Evaluating the impact of display medium (Virtual Reality Head Mounted Display vs. Screen Display) on perceived user experience and aesthetic value within virtual hotel environments.

A Masters Thesis Presented to

The Faculty of the Graduate School at

Cornell University

In Partial Fulfillment of

The Requirements for the Degree

Master of Science in Human Environment Relations

Concentration: Human Factors Ergonomics

By

Ethan H. Arnowitz

Dr. So-Yeon Yoon, Committee Chair

Dr. Erik Andersen, Minor Member

January 2017

Dedicated to the thinkers of tomorrow who

dare to redefine the impossible

Acknowledgements

I would first like to thank my parents, Richard and Shari Arnowitz, for their support and guidance throughout both my academic career and personal life. Without you, none of this would be possible.

I would also like to thank Dr. So-Yeon Yoon for her guidance and patience as my advisor. I was extremely fortunate to have someone who supported my goals, while giving me the creative freedom to explore. Dr. Yoon has made this journey an incredibly rewarding experience. I would also like to acknowledge my minor member, Dr. Erik Andersen, for his advice and guidance throughout development and analysis of my thesis project. Finally, I would like to thank Dr. Don Greenberg for introducing me to new technologies, and his inspiring advice.

Table of Contents

Abstract

This study aims to understand the impact of display medium on affective and cognitive perceptions within virtual environments. Virtual reality (VR) has become a popular design visualization tool, and head-mounted displays (HMDs) offer a potentially distinctive way of displaying VR environments. As HMDs become increasingly accessible, industry leaders are beginning to explore their use as a design prototyping, visualization, and communication tool.

Before HMDs become a new medium for visualization, it is imperative to understand the perceptual and user experience (UX) differences between HMDs and screen displays. In the present study, participant's rate perceived aesthetic value, spatial presence, and affective reactions (i.e. emotions) after touring two virtual hotel environments, one on a stereoscopic HMD, and one the second on a 65" 4K Ultra HD (UHD) screen display.

Using a 2 (display medium types) X 2 (displayed environments) within subject's design, this study was conducted in a lab experiment with 80 participants, split between design and non-design backgrounds. Randomization of both the hotel environment and display medium reduced any potential carryover effects between environments and displays.

Within each hotel environment, participants toured a series of three interior perspectives. After exiting each hotel, participants completed a questionnaire measuring UX and perception variables. UX, i.e. process, variables (spatial presence and affective reactions) were measured through the MEC-Spatial Presence Questionnaire, self-assessment manikin (SAM), and semantic differential. An expanded version of the visual properties questionnaire measured aesthetic value. The effect of display medium across all variables, environments, and disciplines are explored and recommendations for future HMD specific applications are suggested.

Keywords: virtual reality, spatial presence, affective reactions, aesthetic value, head mounted display (HMD), perception, prototyping

Evaluating the impact of display medium (Virtual Reality Head Mounted Display vs. Screen Display) on user experience and perceived aesthetic value within virtual hotel environments.

Introduction

This study aims to understand the impact of display medium on perceptions within virtual reality (VR) environments. Although there is a resurgence of commercial head-mounted display's (HMDs), there is a fundamental lack of novel HMD-specific applications. These applications are critical to the success of HMDs at a consumer and enterprise scale. Advances in HMD hardware has reached a quality threshold that improvements in the immediate future will not drastically affect user-experience. Because of this, many industries and academic institutions are now shifting their attention to developing novel HMD specific applications.

The design industry is particularly interested in HMDs as a design and visualization tool. Designers, including interior design, industrial design, architecture, landscape architecture, and apparel design, use various forms of computer-aided design (CAD) throughout their design process. Studies across multiple industries have suggested that directly integrating digital design tools into the process, rather than just at the production level, can lead to a more successful product (Wheelwright & Clark, 1994).

Design industries typically use computer generated renderings and interactive virtual experiences to prototype and receive feedback through an iterative design process. These visualizations allow designers to gather feedback from team members, clients, and additional stakeholders. Although historically VR is viewed on a screen display, recent innovations in HMDs increase the potential for its use as a VR visualization tool. HMDs offer a potentially unique digital prototyping medium.

Before design-specific HMD applications can be developed, it is imperative to understand fundamental questions centered on the perception and experience while using these headsets. This study explores affective and cognitive perceptions between HMDs and more ubiquitous display mediums i.e. screen displays.

First, literature on the importance of prototyping and visualization throughout the design process is discussed, with a focus on virtual reality as a cost-effective alternative to traditional prototyping mediums. Next, the history of virtual reality as a design prototyping tool is analyzed, focusing on the evolution of different VR display mediums. Discussion on the history of HMDs, with a focus on how current headsets overcome failures of the past, explains the current feasibility of creating HMD applications. This increase in accessible HMD technology leads to the current study and methods.

