance compression for any of the displays. There was also no observed performance difference across the three display sizes (Riecke et al., 2009). The results of this study suggest that distance compression may be overcome by the use of realistic visual stimuli; however, other research using real world images and live video stream of real world environments did not show such improvement (Thompson et al., 2004; Willemsen, Colton, Creem-Regehr, & Thompson, 2004). Visual fidelity and its impacts on spatial cognition will be discussed in more detail in the next section.

In another study Arns et al. (1999) compared a four-sided CAVE environment to a desktop display for analyzing statistical data in three dimensions. They found the CAVE was better than the desktop for identifying clusters of data and radial sparseness. However a third task which had participants view a dataset containing hundreds of data points and asked them to determine if the data was intrinsically 1, 2, or 3 dimensional was performed equally well using CAVE and desktop displays(Arns et al., 1999). Similarly, Schuchardt et al (2007) compared an immersive 4 sided CAVE-like environment to a less-immersive single screen wall-projected environment and found that the CAVE environment was better for identifying key features of an underground cave structure developed using surveying data (Schuchardt & Bowman, 2007).

Previous literature has also focused on the impact of immersive displays on performance for navigation or "travel" tasks. Gruchalla et al (2004) performed a comparative study to quantify whether or not a large CAVE-like immersive virtual environment provided a performance benefit over a stereoscopic desktop workstation for oil-well path planning. The oil-well planning task involved navigating through a VE to an oil well and then virtually manipulating and improving the well path to avoid interference with other wells.

Out of the sixteen participants, fifteen completed the task faster and more accurately using the larger, more immersive CAVE-like virtual environment (Gruchalla, 2004). Similarly, Ni et al. (2006) evaluated the combined effects of display size and resolution on a navigation task which required participants to virtually navigate to a target room and then perform search and comparison task of textual labels within the target room.  The experiment compared 4 different size-resolution display conditions: low resolution/small size, low resolution/large size, high resolution/small size, and high resolution/large size. An IBM flat panel LCD monitor (approx. 22 inches) was used to produce all "small size" conditions.  For the low resolution/large size condition a rear projected 48" × 27" screen with a resolution of 1280 × 720 pixels was used, and for the high resolution/large size condition a tiled high resolution display module called VisBlocks was used with four blocks arranged in a 2×2 array which created a screen size of 48" by 27" and a resolution of 2560×1440 pixels. Authors found that the high resolution/large size condition resulted in the best task performance, while the low resolution/small size condition was the worst (Ni, Bowman, & Chen, 2006). Furthermore, a series of experiments by Tan (2006) comparing performance on navigation, as well as mental map formation and spatial memory tasks, between standard desktop displays and large projected wall displays also found that physically large, immersive displays significantly improved performance.  Physical field of view was held constant for all conditions and was therefore not a factor. Authors gave users four minutes to explore a virtual environment and learn the location of seven targets.  Users were subsequently placed in a random location in the VE and instructed to return to the location of specified targets. The results showed that users moved shorter distances when viewing the VE on the large display.  This finding persisted even when additional visual cues were added to the desktop conditions.   The authors proposed that these results were due to display size automatically biasing a user into adopting either an egocentric or exocentric strategy. The authors posit that if a user is not provided with an explicit strategy, then large displays cause the adoption of egocentric strategies, in other words behaving as if they are within the virtual environment, whereas small displays cause the adoption of exocentric ones, or behaving as if they have an outside or birds-eye-view of the environment but are not physically within it.  They further supported this theory by showing that physically large displays do not provide a performance benefit for exocentric tasks. Authors used three common mental rotation tests (card test, cube test and Shepard-Metzler test) which required the user to look at two images of objects and determine if they were images of the same object or different objects. Tan et al (2006) found that there was no significant performance difference between the large display and small display for these small scale, exocentric tasks (Tan et al., 2006). Another study looking at the impact of display size and FOV on gender differences in navigation performance found that both wider displays and wider FOVs resulted in lower pointing errors, and faster trial times (Czerwinski et al., 2002). Users were required to find target objects (numbered cubes) within a VE and place them at their corresponding target locations. This task was followed by a brief pointing task in which users were asked to point back to the location of objects once they had been removed from the environment. These tasks were meant to be representative of direct manipulation and navigation within a VE. The task was performed across two display widths (18" and 36") and two levels of FOV (32.5° and 75°). Results showed that, on average, wider displays and wider FOVs resulted in lower pointing errors and faster trial times for both men and women. Interestingly, this study found that while males were faster than females on average, females benefited more than males from wider fields of view (Czerwinski et al., 2002).

Conversely, two studies which looked at the impact of different displays on navigation and wayfinding within 3D models of Boeing aircrafts found no significant performance benefit for large displays on navigation tasks. In the first study Kasik et al. (2002) 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" (p. 3). Participants were thirty-two Boeing Commercial Airplane engineers, and they each performed two tasks. First, an active navigation task called "Where's Waldo," in which they had to search the airplane model for a specific part. Second, a wayfinding task called "Hansel and Gretel", in which the test administrator "flew" the subject on a precomputed path that led to the target object, then returned the subject to the starting point and asked the participant to find their way back to the target part. These two tasks were developed by Boeing subject matter experts to represent standard tasks performed in design review sessions. They were meant to evaluate a person's ability to find an object by relying on visual reference and on visual memory. Kasik et al (2002) presented each task on three different display devices: a 20-inch monitor with a 48º FOV and 1280 × 1024 resolution, an Elumens 1.5m VisionStation with 135° FOV and 1024 × 768 resolution, and a 50-inch Pioneer plasma panel with an 85° FOV and 1365 × 768 resolution. Results showed that participants all performed both tasks best on the 20-inch monitor. In other words, larger display size did not provide a performance benefit for these navigation and wayfinding tasks. A follow up study was performed by researchers at the University of British Columbia in 2004 (Swindells & Po, 2004). This study had the same general goal as the previous Kasik et al. 2002 study: to determine "the importance of the [display size] in the visualization of complex 3D models" (p.1). This study used the same data and tasks, but different display conditions. Performance on the Where's Waldo and Hansel and Gretel tasks were compared across a 19" desktop monitor, a tiled wall display that was 8' high and 10' wide, and an immersive room (i.e. CAVE-style environment) that was 8'high, 10' wide and 10' deep. Resolution for all displays was 1024 × 768 pixels and the displays were configured so that similar viewing angles were maintained. In addition to new display environments, this study also expanded on the previous research by including conditions that had stereo and head tracking (Swindells & Po, 2004). Results from this follow up study showed wide variation between participants in their ability to complete the tasks, and that display type had little influence on overall navigation performance in terms of completion time. Within subjects ANOVAs testing completion time against display type (wall or immersive) and rendering type (stereo/headtracked or mono/non-headtracked) for both tasks yielded no significant main effect or interactions. However, evaluation of a preference questionnaire that was given post-experiment indicated that a majority of participants preferred the immersive displays to the desktop display. Discussion from both studies support the hypothesis that other factors, aside from display type and rendering type (in this case referring to stereo vs. mono), may actually have a greater influence on user performance for 3D visualization tasks involving complex structured geometric models. For example, the high variation between participants in the Swindells et al. (2004) study suggests individual differences may be a greater contributing factor. The authors also suggest that other factors including interaction method, user experience level with 3D navigation, as well as visual fidelity characteristics of the models such as photorealism, lighting or shadows, may influence navigation performance.

In sum, there is little consensus within the field on whether or not display type has a significant influence on performance for spatial tasks. There are many probable reasons for this including confounding variables which may average out the effect of display, high variance between individuals which may overpower the effect of display, and little to no consistency across studies in terms of displays used, visual stimuli presented and tasks performed. Furthermore, even studies having seemingly related tasks often produce contradictory results, for example the navigation/search tasks in Kasik et al. (2002) and the Czerwinski et al (2002).

There is also wide variation in terms of control conditions. Some studies measure immersive displays against other immersive displays (for example CAVE vs HMD), while others measure immersive displays vs. standard desktop monitors, and others still compare immersive displays against the real world (Mcmahan, Bowman, Brady, North, & Polys, 2011). Consequently, it would be very difficult for a potential user of VR to determine whether or not large, immersive displays are worth the monetary and computational expense that they demand.

This provides further motivation for applied, user and task specific, evaluations for their intended use. As mentioned earlier, both the Kasik et al (2002) study and the follow up Swindells et al (2004) study failed to find significant effects of display size or rendering technique (mono vs. stereo) on navigation performance. However, both studies suggest that other experimental and individual factors might have a greater influence. These findings were the primary motivation for the current research. Kasik et al (2002) provide an example of an industry application comparing displays; however there were a number of confounding variables which may have limited the conclusions. For example, all participants in the study were experienced Boeing engineers who used 3D visualization software (for example, CAD/CATIA) as a part of their everyday work. In other words, the participants were highly familiar with the program as well as the stimuli (i.e. Boeing airplane geometry). Since, self-assessed familiarity has been shown to correlate highly with performance on spatial tasks (Waller, 1999a) this level of experience could either provide a performance benefit to participants for 3D tasks in general regardless of display, or it could prove problematic if the display and interaction method was not what they were accustomed to using on a regular basis. While the level of experience was broken down between low, medium and high expertise, these levels were not measured quantitatively and statistics were not reported. Swindells et al (2004) study expanded the experimental design by including larger more immersive displays however it still did not address individual characteristics aside from gender and age which both had no effect. Both studies present an opportunity for follow up research which controls for and measures individual characteristics as well as display type and the other design variables suggested in the study such as interaction method, and photorealism.

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 measured were subsequently used as covariates in analysis to disambiguate the relative effects of environmental factors (display type and rendering method) and individual differences.

#### 2.2.1.2. Fidelity

Environmental fidelity, a concept introduced by Waller, Hunt and Knapp (1998), refers to how accurately a VE resembles its real world counterpart. Although environmental fidelity encompasses many different factors of the system, including vestibular cues and structural integrity of the VE, the present research focused primarily on the visual aspects.

Visual fidelity refers to the degree to which visual features in the VE conform to visual features in the real world. Like environmental fidelity, visual fidelity has multiple components including; geometry, lighting, and material properties. There are generally accepted to be three *varieties* of visual fidelity or visual “realism” in computer graphic; a definition inspired by Margaret Hagen’s book <u>Varieties of Realism</u> (Hagen, 1986) which focused primarily on the geometric aspects of images. Extrapolated to computer graphics, the three varieties are (Ferwerda, 2003):

1. **Physical realism**- the image provides the same visual stimulation as the scene. An accurate point by point representation that has accurate descriptions of the shapes, materials and illumination properties of the scene.
2. **Photorealism**- the image produces the same visual response as the scene; i.e. the rendered image is indistinguishable from a photograph of the scene.
3. **Functional realism**- the image provides the same visual information as the scene. Including shapes, size, positions, motions, and materials. If the rendered image lets you do the task at hand as well as you could in the real world then it is "functionally realistic" for that task.

These varieties provide a set of benchmarks that can be used to evaluate a given rendering technique. Although physical realism has been, and is still, a popular goal in computer graphics, it is impossible to implement in any functional way because this level of realism requires a point by point replication of the spectral energy of the scene that conventional displays are not capable of doing. This limitation also effects the capability to achieve photorealism, however since the standard is for the stimuli to produce the same visual response rather than the same physical stimulation, it is slightly more realizable. The main obstruction for photorealism is the efficiency of rendering algorithms. A scene could take minutes or hours to render at this level of realism, making photorealism unrealistic for any interactive use. For this reason, researchers have been working on more perceptually based algorithms that only render full complexity for parts of the scene which are being attended to (Yang, 2005). However most are still in the development stage and have not been widely applied. Due to these limitations, the most commonly used target for visual fidelity is the third variety, functional realism. Since the criteria for functional realism is basically that the images provide useful information that allows a user to perform a given task as well as they could in the real world, this allows that potentially many rendering styles can be used to produce functionally realistic information.

There hasn't been an overwhelming amount of research done on the impact of visual fidelity on performance in virtual environments. Studies that have looked at the impact of fidelity on spatial cognition have found that, more realistic visuals benefit: reverse route reproduction (Wallet et al., 2011), scene recognition (Meijer et al., 2009), facilitation of circular vection (Schulte-Pelkum et al., 2003), and presence under stressful conditions (Slater, Khanna, & Mortensen, 2009). Lower fidelity conditions on the other hand, have been shown to be better for distance estimation, and "remembering" (Mania et al., 2006). Tasks and the interpretation of "high fidelity/visual realism" differ greatly between studies, for example high fidelity in one sense was used to refer to a virtual environment that was rendered using high quality ray tracing, while in another instance it was used to refer to the addition of landmarks to an otherwise bare “low fidelity” scene. An example from Meijer 2009 is shown below.

***Figure 3 Example of “low fidelity” (left) vs “high fidelity” (right) conditions from Meijer et al. (2009).***

Despite these inconsistencies, in general studies suggest that low fidelity visualizations are more beneficial for tasks that require a cognitive map or "survey level" knowledge of an environment, whereas high fidelity representations are better for construction of egocentric spatial representations or "route level" knowledge (Wallet et al., 2011). It is suggested that this difference can be attributed to low fidelity environments providing only the visual information that is absolutely necessary for gaining high level environmental knowledge. For example, maps provide information about streets but generally not specific buildings, houses, etc. High fidelity environments on the other hand provide greater visual detail and landmarks which would provide useful reference points for an individual trying navigate on a given route through an environment. It is interesting to compare these descriptions to new mobile map applications which people are relying on more frequently for everyday navigation. Google maps for example has begun to incorporate outlines of notable buildings into its “basic” map view to be used as landmarks. Google street view of course offers a high fidelity environment that offers photorealistic visuals with landmarks, street signs etc. It would be interesting to compare these different levels of fidelity offered by Google Maps to see if the additional information actually offers navigational benefits.

Most research on VE navigation tends to use simplistic visuals and focus on other aspects of environmental fidelity. Only a few studies have looked at the impact of visual realism on navigation specifically. Wallet et al. (2011) were focused primarily on the transfer of knowledge between virtual and real environments, and found that learning a route in a high fidelity virtual environment improved subsequent recall in the real world for route retracing, sketch mapping and image sorting. Meijer et al (2009) assessed spatial learning of either a photorealistic or non realistic supermarket presented on a large screen. Participants were given four pen and paper tests after exploring a virtual environment and results showed that the participants who had been exposed to the photorealistic VE, more accurately memorized the environmental layout than those who had seen the non realistic VE. It is important to note that the studies mentioned did not look at the performance impact of visual fidelity on a structured *virtual* navigation or wayfinding task; they instead looked at *real world* navigation performance and environmental recall after virtually learning an environment. Kirschen et al (2000) (Kirschen, Kahana, Sekuler, & Burack, 2000) found that visual motion cues in the form of optic flow promoted faster learning, mainly by preventing disorientation and backtracking in a maze navigation task. Lessels et al (2004) showed that increasing visual fidelity by using images of a real world environment as wall textures, as well as using a wide FOV increased the similarity between virtual and real world navigation performance.

In the past, hardware limitations have rendered it nearly impossible to visualize large-scale CAD datasets in their full complexity. As a result, geometric simplifications and low fidelity rendering techniques were common practice. However, 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 & 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. Furthermore, as mentioned previously, many researchers looked at *real world* performance after a period of virtually learning an environment, making it impossible 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 as to how to display virtual environments on immersive displays within context for their intended use, for virtual navigation applications. To this end we compared two rendering methods, flat and Gouraud, and used the same navigation and wayfinding tasks as Kasik et al. (2002) and Swindells et al. (2004) (discussed in Section 2.2.1.1) to measure virtual navigation performance in the current study.

### *2.2.2. Factors Pertaining to Individual Differences*

#### 2.2.2.1. Spatial Ability

Many spatial cognition researchers have found that small scale paper and pencil tests are generally not predictive of environmental learning; however there have been a few exceptions. For example, Cutmore (2000) looked at cognitive and gender differences that influence navigation ability, using small scale psychometric tests and EEG frequencies to show neural activation patterns. Since the EEG has been used to show asymmetry of neural activation during cognitive tasks, Cutmore et al. hypothesized that this may appear as greater activation in the right hemisphere during VE navigation. Results showed that small scale ability correlated with ability on large scale tasks. Using subtests from the Wechsler Adult Intelligence Scale (WAIS-R), which have been validated to correlate highly with visual spatial IQ, they split people into “high ability” and “low ability” groups. Through a series of maze navigation and distance estimation tasks, they found that people with lower spatial abilities, were less accurate on the distance estimation, and that individuals with high spatial ability developed superior survey level knowledge. They also showed using EEG measurements that people with low spatial ability exhibited greater power (more cognitive effort) in the right hemisphere during VE navigation (Cutmore, Hine, Maberly, Langford, & Hawgood, 2000). Likewise, in 2000 Waller performed a comprehensive examination of individual differences and their impact on VE spatial knowledge using a number of verbal and spatial psychometric tests and found that spatial ability was predictive of individual’s performance in a series of real and virtual maze tasks. Spatial ability was assessed by a pointing task (called WALKABOUT), as well as the Guilford Zimmerman Spatial Orientation task. The WALKABOUT test proved to be a valid predictor of environmental knowledge. This study found that environmental spatial ability was reliably associated with interface proficiency, ability to transfer knowledge to the real world, and a VE projective convergence factor which included a measure of consistency and angle error. It is plausible, however, that the VEs used were perceived by participants as being “small scale stimuli,” since both studies used desktop VR. Neither study directly measured participant’s perception of the size of VE.

In sum, small scale spatial ability has been shown to significantly mediate performance on some spatial tasks, such as mental rotations. It is less clear how to measure environmental spatial ability, and how exactly it relates to small scale ability. To date, the Guilford Zimmerman Spatial Orientation test (GZ-SO) is the only test that seems to address large scale ability (Guilford & Zimmerman, 1948). Neuroscience research indicates that small and large scale abilities are at least partially dissociated. Some variability in environmental spatial ability can be attributed to familiarity or exposure to the environment, which will be discussed in the next sections, but there is also evidence that navigation ability has genetically predisposed factors. For example, genes regulate certain prefrontal cortex functioning which would influence navigation ability. There is also evidence of genetic predispositions that control hippocampal neurogenesis and have direct effects on navigation abilities due to differences in pattern separation (Wolbers & Hegarty, 2010). Furthermore, Hegarty et al also hypothesize three main sources of individual differences in large scale spatial cognition:

1. The ability to encode spatial information from sensory experience
2. The ability to maintain an internal representation of that information
3. The ability the perform spatial transformation

While not necessarily genetically predisposed, these three sources involve the visual and vestibular senses as well as working memory, which can all vary greatly between individuals.

Regardless of the relationship between small scale and large scale abilities it is clear from previous research that spatial ability in general plays a significant role in performance on large scale tasks like navigation and wayfinding (Waller, 1999a). It is also clear that there is significant variance between individuals in terms of these abilities (Wolbers & Hegarty, 2010). As a result, it is important for users and practitioners of VR to take these individual differences into account to gain effective and efficient use of their systems. For example, if those with low spatial ability scores perform worse on large scale environmental tasks, as suggested by (Waller, 2000), it could provide insight and direction for training opportunities. For the purposes of measuring spatial ability for the current study, participants were given the Guilford-Zimmerman Spatial Orientation at the beginning of the study, and these results were used as covariates in the analysis of navigation and wayfinding performance.

#### 2.2.2.2. Gender

Many studies have shown that gender plays a significant role in the acquisition and encoding of spatial knowledge. Prior studies indicate consistent differences favoring men in mental rotation and spatial perception tasks such as distance estimation and pointing (Moffat, Hampson, & Hatzipantelis, 1998). While differences have been shown in both real and virtual environments, they are more pronounced when tested in VEs and most often appear to be a result of variation in the strategies used to encode environmental information. Several researchers have shown that women tend to use landmark information and strategies which are consistent with route knowledge, whereas men rely more on survey knowledge strategies and on geometric and directional information (Bever, 1992; Lawton, 1994; Sandstrom, Kaufman, & Huettel, 1998). In addition to differences in strategy, it has also been suggested that the observed gender differences might actually be explained by familiarity. Men, for example have been shown to have more experience with navigation, more experience playing video games, and also higher joystick interface proficiency which could in turn provide benefits for spatial tasks (Chai & Jacobs, 2009; Lawton, 1994; Waller, 1999a). On the other hand, Kimura (1992) has argued that these differences might have biological basis in both brain structure and hormone levels. High levels of estrogen, for example, have been linked to decreased spatial ability while low-normal male levels of testosterone provide "maximal spatial ability". There have also been a number of evolutionary explanations proposed. The Fertility and Prenatal Care hypotheses posits that female reproduction is enhanced by reduced mobility by means of lower energy expenditure and fewer opportunities for accidents (Sherry & Hampson, 1997). This theory would also explain the negative correlation between estrogen levels and spatial ability since estrogen levels are high during reproduction. The Male Foraging theory, however, posits that historically there was a division of foraging labor in humans, and therefore men will be better at any spatial skills involved in hunting (Jones & Braithwaite, 2003; Silverman & Eals, 1992).

As experience has been posited as a potential explanation for these observed gender differences it is possible that they may not play a significant role for industry applications as most industry VR users would have extensive experience using 3D technology regardless of gender. On the other hand, if biological theories are more accurate, research showing that even females with extensive experience are still at a disadvantage, could provide opportunities for training or alternative instructions which would bias women towards a more advantageous task strategy. To begin to explore these potential circumstances the current study uses gender as a covariate in analysis of VE performance, with participants from various backgrounds and fields of expertise.

#### 2.2.2.3. Background Experience and Familiarity

As mentioned in the previous section, another evident source of individual differences in navigation performance is familiarity. A benefit of using VR that was mentioned in the introduction, Chapter 1, is the ability to control environmental exposure. This is true mainly for VEs that do not have a real world counterpart, in which case an experimenter can give all participants the same amount of practice before beginning test trials and without having to account for real world environmental knowledge that could be used to aid in VE navigation. However, for most training (and other industry) applications this is not a realistic assumption to make. Most training applications are modeled after real world counterparts, so previous real world environmental experience comes into play as a confounding performance variable, and as mentioned previously is quite hard to measure. Additionally, there are other familiarity factors beyond experience with the specific environment, including interface familiarity/proficiency.

In an in depth examination of individual differences in spatial knowledge in real and virtual environments, Waller (1999) lead participants through a battery of paper and pencil tests and questionnaires, followed by a real world and VE navigation tasks. The tests included measures of large scale and small scale spatial ability, background experience and attitude toward computers, interface proficiency and environment familiarity. There were several experimental tasks including map retracing, active navigation to known locations, pointing and distance estimations. Waller (1999) found that self-assessed environmental familiarity and interface proficiency were both highly correlated with performance on projective convergence measures, map accuracy, and distance estimation.

Another study by Thorndyke and Hayes-Roth (1982) compared a group of participants who worked in a building, and therefore had navigation experience, to a group of people who had seen a map of the building but never been inside. Each of these groups was separated into three subgroups. The people who worked in the building were separated into groups based on the amount of time they had worked there (1-2 months, 6-12 months, 13-24 months), while the map group was subdivided by the amount of time they were given to learn the map (until they could reproduce the map without error, until they could reproduce without error plus 30 additional minutes, and until they could reproduce the map plus an addition 60 minutes). Interestingly, map learning resulted in equivalent performance to the group who had worked in the building for at least 6 months. However, the group who had worked in the building for 13-24 months had the best overall performance on distance estimations and there was no significant difference found between the three map subgroups(Thorndyke & Hayes-Roth, 1982).

Richardson et al (2011) examined the impact of video game experience on virtual and real navigation performance and found that gaming experience was predictive of performance in both desktop and immersive VEs but not in the real world. Those with more gaming experience had faster response times and were more accurate in pointing to unseen targets. These results are most likely partially explained by familiarity/proficiency with the interaction device(Richardson, Powers, & Bousquet, 2011)

#### 2.2.2.4. Individual Differences Summary

There are a number of individual user characteristics and environmental variables that impact the ability to learn and use spatial information in virtual environments. This can result in qualitative differences in mental representations of VEs that are formed. While VE aspects are often examined in research for potential influence, individual differences are often not, and the studies mentioned in this section suggest that individual differences account for more variance in VE spatial performance than do system design variables. This has implications for the use of virtual reality systems within industry settings, particularly for training and decision making applications. If the application requires the ability to quickly form and use survey level knowledge of a VE, then it may not be suitable for people who have lower spatial ability and need more training time and exposure to gain survey level knowledge.

For researchers who are more concerned with VE/VR aspects that effect spatial knowledge and performance, it is clear that individual differences are a significant factor and that they should be controlled for whenever possible, either statistically as covariates, or by using within subjects designs whenever possible. For this reason, the current research is a within subjects design (2 Display (immersive, desktop) × 2 Shading (flat, smooth)) primarily looking at the impact of both design (display size and shading) and individual differences on navigation performance in VEs. Prior experience with computers, attitude towards computers, gender, and spatial ability were all measured as used as covariates during analysis.

## 2.3. Virtual Reality for Design and Manufacturing

As mentioned at the outset of this thesis, the industry applications of virtual reality continue to increase exponentially. The tasks and context for which they are used vary greatly, so for the purposes of this research, this section will focus on research that has been done in the manufacturing and/or assembly domain.