Definitions

"Virtual reality" (VR) is often used as an umbrella term to refer to a range of devices and computer generated experiences. For this study, virtual reality is defined as a computer-generated virtual representation of an object or environment. Various mediums, including print renderings, screen displays, and head mounted displays (HMDs), have visualized VR experiences. Print renderings are images created with computer-generated 3D modeling and rendering packages. A screen display is a monoscopic display with external user input for interactions. HMDs are a wearable stereoscopic display with integrated tracking technology. Both HMDs and screen displays can visualize interactive VR experiences, but offer key distinctions explored throughout this paper.

Prototyping in the Design Process

Multidisciplinary teams including designers, engineers, product managers, and end-users can significantly improve the overall successes of a product (Cooper & Kleinschmidt, 1986; Srinivasan et al., 1997). Studies have shown that an integrated approach involving all stakeholders, can lead to a more efficient design process (Wheelwright & Clark, 1994). Different disciplines bring a unique perspective to the complex process of bringing a product or building from concept through construction. Throughout the design process, communication between team members is critical to identify problems, respond to design preferences, and meet project goals (Koskinen, 2012; Söderman, 2005).

The challenge is that communication between stakeholders is fragmented due to differences in background, knowledge, and priorities. Different disciplines often have specialized languages that make it hard to communicate with others outside their discipline, even if they share a common goal (Canter, 1969; Droz, 1992). This lack of communication is especially salient in the early stages of the design process when a product is not visualized, leaving team members with nothing to see, hold, and focus on (Söderman, 2005). This phase of the design process is critical because most of a product's manufacturing costs are determined during the initial phases of its design process (Hayes & Wheelwright, 1976).

Additionally, multidisciplinary team members have different priorities that affect their ability to make objective decisions regarding cost, usability, and aesthetic preferences (Droz, 1992). Typically, designers focus on aesthesis and usability; engineers focus on functionality and production, and product manager's focus on cost and marketing.

A common way for different members to communicate their priorities is through prototypes or design visualizations. Studies indicate that visual prototypes are better than verbal descriptions as communication tools during the design process (Holbrook & Moore, 1981). Prototypes provide a visual representation of an object in a neutral language and background allowing team members to communicate and increase the likelihood of developing consensus in the critical early stages of the design process. Prototypes often confirm initial ideas that were unrealistic, ill-conceived, or not imaginative enough (Schrage, 2006). Additionally, prototypes can keep different disciplines focused on the user experience, and not on their point of view (Droz, 1992). The method, quality, and type of prototype often depends on the design process. Different design processes provide different time allotments to the prototyping process.

Studies have supported that creating multiple prototypes before finalizing a concept to commercialize can ultimately lead to a more successful product (Kaulio, 1998; Schrage, 2006; Stevens & Burley, 1997; Srinivasan et al., 1997; Wheelwright & Clark, 1994). Iterative prototypes give team members the freedom to make mistakes and eliminate alternative ideas when there is little time and equity in the design (Droz, 1992). The importance of iterative prototyping applies in particular to the high risks associated with bringing a new product to market or constructing a building.By testing multiple ideas in parallel, prototypes can help reduce this uncertainty and cost (Dahan & Srinivasan, 1998).

It is imperative to consider various ideas because only a small percentage of new ideas ultimately prove to be profitable (Stevens & Burley, 1997). Additionally, the iterative prototyping process generates new product-specific vocabulary, identifying new and critical features. The development of new vocabulary provides an opportunity to analyze the relative value of specific design elements, reducing the likelihood of adding unnecessary features (Schrage, 2006).

Documentation and testing of multiple interactive prototyping processes across a range of industries provide structure for companies to build better products through better prototypes (Schrage, 2006). The user-centered product development is a human factors/ergonomic approach applied to industries including apparel design, interior design, industrial design, and urban planning projects (Kaulio, 1998). This method allows designers to make an initial prototype from a developed set of user requirements. End-users can then provide feedback by interacting with prototypes, and designers can transform this feedback into new iterative prototypes until reaching a final design.

Alpha, i.e. concept, and beta testing allows designers to involve end-users in the initial and final design phases. Alpha testing is when a designer displays a prototype to analyze specific responses while finalizing the initial concept. Respectively, beta testing is a similar approach applied in the later phases of the design process that aims to determine if the product does what it was designed to do (Kaulio, 1998). Both alpha and beta testing are dependent on high-quality prototypes to elicit feedback that further informs the iterative design process.

The information acquisition process addresses uncertainties and gathers information about a product through iterative prototyping. This process allows companies to develop a product from an initial conceptual idea, with few specific goals, to a finalized product with materials, tolerances, and requirements stated in high detail (Srinivasan et al., 1997).