Computer Aided Design (CAD) has become ubiquitous in today's large scale manufacturing industries. More and more frequently, these CAD methods are being employed in conjunction with immersive Virtual Reality (VR) displays for every step of the design and production process. From virtual prototyping, to virtual assembly, to virtual maintenance, nearly all companies already define their products digitally. In lieu of more expensive and time consuming physical mock-ups, using computer based engineering benefits companies by reducing costs, shortening design and evaluation schedules, standardizing analysis methods, and the ability to replicate and/or reuse information

Several major companies have incorporated VR technology into their design or production processes; For example, Ford, Volkswagen and Chrysler all use some form of VR for design validation. Yet, while there are numerous examples of academic papers validating the concept of using virtual reality for assembly, manufacturing, and design, there are few examples of actual use case industry validation, with a few notable exceptions. In a 1999 study, BMW explored the potential use of VR for verification of assembly and maintenance processes (Gomes de Sá & Zachmann, 1999). They performed a survey of prospective users of a VR system for Virtual Prototyping. Participants were asked to perform an installation task and rate the VR system in terms of ease of use and satisfaction for ergonomics, collision feedback, interaction, and navigation. The results of the survey showed that overall users found VE navigation intuitive and were satisfied with it as an alternative to the real world. Interestingly, the survey showed that interaction method preference varied with a user's experience level. Users who had no experience with CAD or IT technologies had no preference, but those with high experience preferred voice input. The survey concluded with optimistic results that VR for Virtual Prototyping has the potential to reduce the number of physical mock ups, and improve overall product quality. It is important to note, however, that this study did not measure actual ***performance*** on the installation task, just satisfaction with the system for that task (Gomes de Sá & Zachmann, 1999).

In another example of VR validation from industry, Motorola used HMD and desktop VR systems to train employees on a pager assembly process (Wittenberg, 1995). They found that employees trained in the Virtual Environment performed better at the task than employees who were trained in the real environment, in terms of number of problems encountered and test scores. Authors did not provide detail about the tests or problems that were encountered and offered no explanation for the improved performance other than to say that more comprehensive evaluations, with larger numbers of participants, were in work to verify results.

Another example of an industry validation from Lockheed Martin, provided an overview of a VR system that was implemented to perform Virtual Maintenance and ergonomics evaluation of the F16 design (Abshire & Barron, 1998). The tool incorporated both human models and CAD geometry and allowed engineers to accurately simulate assembly and maintenance tasks, which in turn provided cost and time savings to the company. As in the BMW study, this study did not look at performance in terms of the time it took to perform tasks, and the authers noted that this area requires further research.

Today, nearly all major manufacturing industries are using some form of virtual reality in their design, verification, production and maintenance processes. Most have at least validated that their VR systems can accurately simulate tasks as they would be performed in the real world. These industry validations answer questions such as; Do these virtual prototyping systems provide accurate physical constraints for installation processes? Do they accurate predict collisions or interference? However, very few have taken the step to determine if use of advanced visualization techniques provide a benefit when answering these questions. Is it necessary or beneficial to invest in high fidelity rendering, immersive displays, stereo and headtracking systems, haptic feedback? Do these sophisticated systems help intended users perform tasks faster or more accurately? These questions are important and need to be addressed with quantitative research. The study presented here is intended to be an example of such quantitative research and a step in the direction of use-case validation for industry applications.

# 3. Methods

## 3.1. Introduction

The intent of this study is to demonstrate the importance of performing use case evaluations for virtual reality applications by showing the impact of user and system characteristics for one particular system. The present research is an extension of two previous studies which considered the effect of display size on performance in navigation tasks where the virtual environment was a Boeing 777 aircraft. The prior studies by Kasik et al (2002) and Swindells et al (2004) used two navigation and wayfinding tasks called *Where's Waldo* and *Hansel and Gretel*, respectively. These tasks were originally developed by Kasik, a computer graphics and visualization expert at Boeing, because they are typical of tasks that are performed in design and review sessions at Boeing, while still being general enough to be used in other evaluations. They require the user to orient themselves in a highly complex, geometric virtual environment, in this case an airplane, and to locate specific target objects while being measured for speed and accuracy. For consistency and comparability of results, the current study used the same two tasks, Boeing 787 geometry and different target objects than the previous studies.

As discussed in Section 2.2.1.1, these studies presented an opportunity for further research controlling for individual differences as well as display size and other design variables suggested by the authors (interaction method, photorealism). For that reason the present study extends the display research by Kasik (2002) & Swindells et al (2004) in the following ways:

1. In addition to a comparison of display size, two rendering methods (flat shading and smooth/Gouraud shading) were also compared.

2. Information on participant's spatial ability, computer experience and attitude towards computers (using the Guilford-Zimmerman Spatial Orientation Test and a computer use questionnaire and the Computer Experience Questionnaire developed by Waller 1998) was collected prior to starting the navigation tasks, and these measures were used as covariates during analysis.

## 3.2. Hypotheses

The fundamental goal of this study was to determine the influence of display type, visual fidelity (in this case shading), and individual differences on navigation performance for two specific tasks. Explicitly, the research questions were:

**Q1: What is the effect of display type (in this case desktop vs immersive display) 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?**

In regards to the first question, my hypothesis is that the results will be in line with both the Kasik 2002 and Swindells 2004 studies which found that display type by itself had no significant effect. While not significant by itself, I did expect to find the interaction between display type and individual characteristics significant. Specifically, I think that engineers and other expert users who have many years of computer experience with programs like CATIA, will perform better under familiar conditions; i.e. the desktop + flat shading condition. In regards to fidelity, since route level knowledge is needed for successful navigation and wayfinding and high fidelity visuals have been shown to facilitate the acquisition of route level knowledge (Satalich, 1995; Wallet et al., 2011), I expect to find that people who do not have previous experience/bias with flat shaded models, will perform better under smooth shaded (high fidelity) conditions.

A secondary purpose of this study was to provide a model for a usability study for an industry application of VR. Studies of this nature, that take into account not only display and environmental parameters but also individual differences in the end user, will help Boeing and other companies who use (or are interested in using) VR begin to understand under what conditions and for what tasks VR is beneficial and how to quantify those benefits.

## 3.3. Equipment

#### 3.3.1.1. 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 large amounts of engineering data to produce datasets that represented *partially complete* cross-sections of the 787 Dreamliner airplane. This software also allows the user to export data in either flat shaded or smooth shaded form. This functionality was used to create the two shading conditions for the experiment.

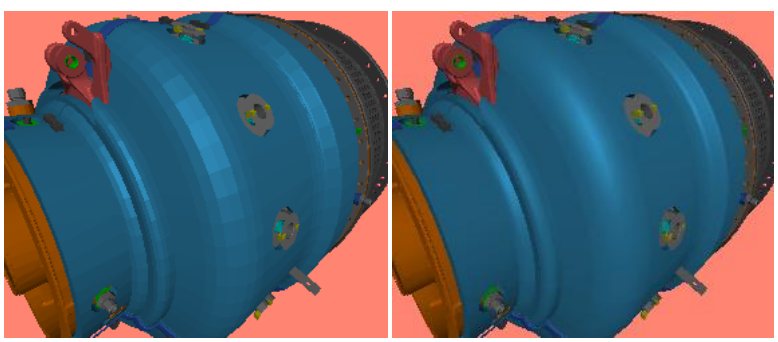


***Figure 4 Single airplane part flat shaded (left) and smooth shaded (right) using IVT***

All 8 test datasets were subsets 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 (ICIDO, 2010) which was used to visualize the airplane sections on either a standard desktop display or on a large immersive display described in detail below.

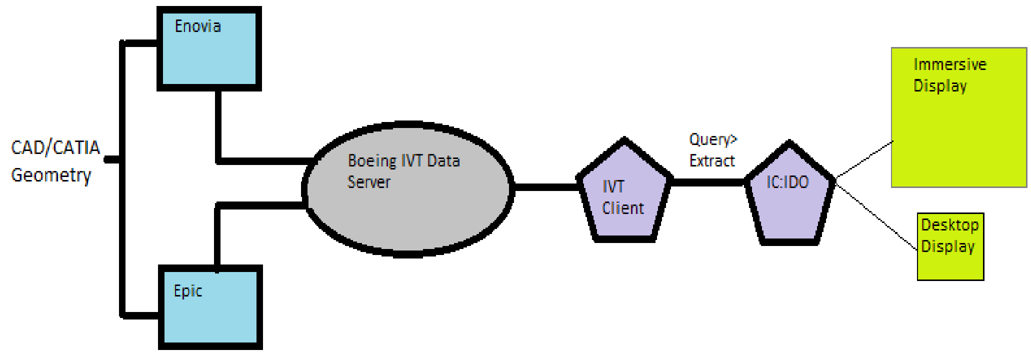



***Figure 5 Data flow from source databases through IVT & IC:IDO to the display devices***

### 3.3.1.2. Hardware

The ICIDO 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 ICIDO immersive display (approx. 57” × 75”) using 2 Christie Digital 1280 ×1024 resolution projectors.

The desktop navigation was performed using a standard mouse and keyboard setup. The immersive navigation was performed using a FlyStick2 from ART, approximately 20 ×10 cm in size, depicted below.

***Figure 6 FlyStick 2 interaction device used for the immersive conditions***

The setup for the two displays is shown in Figure 7. The immersive display is just to the left of the desktop, and both conditions were run off of the same workstation. Users were given the option of either standing or sitting for the desktop condition. Users stood approximately five feet back from the center of the screen for the immersive conditions

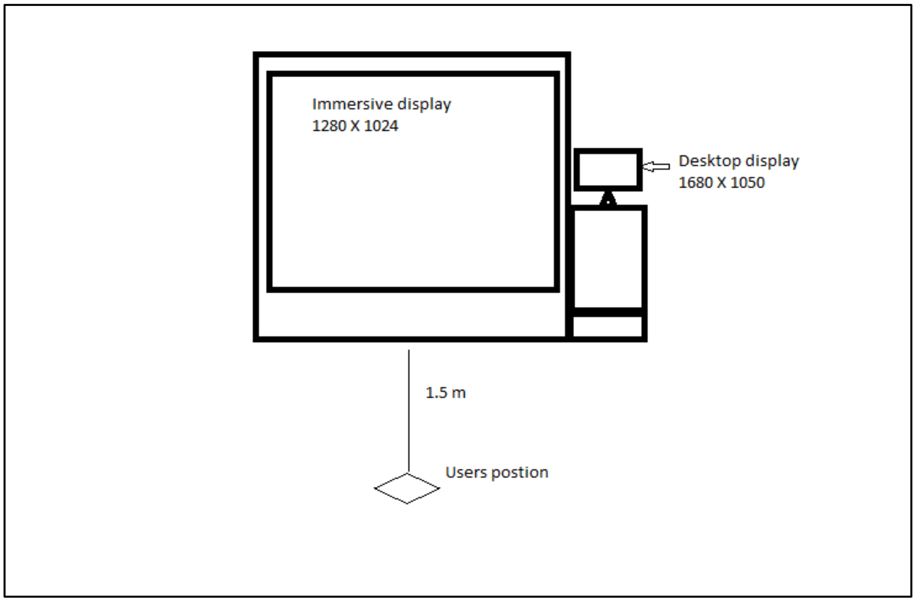



***Figure 7 Experimental Setup showing the location of the displays relative to one another and relative to the user***

### 3.3.2. Participants

The participants were 28 Boeing Employees (17 male, 11 female) working at the Everett Site, recruited via poster and referrals from other participants. Participants intentionally came from a wide range of backgrounds including: engineering, IT, finance, administration, and manufacturing. There was no prior experience required to take part in the study, and $10 Tullys cards were given in exchange for participation.

### 3.3.3. Navigation Mode

For both the immersive and desktop conditions, the IC:IDO system was set to the “Fly” navigation mode. For the desktop condition translation (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 (Figure 8) 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.

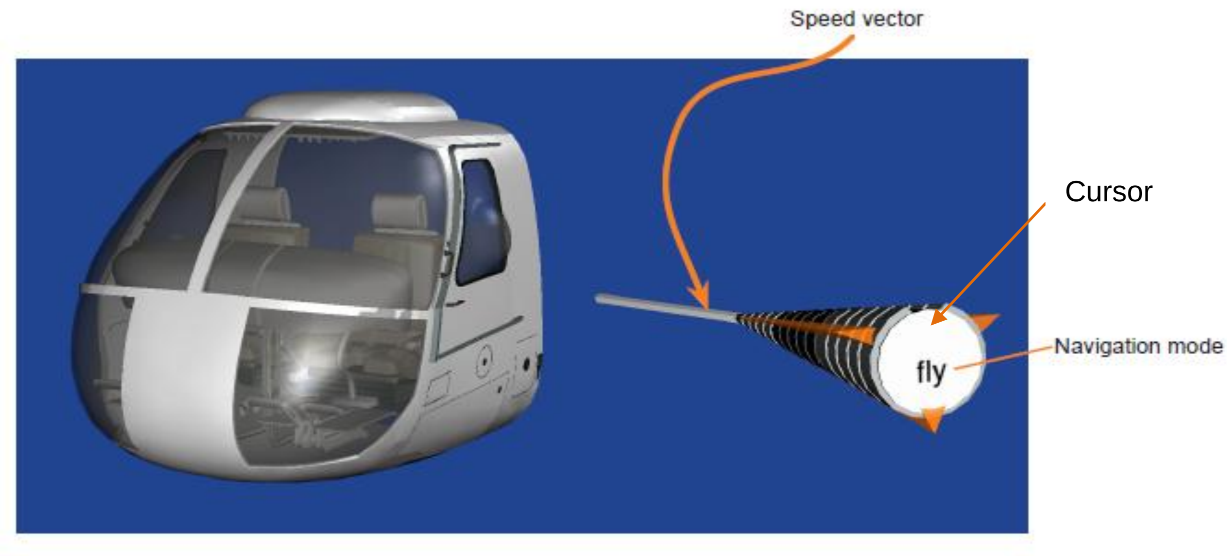



***Figure 8  Depiction of cursor and speed vector in virtual environment in IC:IDO***

## 3.4.  Experimental Design

| | Flat Shading | Smooth Shading |
|---|---|---|
| Desktop Display | 1. Where's Waldo?<br>2. Hansel and Gretel | 1. Where's Waldo?<br>2. Hansel and Gretel |
| Immersive Display | 1. Where's Waldo?<br>2. Hansel and Gretel | 1. Where's Waldo?<br>2. Hansel and Gretel |

***Table 9 The 2 (Display)× 2 (Shading) within subjects experimental design***

The study used a 2 Display (Immersive/Desktop) × 2 Shading (flat/smooth) Design shown in

Table 9.  Half the participants began the experiment on the desktop display and half began on the immersive display to control for carry over effects. Participants completed a total of four tasks per display, two Where's Waldo and two Hansel and Gretel, resulting in a total of 8 tasks per participant, and 224 total for the experiment. The two tasks

(Hansel and Gretel and Where's Waldo) alternated within each display condition, always beginning with Where's Waldo.

## 3.5. Procedure

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 & Zimmerman, 1981; Kyritsis & Gulliver, 2009). The remainder of the study consisted of four conditions.

Display and shading conditions were counterbalanced between participants to control for learning and carry over effects. After completion of the GZ-SO test, before beginning the experimental tasks participants were given verbal instructions on how to use the interaction device for whichever display they were starting on. For example, if they began on the desktop display they were given instructions on how to navigate using the four arrow keys and the mouse. 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 in terms of, for example, the experiment, navigation technique, and display preference.

#### 3.5.1.1. Computer Use Questionnaire

Please rate indicate the degree to which each statement applies to you (1 = Completely disagree; 7 = Completely agree).

1. Computers dehumanize society by treating everyone as a number.
2. I am able to learn about computers very quickly.
3. Computers are beyond the understanding of the typical person.
4. I know how to program computers.
5. I feel at ease when I am around computers.
6. I have played a lot of computer games.
7. I feel comfortable when a conversation turns to computers.
8. Kids these days know more about computers than I do.
9. I have a lot of self-confidence when it comes to computers.
10. I think working with computers would be enjoyable and stimulating.

***Figure 10 Computer Use Questionnaire (Waller, 2000)***

Items on the computer use questionnaire assess both a person's attitude toward computers (items 3, 5, 7, 9, and 10) and prior experience with computers (items 2, 4, 6, and 8). Previous analyses have shown that these scales represent two associated factors of computer use (Waller, 2000). Responses on items 3 and 8 were reflected so that high answers indicated either a more positive attitude toward or more experience with computers. Attitude towards computers and prior experience with computers were then measured as the average score on items from their respective scales.

#### 3.5.1.2. 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 & Zimmerman, 1948). This includes an awareness of whether objects are to the right or left of each other, nearer or farther, etc. Unlike other psychometric tests of spatial ability, the GZ-SO test 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 (land, another smaller sail boat) that give information as to the current position (Figure 11). 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 had an answering system slightly different from the original paper based test, a screen shot of the online version is shown in Figure 11. In the online version, participants answer by clicking on arrow symbols indicating the directions of change. For example if the participant thought the boat had moved forward, and to the right, they would choose (↑, →) with the mouse and then hit “ok”. The developers conducted a study validating that the pattern of Spatial Orientation scores would be the same regardless of the medium (Kyritsis & Gulliver, 2009).

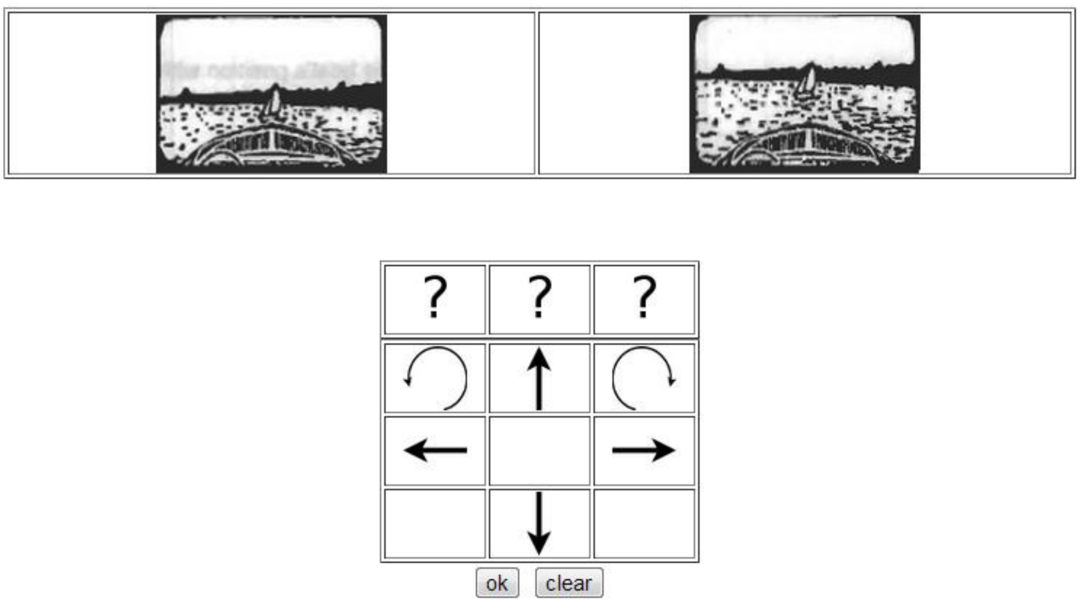



***Figure 11 Screenshot from the online version of the Guilford-Zimmerman test of Spatial Orientation***

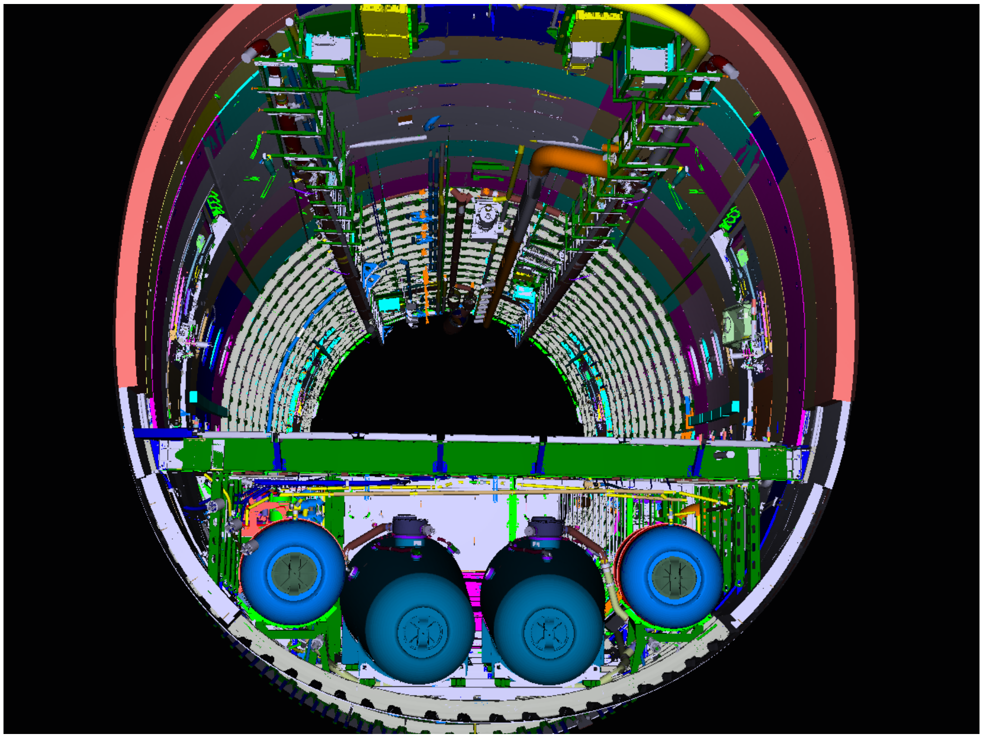


***Figure 12 Sample of a trial starting position***

***Figure 13 Sample of clipboard images showing the target part from three orientations (Front (Top), Left (Middle), Top (Bottom) )in gray scale***

### 3.5.1.3. "Where's Waldo" aka Navigation Task

Participants began each test condition with a navigation 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)[2]. The shading technique for the clipboard images always matched the shading of the test airplane. The participant was then positioned at either the front or back of the airplane, and instructed that they had five minutes to find the target part. An example of the view participants saw from the "starting point" is shown in Figure 12.

The target was positioned in its normal location, orientation, and context within the airplane. 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 participant had identified the correct part or the five minutes had expired.

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

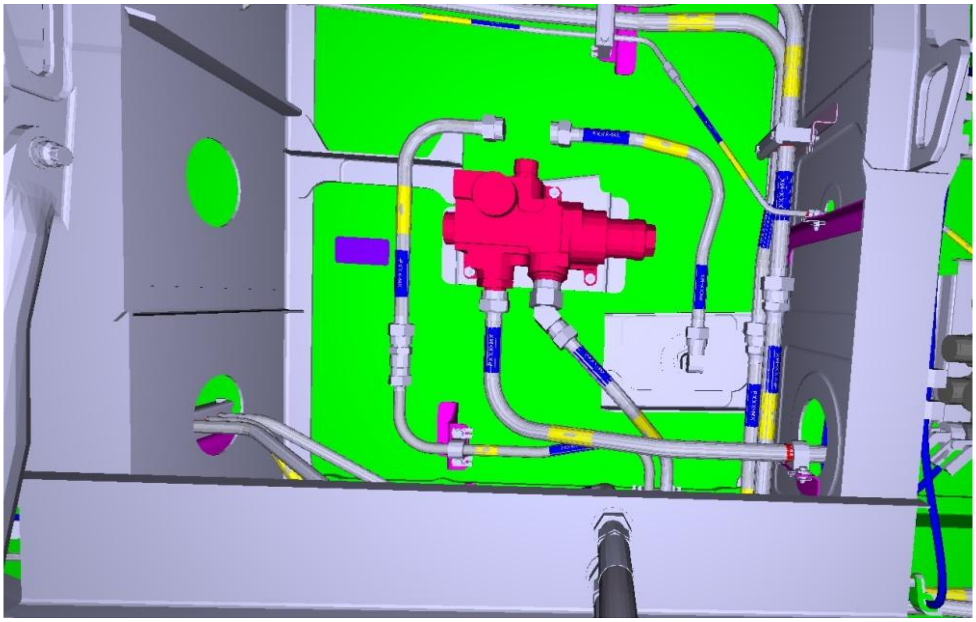


***Figure 14  Example of a target part within its context***

### 3.5.1.4. “Hansel and Gretel” aka Wayfinding Task

The second task was a wayfinding task called “Hansel and Gretel.” The task began with a short clip (60-90 seconds), of the experimenter navigating indirectly to a target part. Participants watched the clip and were told that they would not be required to follow the same path.  At the end of the clip, the participant was then placed back at the start location (front or back of the airplane section) and the task from there was the same as in Where’s Waldo; i.e. to navigate to the target part.  The same “Fly“navigation mode was used for both tasks. The task ended when the participant had identified the correct part or the five minutes had expired.

# 4. 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. Mean completion times for all conditions are shown in Figure 15. Partial eta squared ($\eta_p^2$) is reported in the SPSS output for each ANCOVA and used as a measure of statistical power (Cohen, 1973). It represents the proportion of the total variance that can be attributed to the variable.

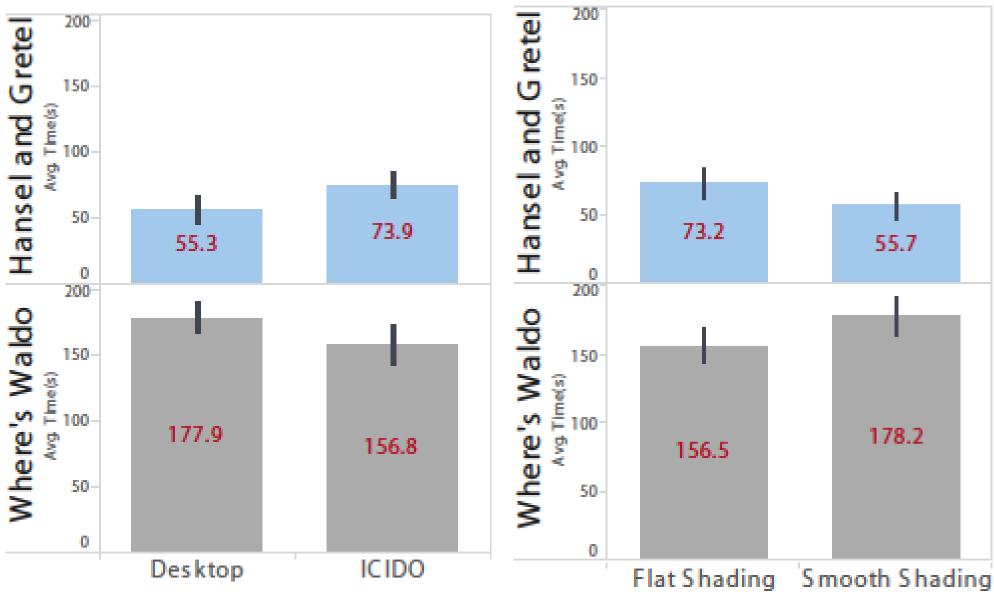



***Figure 15 Mean completion times for display (left) and shading (right) test conditions with error bars showing +- 1standard error of the mean***

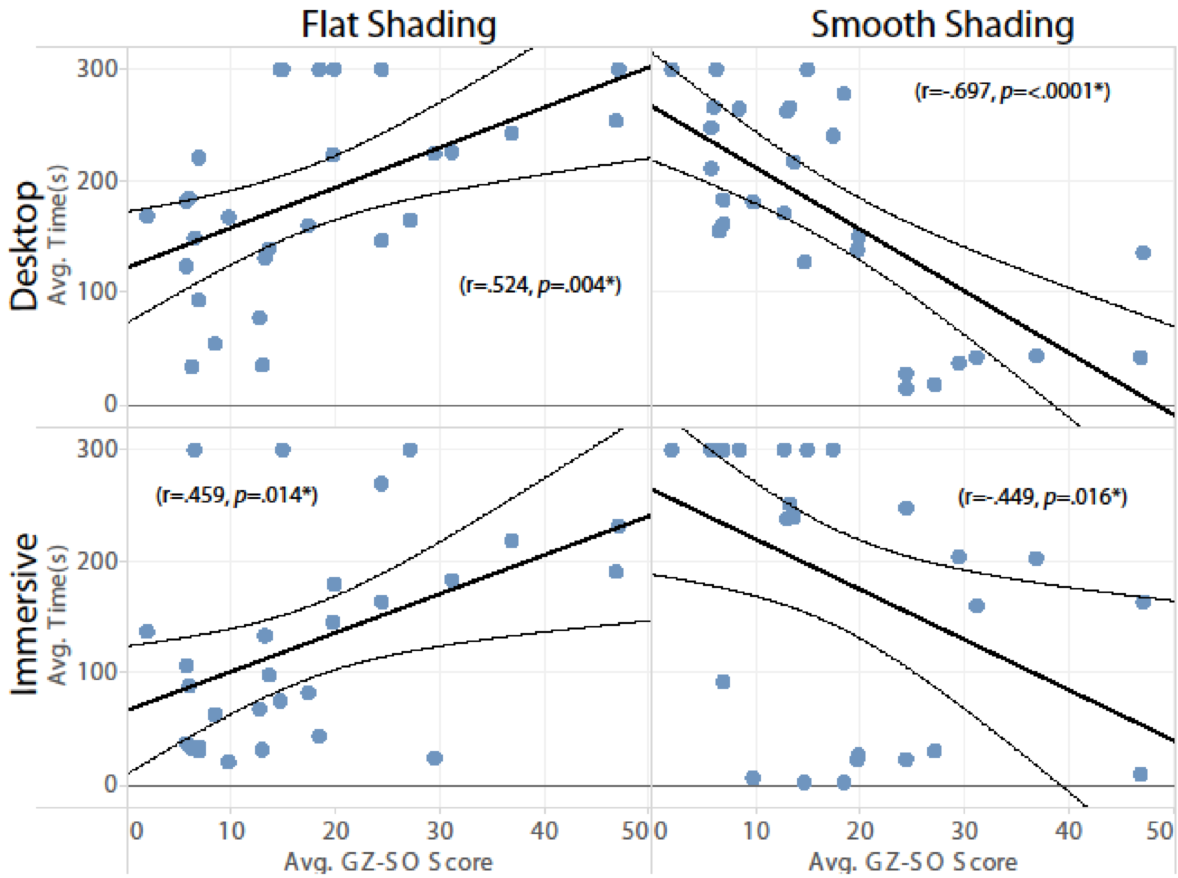



***Figure 16 Correlation between completion time and spatial ability (GZ-SO score) (with confidence intervals) for all four display × shading conditions for the Where's Waldo Navigation task.***

## 4.1. Where's Waldo

The ANCOVA output for the Where's Waldo navigation task is shown in Table 1. The analysis revealed a significant main effect of shading ($F(1,23)=4.15$, $p=.05$, $\eta_p^2=.153$), shown on the right side of Figure 15. There was no main effect of display ($F(1,23)=.889$, $p=.33$, $\eta_p^2=.037$), however results from debriefing discussions showed that, while immersive displays did not result in faster performance, 27 out of the 28 participants preferred the immersive display to the desktop display. Descriptive statistics for the task show that on average participants performed faster in the flat shading condition than the smooth shading condition. Significant interaction effects were also found between

Shading × Spatial Ability ($F(1,23)=11.62$, $p=.002$, $\eta_p^2=.336$), Shading × Gender ($F(1,23)=8.25$, $p=.009$, $\eta_p^2=.264$), and Display × Shading × Spatial Ability $F(1,23)=14.90$, $p=.001$, $\eta_p^2=.393$). Post hoc analysis of each of these interactions revealed that participants with high spatial ability performed better than those with low spatial ability in the smooth shading conditions and that the opposite correlation was true for the flat shading conditions, this trend is shown in Figure 16. Breaking the correlation down into groups based on low, medium, and high GZ-SO scores in Figure 17 shows that the unexpected correlation is due to a strong positive correlation for the "medium ability" group, particularly for the desktop condition. This is still an inexplicable result and more research needs to be done to address the potential causes.

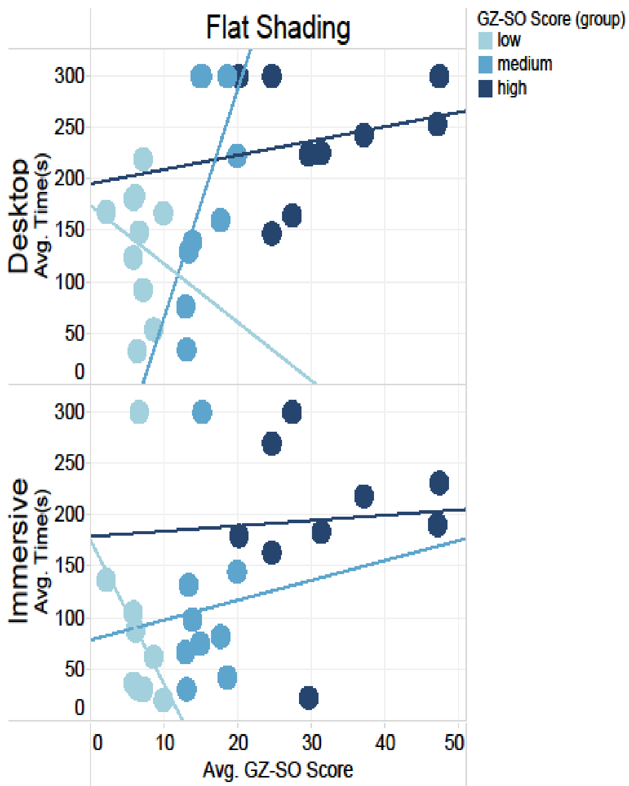



***Figure 17 Inverse correlation between Completion time and Spatial Ability for the Where's Waldo task in the flat shading condition broken down by gz-so level***

This pattern was true for both display conditions. Post hoc analysis of the interaction between shading and gender revealed that women performed significantly better in smooth shading conditions while the opposite pattern was true for men, shown in Figure 18.

***Table 1 Where's Waldo SPSS ANCOVA output***

*Multivariate Tests*

| Effect | Value | F | Error df | Sig. | Partial Eta Squared | Observed Power[b] |
|---|---|---|---|---|---|---|
| Display | .440 | .889[a] | 23.000 | .330 | .037 | .137 |
| Shading | .847 | 4.150[a] | 23.000 | .053 | .153 | .497 |
| **Display * Avg.GZSOScore** | **.737** | **8.214[a]** | **23.000** | **.009** | **.263** | **.784** |
| Display * Avg.attitude | 1.000 | .000[a] | 23.000 | 1.000 | .000 | .050 |
| Display * Avg.experience | 1.000 | .000[a] | 23.000 | 1.000 | .000 | .050 |
| Display * gender | 1.000 | .000[a] | 23.000 | 1.000 | .000 | .050 |
| **Shading * Avg.GZSOScore** | **.664** | **11.616[a]** | **23.000** | **.002** | **.336** | **.904** |
| Shading * Avg.attitude | .951 | 1.192[a] | 23.000 | .286 | .049 | .182 |
| Shading * Avg.experience | .981 | .445[a] | 23.000 | .511 | .019 | .098 |
| **Shading * gender** | **.736** | **8.247[a]** | **23.000** | **.009** | **.264** | **.785** |
| Display * Shading | .882 | 3.067[a] | 23.000 | .093 | .118 | .389 |
| **Display * Shading * Avg.GZSOScorecore** | **.607** | **14.904[a]** | **23.000** | **.001** | **.393** | **.959** |
| Display * Shading * Avg.attitude | .983 | .395[a] | 23.000 | .536 | .017 | .093 |
| Display * Shading * Avg.experience | .994 | .140[a] | 23.000 | .711 | .006 | .065 |
| Display * Shading * gender | .978 | .515[a] | 23.000 | .480 | .022 | .106 |

a. Exact statistic b. Computed using alpha = .05

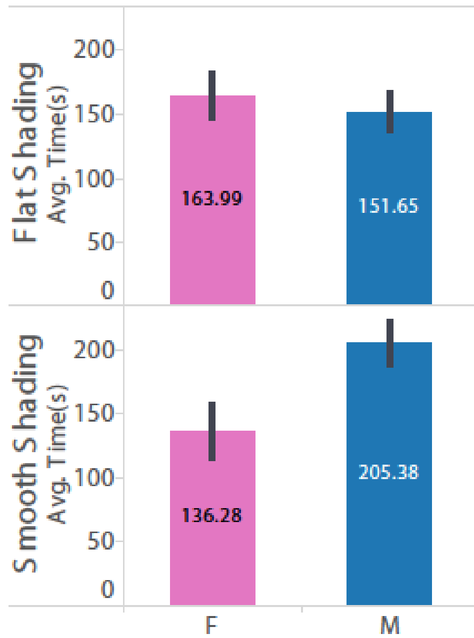



***Figure 18 Significant interaction between Shading × Gender (F(1,23)=8.25, p=.009, $\eta_p^2$=.264) for the Where's Waldo task, with error bars displaying +- 1 standard error of the mean***

## 4.2. Hansel and Gretel

The SPSS output for the 2 × 2 ANCOVA for the Hansel and Gretel task is shown in Table 2. The analysis revealed no main effects of display (F(1,23)=.678, *p*=.419) or shading (F(1,23)=.087, *p*=.771). However, there were three significant interactions: Shading × Spatial Ability (F(1,23)=7.73, p=.011, $\eta_p^2$=.252), Display × Shading × Spatial Ability (F(1,23)=4.74, *p*=.040, $\eta_p^2$=.171), Display × Shading (F(1,23)=4.63, *p*=.042, $\eta_p^2$=.168). Post Hoc analysis revealed that participants performed significantly faster in the flat shading condition on the immersive display, shown in Figure 20, and slower in the flat shading condition for the desktop. That pattern was particularly true for those with high spatial ability. There was no significant correlation between spatial ability and completion time on either display for the smooth shading conditions. Correlations between completion time and spatial ability for all Display × Shading conditions are shown in Figure 19.

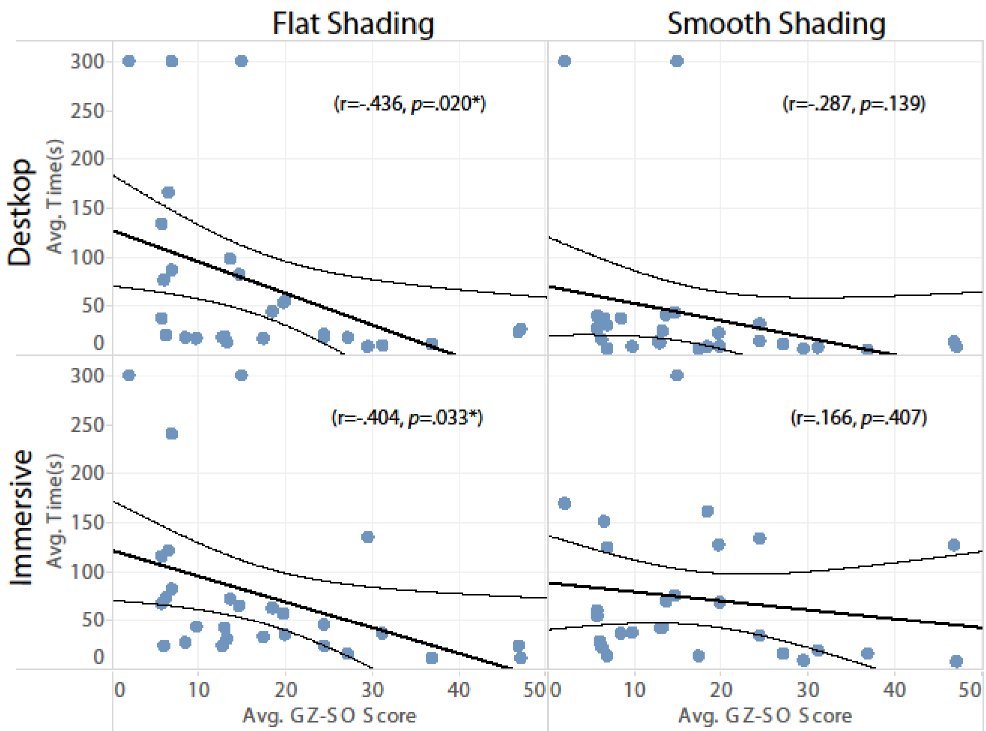



***Figure 19 Correlation between completion time and spatial ability (GZ-SO score) (with confidence intervals) for all four display × shading conditions for the Hansel and Gretel task.***

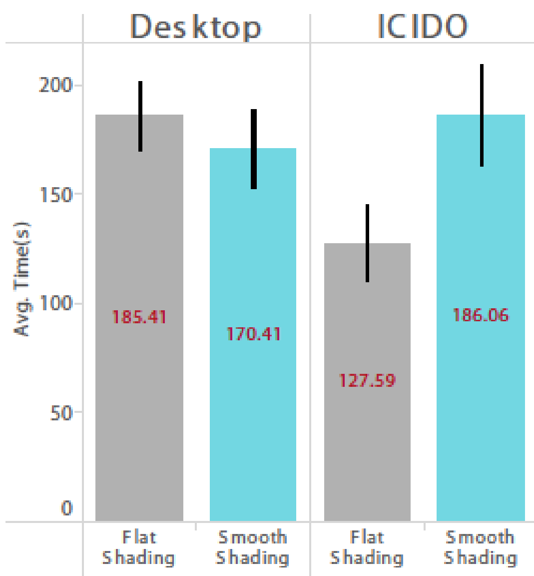



***Figure 20 Significant interaction between Display × Shading (F(1,23)=4.63, p=.042, $\eta_p^2$=.168) for the Hansel and Gretel task, with error bars displaying +- 1 standard error of the mean***

*Table 2 Hansel and Gretel SPSS 2×2 ANCOVA output*

| *Multivariate Tests* | | | | | | |
|---|---|---|---|---|---|---|
| Effect | Value | F | Error df | Sig. | Partial Eta Squared | Observed Power[b] |
| Display | .971 | .678[a] | 23.000 | .419 | .029 | .124 |
| Shading | . 996 | .087[a] | 23.000 | .771 | .004 | .059 |
| Display * Avg.GZSOScore | . 992 | .183[a] | 23.000 | .673 | .008 | .069 |
| Display * Avg.attitude | .971 | .682[a] | 23.000 | .417 | .029 | .124 |
| Display * Avg.experience | .972 | .660[a] | 23.000 | .425 | .028 | .122 |
| Display * gender | . 974 | .613[a] | 23.000 | .442 | .026 | .117 |
| **Shading * Avg.GZSOScore** | **.748** | **7.730[a]** | **23.000** | **.011** | **.252** | **.759** |
| Shading * Avg.attitude | . 995 | .121[a] | 23.000 | .732 | .005 | .063 |
| Shading * Avg.experience | . 970 | .700[a] | 23.000 | .412 | .030 | .126 |
| Shading * gender | . 991 | .212[a] | 23.000 | .650 | .009 | .073 |
| **Display * Shading** | **. 832** | **4.634[a]** | **23.000** | **.042** | **.168** | **.541** |
| **Display * Shading * Avg.GZSOScore** | **. 829** | **4.742[a]** | **23.000** | **.040** | **.171** | **.550** |
| Display * Shading * Avg.attitude | . 903 | 2.474[a] | 23.000 | .129 | .097 | .326 |
| Display * Shading * Avg.experience | . 935 | 1.607[a] | 23.000 | .218 | .065 | .229 |
| Display * Shading * gender | . 860 | 3.746[a] | 23.000 | .065 | .140 | .458 |

a. Exact statistic b. Computed using alpha = .05

## 4.3. GZ-SO and Computer Use Questionnaire

A Pearson's correlation coefficient was computed to assess the relationship between spatial ability (as assessed by the GZ-SO) and completion time.  As expected, there was a negative correlation between GZ-SO score and completion time for both tasks across all display and shading conditions, meaning that people with higher spatial ability completed the tasks faster, however this did not reach significance (r(52)=-.35, *p*=.07). An analysis of the two tasks independently revealed that this correlation was stronger for the Hansel and Gretel wayfinding task (r(52)=-.36, *p*=.056) than for the Where's Waldo navigation task (r(52)=-.20, *p*=.31). When results for the two display conditions were separated and analyzed individually, the correlation for the desktop condition became significant (r(52)=-.33, *p*=.047). Attitude towards computers and experience with computers were positively correlated (r(52)=.48, *p*=.009), meaning that those with more self-reported experience with computers also had a more positive attitude about them. Neither experience nor attitude significantly correlated with completion time (r(52)=-.31, *p*=.11) and (r(52)=-.17, *p*=.38), implying that prior experience and attitude towards computers do not impact performance on these navigation and wayfinding tasks.

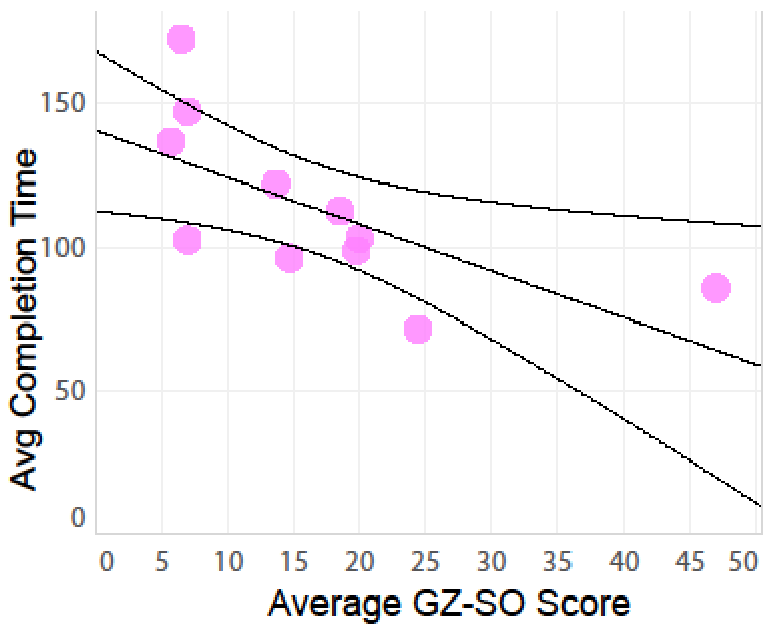



***Figure 21 Negative correlation (r(9)=-.66, p=.025) with confidence bands for women between completion time and spatial ability (GZ-SO Score)***

## 4.4. Gender

While gender did not have a significant main effect, it was a significant covariate with shading in the Where's Waldo navigation task. Men performed better in the flat shaded condition (although not significantly), and women performed significantly better in the smooth shaded condition Figure 22.

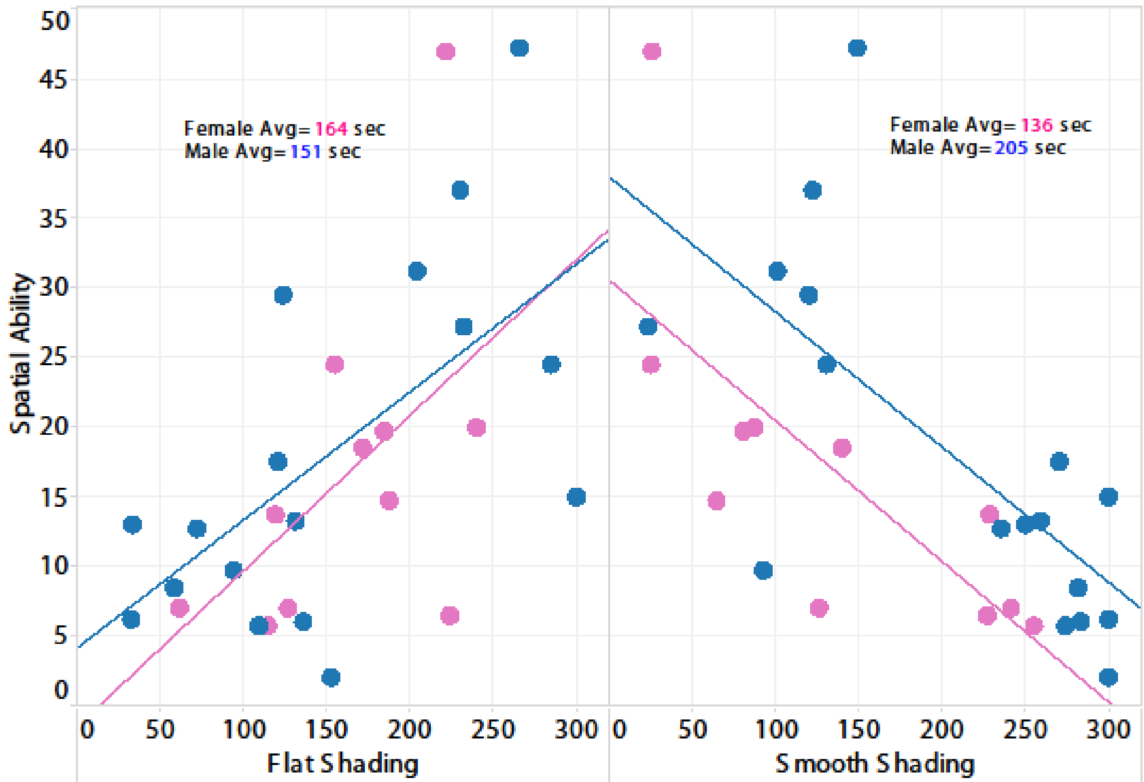



***Figure 22 Significant interaction between spatial ability and completion time across shading conditions for the Where's Waldo task***

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, however this 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), shown in Figure 21, indicating that women with higher spatial orientation scores tended to complete the navigation task faster.

# 5. Discussion

There were two tasks performed in this study. First, an active navigation task called "Where's Waldo," in which participants searched a CAD model of an airplane a specific part and second, a wayfinding task called "Hansel and Gretel", in which participants were passively guided on an indirect route through an airplane to a specific part and then subsequently had to find their way back to that part. These two tasks 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 that I was interested in investigating at the outset of this section. 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. Participants with high spatial ability did have faster completion times on the desktop for the wayfinding task, "Hansel and Gretel." 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 Where's Waldo navigation 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 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.

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 and the Where's Waldo 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 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 ICIDO. For example, participants with high spatial ability most often had engineering roles or other occupations in which they would be using 3D software regularly. ICIDO 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.

Interestingly, even though there was no significant main effect of display type, 27 out of the 28 participants reported that they preferred the immersive display over the desktop. The immersive display had faster completion times only for the navigation 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 Czerwinski et al. (2003) study noted in the background section which found that an increased FOV and wide displays provided a specific benefit to women (Czerwinski et al., 2003). 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 they would travel directly forwards to the desired location. Once there, they used the mouse to look around and search the immediate area for the target. Another technique was to fly around graphically selecting many parts, which highlighted the part bright yellow and allowed the user to get a better look at the exact shape while ignoring the surrounding parts. A technique that was used by participants who were more familiar with the airplane structure was to look at the target on the clipboard and determine 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 (Figure 23). 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.

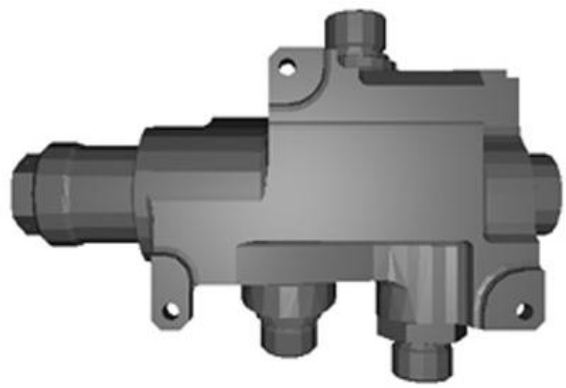


***Figure 23 Example of one of the target parts that had many physical inputs***

## 5.1. 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. In terms of this research, familiarity could have taken on a number of different forms, for example, familiarity with the geometry and structure of an airplane (particularly the 787), or familiarity with the interface or mode of interaction. The most closely related measure that was taken in this study 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 in subsequent tasks. 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, 2000). Participants also should have been tested for stereo blindness prior to starting the experimental tasks, as Swindells et al. (2004) did in their study with the Titmus Stereo Fly Test to ensure that effects could be attributed to display characteristics and rather than physiological (Somers & Hamilton, 1984; Swindells & Po, 2004).

Another limitation and potential confound in this study was the difference in resolution between the two display conditions. As noted in the Procedure section, the Dell Desktop Monitor had a resolution of 1680 × 1050, while the immersive display had a resolution of 1280 × 1024. Previous 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 & Knapp, 2003; Riecke et al., 2009; Satalich, 1995). 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 navigation task. While the impact of the different resolutions cannot be stated conclusively, all previously literature on the impact of resolution on navigation performance suggests that there should have been a negative effect of lower resolution for the immersive screen, and this was not observed. It is possible that the difference in resolution was not large enough to cause a performance difference or that other display factors such as stereo visuals and display size may have averaged out the effect of resolution differences, and this would be an interesting question for future research.

Finally, the study would have been more comprehensive if a desktop/stereo condition was included. Comparing a non-stereo desktop condition to a stereo immersive condition is not ideal because it is not possible to distinguish whether significant results can be attributed to the difference in the display size, the stereo image or both. Ideally, this study would have included eight display conditions instead of four. In addition to the desktop/nonstereo/flat, desktop/nonstereo/smooth, immersive/stereo/flat, and immersive/stereo/smooth conditions I would have liked to have also had desktop/stereo/flat, desktop/stereo/smooth, immersive/nonstereo/flat, and immersive/nonstereo/flat. 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) using the same tasks and stimuli as the present research and still found that display condition appeared to have little influence on task performance (Swindells & 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. Controlling for individual differences or a within subjects study rather than between would have minimized this issue, as the author's themselves noted at the end of the paper.

## 5.2. Implications and Outlook

The stated goal throughout this thesis was to provide a use case evaluation of the IC:IDO Immersive VR system at the Boeing company as a motivational example of an applied case study. 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. As mentioned previously, those tasks include:

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

Both the navigation and the wayfinding task had aspects of **Finding** an Object (1), **Inspecting** and object (2), and visually **scanning** scenes (3). The wayfinding task, Hansel & Gretel, was an example of passive learning followed by active navigation. Despite the learning period in this task being quite short (60-90 seconds), participants were able to not only navigate to the target successfully, but were also observed taking novel paths and shortcuts (unfortunately this was only an informal observation and was not measured directly).

From the literature we know that both of these behaviors are signs of survey level spatial knowledge. This may have been facilitated slightly for those participants who had a working knowledge of the airplane prior to the experiment; however the behavior was not limited to this group. This further supports the notion that the different levels of spatial knowledge develop concurrently rather than sequentially and that survey level knowledge does not necessarily develop over an extended length of time (Richardson et al., 1999).

Results also 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. the Hansel and Gretel Task), probably due to higher interface proficiency, although, again, this was unfortunately not measured.

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 result of display type, 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, Lovelace, Golledge, & Self, 1999; Waller, 1999). Specifically, users with high spatial ability performed better in the smooth shading conditions for the navigation task on both display devices, while the opposite was true for the flat shading conditions. For the wayfinding condition, people with low spatial ability performed worse on both devices and for both shading techniques. 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. The spatial ability tests used in this study, the GZ-SO, as well as the experimental tasks, were meant as measures of large scale spatial ability. As discussed in 2.1.1 it is not clear how these large scale abilities map to the small scale abilities often tested by psychometric tests. Another interesting direction for future research would be to measure both large and small scale abilities prior to experimental tasks to see which model proposed by Hegarty et al (2006) most accurately depicts the relationship.

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 benefit, there are plenty of more visually realistic rendering techniques that should/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 & 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 resulting from this work, some mentioned in earlier sections, including:

1. Evaluating the influence of the interaction method. The IC:IDO system has a number of other "navigation modes" from which users can choose.
2. Evaluating the influence of 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, Shirah, & Jennings, 2004). It would be interesting to see the effect for these navigation and wayfinding tasks.
3. Evaluating the impact of head tracking for navigation and wayfinding. The IC:IDO system includes head tracking in the immersive mode. I chose not to use it in this study for the sake of consistency with the desktop condition, but further research to investigate the benefit of its use would be an interesting extension.
4. Measuring the frequency of users taking novel paths & short cuts and how this effects navigation performance

A matrix of design and participant variables which may affect performance in Where's Waldo and Hansel and Gretel is shown in Figure 24.

| | | | Individual Differences | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Gender | Age | Culture | Experience | | | Stereo Acuity | Spatial Ability | | Spatial Memory | Color Blindness | Search Strategy |
| | | | | | | Environment | Task | 3D | | Small Scale | Large Scale | | | |
| System | Display | Resolution | | | | | | | | | | | | |
| | | Size | | | | | | | | | | | | |
| | | Contrast | | | | | | | | | | | | |
| | | Brightness | | | | | | | | | | | | |
| | | Stereoscopy | | | | | | | | | | | | |
| | | Head Tracking | | | | | | | | | | | | |
| | | Frame rate | | | | | | | | | | | | |
| | | Field of View | | | | | | | | | | | | |
| | | Eye Tracking | | | | | | | | | | | | |
| | | Curvature | | | | | | | | | | | | |
| | Speed | Frame rate | | | | | | | | | | | | |
| | | Screen Refresh Rate | | | | | | | | | | | | |
| | Input Device | Mouse | | | | | | | | | | | | |
| | | Keyboard | | | | | | | | | | | | |
| | | Wand | | | | | | | | | | | | |
| | | Joystick | | | | | | | | | | | | |
| | | Hand | | | | | | | | | | | | |
| | | Trackball | | | | | | | | | | | | |
| Visual Environment | Rendering Style | Hidden Line | | | | | | | | | | | | |
| | | Hidden Surface flat shading | | | | | | | | | | | | |
| | | Hidden Surface smooth | | | | | | | | | | | | |
| | | Ray Tracing | | | | | | | | | | | | |
| | | Radiosity Rendering | | | | | | | | | | | | |
| | | Shadows | | | | | | | | | | | | |
| | | Lighting | | | | | | | | | | | | |

***Figure 24 Matrix of variables with potential to influence performance in Where's Waldo and Hansel and Gretel***

Ideally this matrix would be populated with studies which have looked at the specific design and individual difference variables 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 ICIDO 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.

Although this research looked at a specific system, the lessons and insights can be applied more generally. At the very least, these results show that aspects of the environment and user that affect performance and efficacy are numerous for spatial tasks. And while using virtual reality applications add their own interesting facets to that, they also provide environmental control to an otherwise variable field. Another important take away from this research is further support for the emphasis of individual differences and advocating within subjects design whenever possible.

# 6. Conclusion

There were a number of questions asked at the outset of this thesis.

> *Is using virtual reality a better alternative than a standard desktop approach? Is the answer to that question task dependent? Does it change depending on the type of display used? Or the interaction method? The answers to these kinds of questions are crucial for the development and deployment of Virtual Reality systems. With so many possibilities for advancements and applications, how can researchers or industry users know where to invest their time and resources?*

The study presented was not a comprehensive look at these questions for VR in general, but rather an example of how to begin to address these queries for one industry application. The most important conclusion of this research is that *there are individual and system characteristics that significantly influence the effectiveness of using virtual reality for spatial tasks*. Specifically, user's spatial ability was a predictor of their task performance across all display and shading conditions, although this only reached significance for the desktop condition. Additionally, participants with high spatial ability performed better than those with low ability in the smooth shading conditions and the opposite correlation was true for the flat shading conditions. This pattern was true for both the immersive and desktop displays. Women in particular saw a performance benefit from higher fidelity visuals.

The momentum has swung rapidly from the use of expensive, time consuming physical mock-ups to the realm of digital representations and simulation. While research has shown that it is possible for people to acquire spatial knowledge from Virtual Environments (Lessels, 2005), the rate of learning is slower and information less accurate than that acquired from the real world. In addition, research has also shown that certain environmental characteristics can facilitate this acquisition, but that this is largely task dependent. It is also clear that there is profound ***variation between individuals*** in the ability to acquire and use spatial information in the real world and which is only more pronounced in VEs. These individual differences exist regardless of the task or application. Indeed, the results of this study as well as previous research support the notion that individual differences account for greater variance in VE performance than environmental and physical characteristics of the VR system (Hegarty et al., 2006; Swindells & Po, 2004; Waller, 1999; Waller & Knapp, 2001). Industries that use virtual reality, predominantly manufacturing, have very complex geometric data, as well as highly skilled workforces. There has been plenty of work done in the domain of system validations, in terms of ergonomic constraints and physical parameters that they produce (Chryssolouris, Mavrikios, Fragos, & Karabatsou, 2000; Mavrikios, Karabatsou, Fragos, & Chryssolouris, 2006). However, in order to effectively implement VR within their processes, industries also need to ask and verify if the intended users of the system ***perceive*** the depicted process accurately. Variation due to individual differences and environmental aspects need to become a research and industry priority. This will help create more accessible and appropriate VR systems for applications.

## References


Abshire, K., & Barron, M. (1998). Virtual maintenance real-world applications within virtual environments. In *Reliability and Maintainability Symposium, 1998. Proceedings., Annual* (pp. 132–137). Anaheim, CA: IEEE.

Arns, L., Cook, D., & Cruz-Neira, C. (1999). The benefits of statistical visualization in an immersive environment. In *In Virtual Reality, 1999. Proceedings, IEEE* (pp. 88–95). IEEE.

Bever, T. (1992). Logical and extrinsic sources of modularity. In M. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and Cognition* (25th ed., Vol. 25, pp. 179–212). Hillsdale, NJ: Lawrence Earlbaum and Associates.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer*, *40*(7), 36–43.

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

Carlson, W. (2003). A Critical History of Computer Graphics and Animation. *Ohio State University*. Retrieved from http://accad.osu.edu/~waynec/history/lesson17.html

Chai, X. J., & Jacobs, L. F. (2009). Sex differences in directional cue use in a virtual landscape. *Behavioral neuroscience*, *123*(2), 276–83. doi:10.1037/a0014722

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic bulletin & review*, *19*(1), 1–23. doi:10.3758/s13423-011-0182-x

Chryssolouris, G., Mavrikios, D., Fragos, D., & Karabatsou, V. (2000). A virtual reality-based experimentation environment for the verification of human-related factors in assembly processes. *Robotics and Computer-Integrated Manufacturing*, *16*(4), 267–276.

Cohen, J. (1973). Eta-squared and partial eta-squared in communication science. *Human Communication Research*, *28*(56), 473–490.

Cutmore, T. R. H., Hine, T. J., Maberly, K. J., Langford, N. M., & Hawgood, G. (2000). Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human-Computer Studies*, *53*(2), 223–249. doi:10.1006/ijhc.2000.0389

Czerwinski, M., Smith, G., Regan, T., Meyers, B., Starkweather, G., & Way, O. M. (2003). Toward Characterizing the Productivity Benefits of Very Large Displays, (c), 9–16.

Czerwinski, M., Tan, D. S., & Robertson, G. G. (2002). Women take a wider view. In *Proceedings of the SIGCHI conference on HHuman factors in computing systems* (pp. 195–202). New York, New York, USA: ACM Press. doi:10.1145/503411.503412

Demiralp, C., Jackson, C. D., Karelitz, D. B., Zhang, S., & Laidlaw, D. H. (2006). CAVE and fishtank virtual-reality displays: a qualitative and quantitative comparison. *IEEE transactions on visualization and computer graphics*, *12*(3), 323–30. doi:10.1109/TVCG.2006.42

Demiralp, C., Laidlaw, D., & Jackson, C. (2003). Subjective usefulness of CAVE and fish tank VR display systems for a scientific visualization application. In *Proceedings of the 14th IEEE Visualization (VIS'03)* (p. 86). IEEE Computer Society.

Dietrich, A., & 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).

Difede, J., & Hoffman, H. G. (2002). Virtual Reality Exposure Therapy for World Trade Center Post-traumatic Stress Disorder: A Case Report. *CyberPsychology & Behavior*, *5*(6), 529–535. doi:10.1089/109493102321018169

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. *Princeton NJ Educational Testing Service* (Vol. 102, pp. 109–113). Princeton, New Jersey: Educational Testing Service. doi:10.1073/pnas.0506897102

Espinoza, M., Banos, R. ., Garcia-Palacios, A., Alcaniz, M., Botella, C., Cervera Grau, J. ., … De Sanctis, R. (2012). Promotion of emotional wellbeing in oncology inpatients using VR. *Studies in Health Technology and Informatics*, *181*, 53–57.

Ferwerda, J. A. (2003). Three varieties of realism in computer graphics. *Proceedings SPIE Human Vision and Electronic*, *3*, 290–297. doi:10.1117/12.473899

García-Rodríguez, O., Pericot-Valverde, I., Gutiérrez-Maldonado, J., Ferrer-García, M., & Secades-Villa, R. (2012). Validation of smoking-related virtual environments for cue exposure therapy. *Addictive behaviors*, *37*(6), 703–8. doi:10.1016/j.addbeh.2012.02.013

Goldin, S., & Thorndyke, P. (1982). Simulating navigation for spatial knowledge acquisition. *Human Factors: The Journal of the the Human Factors and Ergonomics Society*, *4*(4), 457–471. doi:10.1177/001872088202400407

Gomes de Sá, A., & Zachmann, G. (1999). Virtual reality as a tool for verification of assembly and maintenance processes. *Computers & Graphics*, *23*(3), 389–403. doi:10.1016/S0097-8493(99)00047-3

Gruchalla, L. (2004). Immersive well-path editing: investigating the added value of immersion. In *In Virtual Reality, 2004. Proceedings, IEEE* (pp. 157–164). IEEE.

Guilford, J., & Zimmerman, W. (1948). The Guilford-Zimmerman Aptitude Survey. *Journal of applied Psychology*, (7), 24–34.

Guilford, J., & Zimmerman, W. (1981). The Guilford-Zimmerman Aptitude Survey manual of instruction and interpretation. *CA: Consulting Psychologists Press*.

Hagen, M. (1986). *Varieties of Realism: Geometries of representational art*. (C. U. Press, Ed.). New York: CUP Archive.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. L. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176. doi:10.1016/j.intell.2005.09.005

ICIDO. (2010). ICIO: Explore. Retrieved June 12, 2013, from http://www.icido.de/en/Products/VDP/IDO_Explore.html

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

Ishikawa, T., & 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), 93–129. doi:10.1016/j.cogpsych.2005.08.003

Johnson, D. M., & Stewart, J. E. (1999). Use of virtual environments for the acquisition of spatial knowledge: Comparison among different visual displays. *Military Psychology*, *11*(2), 129–148.

Jones, C. M., & Braithwaite, V. a. (2003). The evolution of sex differences in spatial ability. *Behavioral neuroscience*, *117*(3), 403–411.

Kasik, D., Troy, J., & Amorosi, S. (2002). Evaluating graphics displays for complex 3d models. *Computer Graphics and Applications, IEEE*, *22*(3), 56–64.

Kimura, D. (1992). Sex differences and the brain. *Scientific American*, *267*(3), 118–125.

Kirschen, M. P., Kahana, M. J., Sekuler, R., & Burack, B. (2000). Optic flow helps humans learn to navigate through synthetic environments. *Perception*, *29*(7), 801–18.

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cincinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *Journal of Motor Behavior*, *22*, 19–43.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, *129*(5), 723–746.

Kuntze, M. F., Stoermer, R., Mager, R., Roessler, A., Mueller-Spahn, F., & Bullinger, A. H. (2001). Immersive virtual environments in cue exposure. *Cyberpsychology and Behavior*, *4*(4), 497–501.

Kyritsis, M., & Gulliver, S. R. (2009). Gilford Zimmerman orientation survey: A validation. In *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on* (pp. 1–4). IEEE. doi:10.1109/ICICS.2009.5397592

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

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

Li, A., Montaño, Z., Chen, V. J., & Gold, J. I. (2011). Virtual reality and pain management: current trends and future directions. *Pain*, *1*(2), 147–157. doi:10.2217/pmt.10.15

Loomis, J. M., & 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* (pp. 21–46). Mahwah, NJ: Lawrence Erlbaum Associates.

Mania, K., Wooldridge, D., Coxon, M., & 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), 396–404. doi:10.1109/TVCG.2006.55

Mavrikios, D., Karabatsou, V., Fragos, D., & Chryssolouris, G. (2006). A prototype virtual reality-based demonstrator for immersive and interactive simulation of welding processes. *International Journal of Computer Integrated Manufacturing*, *19*(3), 294–300. doi:10.1080/09511920500340916

Mcmahan, R. P., Bowman, D. A., Brady, R. B., North, C., & Polys, N. F. (2011). Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games Interaction for Virtual Reality Games.

Meijer, F., Geudeke, B. L., & Van Den Broek, E. L. (2009). Navigating through virtual environments: visual realism improves spatial cognition. *Cyberpsychology and Behavior*, *12*(5), 517–521.

Miyahira, S., Folen, R., Hoffman, H. G., Garcia-Palacios, A., Spira, J., & Kawasaki, M. (2012). The effectivness of VR exposure therapy for PTSD in returning warfighters. *Annual Review of Cybertherapy and Telemedicine 2012: Advanced Technologies in the Behavioral, Social and Neurosciences*, *181*, 128–132.

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a “Virtual” Maze: Sex Differences and Correlation With Psychometric Measures of Spatial Ability in Humans. *Evolution and Human Behavior*, *19*(2), 73–87.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. Frank & I. Campari (Eds.), *Spatial Information Theory A Theoretical Basis for GIS* (Vol. 716, pp. 312–321). Springer Berlin Heidelberg. doi:10.1007/3-540-57207-4_21

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* (pp. 143–154). New York: Oxford University Press.

Montello, D. R., Lovelace, K. L., Golledge, R. G., & Self, C. M. (1999). Sex • Related Differences and Similarities in Geographic and Environmental Spatial Abilities. *Annals of the Association of American Geographers*, *89*(3), 515–534.

Montello, D. R., & Sas, C. (2006). Human factors of wayfinding in navigation. *In: Karwowski W (ed) International encyclopedia of ergonomics and human factors*, *3*(2), 2003–2008.

Ni, T., Bowman, D. A., & Chen, J. (2006). Increased display size and resolution improve task performance in Information-Rich Virtual Environments. In *In Proceedings of Graphics Interface 2006* (pp. 139–146). Canadian Information Processing Society.

Philbeck, J. W., Klatzky, R. L., Behrman, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilitates nonvisual navigation. *Experimental Psychology: Human Perception and Performance*, *27*(1), 141–153.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, *23*(12), 1447–1455.

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & cognition*, *27*(4), 741–50.

Richardson, A. E., Powers, M. E., & Bousquet, L. G. (2011). Video game experience predicts virtual, but not real navigation performance. *Computers in Human Behavior*, *27*(1), 552–560. doi:10.1016/j.chb.2010.10.003

Riecke, B. E., Behbahani, P., & 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* (pp. 15–18). Chania, Crete, Greece: ACM.

Riecke, B. E., Cunningham, D. W., & Bülthoff, H. H. (2007). Spatial updating in virtual reality: the sufficiency of visual information. *Psychological Research*, *71*(3), 298–313. doi:http://dx.doi.org/10.1007/s00426-006-0085-z

Rizzo, A., & Graap, K. (2008). Virtual Iraq: initial results from a VR exposure therapy application for combat-related PTSD. *Studies in Health Technology and Informatics*, *132*, 420–5.

Rodger, J., & Browne, R. (2000). Choosing rendering parameters for effective communication of 3d shape. *Computer Graphics and Applications, IEEE*, *20*(2), 20–28.

Sandstrom, N. J., Kaufman, J., & Huettel, S. a. (1998). Males and females use different distal cues in a virtual environment navigation task. *Brain research. Cognitive brain research*, *6*(4), 351–60.

Satalich, G. (1995). *Navigation and wayfinding in virtual reality: Finding the proper tools and cues to enhance navigational awareness*. University of Washington.

Schuchardt, P., & Bowman, D. A. (2007). The benefits of immersion for spatial understanding of complex underground cave systems. In *In Proceedings of the 2007 ACM symposium on Virtual reality software and technology* (pp. 121–124). ACM.

Schulte-Pelkum, J., Riecke, B. E., von der Heyde, M., & Bülthoff, H. H. (2003). Circular vection is facilitated by a consistent photorealistic scene. In *Proceedings of Presence 2003*.

Sherry, D. ., & Hampson, E. (1997). Evolution and hormonal control of sexually dimorphic spatial abilities in humans. *Trends in Cognitive Sciences*, *1*(2), 50–56.

Sigurdarson, S., Milne, A. P., Feuereissen, D., & Riecke, B. E. (2012). Can physical motions prevent disorientation in naturalistic VR? In *Virtual Reality Workshops (VR), 2012 IEEE* (pp. 31–34). IEEE. doi:10.1109/VR.2012.6180874

Silverman, I., & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. In *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 533–549). Oxford, U.K.: Oxford University Press.

Slater, M., Khanna, P., & Mortensen, J. (2009). Visual realism enhances realistic response in an immersive virtual environment. *Computer Graphics and Applications, IEEE*, *29*(3), 76–94.

Somers, W. W., & Hamilton, M. J. (1984). Estimation of the stereoscopic threshold utilizing perceived depth. *Ophthalmic and Physiological Optics*, *4*(3), 245–50.

Sousa Santos, B., Dias, P., Pimentel, A., Baggerman, J.-W., Ferreira, C., Silva, S., & Madeira, J. (2008). Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimedia Tools and Applications*, *41*(1), 161–181. doi:10.1007/s11042-008-0223-2

Springer, S., & Deutsch, G. (1985). *Left Brain, Right Brain*. *Saturday Review* (Vol. 2). San Francisco: W.H. Freeman.

Steinicke, F., Bruder, G., Hinrichs, K., Lappe, M., Ries, B., & Interrante, V. (2009). Transitional environments enhance distance perception in immersive virtual reality systems (pp. 19–26).

Swindells, C., & Po, B. (2004). Comparing CAVE, wall, and desktop displays for navigation and wayfinding in complex 3D models. In *Computer Graphics International, 2004. Proceedings* (pp. 420–427). IEEE.

Tan, D. S., Gergle, D., Scupelli, P., & Pausch, R. (2006). Physically large displays improve performance on spatial tasks. *ACM Transactions on Computer-Human Interaction*, *13*(1), 71–99. doi:http://doi.acm.org.proxy.lib.sfu.ca/10.1145/1143518.1143521

Thompson, W. B., Gooch, A. A., Creem-regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments ? *Presence: Teleoperators and Virtual Environments*, *13*(5), 560–571.

Thorndyke, P., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, *14*(4), 560–589.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, *47*(2), 599–604.

Waller, D. (1999a). *An assessment of individual differences in spatial knowledge of real and virtual environments*. University of Washington.

Waller, D. (1999b). An assessment of individual differences in spatial knowledge of real and virtual environments.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, *6*(4), 307–321. doi:10.1037//1076-898X.6.4.307

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

Wallet, G., Sauzéon, H., Pala, P. A., Larrue, F., Zheng, X., & 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), 417–23. doi:10.1089/cyber.2009.0187

Weatherford, D. L. (1982). Spatial cognition as a function of size and scale of the environment. *New directions for child development: Children's conception of spatial relationships*, (15), 5–18.

Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization* (pp. 35–38). ACM. doi:10.1145/1012551.1012558

Wittenberg, G. (1995). Training with virtual reality. *Assembly Automation*, *15*(3), 12–14.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in cognitive sciences*, *14*(3), 138–46. doi:10.1016/j.tics.2010.01.001

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