The speed of product development often limits a firm's innovation (Droz, 1992). Due to time and cost, firms are often required to limit the number of prototypes a team can generate. This forces team members to make decisions with incomplete information about the design, quality, and cost of the product. Team members with a higher level of spatial presence, such as designers or architects, may be able to fill in missing space-related information, but not all stakeholders have this capacity (Vorderer et al., 2003).

Virtual reality (VR) has the potential to reduce these limitations by creating iterative high-quality three-dimensional prototypes. Since VR prototypes cost considerably less to build and test than physical prototypes, design teams using VR can develop a higher number of concepts. (Dahan & Srinivasan, 1998).

Virtual Reality as a Visualization Tool

VR has become a universal way for designers to communicate and iterate their ideas throughout the design process. Although VR has become a ubiquitous prototyping and visualization tool, the way in which designers display their virtual designs has radically changed over time. With the introduction of new display mediums, specifically HMDs, the way in which we view, interact, and perceive the virtual world is distinct from the past.

Before VR, designers would use various forms of product visualizations such as clay sculptures or foam models (Schrage, 2006). Models provide a 3D tangible object that cannot translate into a 2D image. This 3D experience allows stakeholders to move around, examine and provides a platform for feedback. This visual platform is of particular importance for stakeholders without design backgrounds to conceptualize how the product will translate from a prototype to a product (Canter, 1969). By having a model to interact with, stakeholders can put themselves in the position of the end-user to identify and remedy potential problems (Droz, 1992). Although these models provide a physical prototyping experience, the process to make them is expensive, time consuming, and often produces untouchable works of art rather than a highly adaptable prototype (Schrage, 2006).

Virtual reality can offer many unique opportunities that are not possible in the real world. Creating a computer generated visualization of an environment or object is often a more cost-effective method than creating a physical prototype, due to the lack of physical materials and resources needed. Studies have also suggested that virtual representations can provide the same visualization results as physical prototypes (Dahan & Srinivasan, 1998).

VR can also allow a designer to iterate and test several prototypes in parallel without increasing cost significantly (Dahan & Srinivasan, 1998). Iterative prototyping leads to fewer changes during the production process, and ultimately saves costs. Additionally, VR allows designers to display their ideas in different scales, orientations, and viewpoints to create a visualization experience that may not be possible with physical prototypes.

Besides visualization, VR has become a tool for researchers to explore perception in real-world experiences. VR has the potential to be a more naturalistic way of creating stimuli that can translate knowledge obtained in from a lab setting to real world applications (Slater et al., 1996).

Before VR, perceptual studies often restricted stimuli to isolate sensory information. These studies were carried out in a way that was not natural to how participants interacted in comparable real-world experiences (Wheelwright & Clark, 1994). A similar problem existed for companies conducting user tests, where researchers would bring customers into a lab and conduct observational studies with prototypes. The constrained environment of these studies offered little control to the researcher and was often carried out in a way that was not generalizable to real-world experiences (Cooper & Kleinschmidt, 1986).

As VR becomes increasingly accessible, the potential for researchers to use it as an experimental stimulus is increasing. VR is now allowing researchers to analyze human perception in a sterile lab setting, with highly realistic, controllable, and interactive stimuli (Scarfe & Glennerster, 2015).

Typically, screen displays visualize VR environments. Head mounted display's (HMDs), have the potential to change the way virtual reality is displayed. HMDs can influence both the design prototyping process, as well as stimuli for researchers to understand perception in real-world applications. Before HMDs become a ubiquitous display medium, it is imperative to understand perceptual differences between HMDs and screen displays.

History of VR Display Mediums

The way in which users view and experience VR has dramatically shifted with advances in technology. With the advent of the personal computer, the design process slowly shifted from pencil and paper to 2D CAD drawings on a screen display. As computational power increased, 3D modeling software and computer-generated images have become a universal way for designers to visualize their ideas. Although these pictures provide a realistic way for designers to translate their thoughts, the ability to manipulate perception and the lack of interactivity leaves a major disconnect between the conceptual image and the physical product.

As computational power increased over time, interactive VR experiences became a popular visualization and prototyping tool. VR provides an opportunity for users to navigate and view a virtual environment with 360° of freedom. These experiences are designed for a large scale (>40") screen displays where the user controls movement through a mouse, joystick, or keyboard. Although this provided a heightened level of interactivity, there is still disconnect between the virtual visualization and the finished product.

Companies began exploring the use of HMDs starting in the early 2010's as a method of portraying virtual reality environments. Due to increases in computational power, display technology, and tracking hardware, this was the first time in history where high quality, low-cost headsets became accessible.

History of Head Mounted Displays

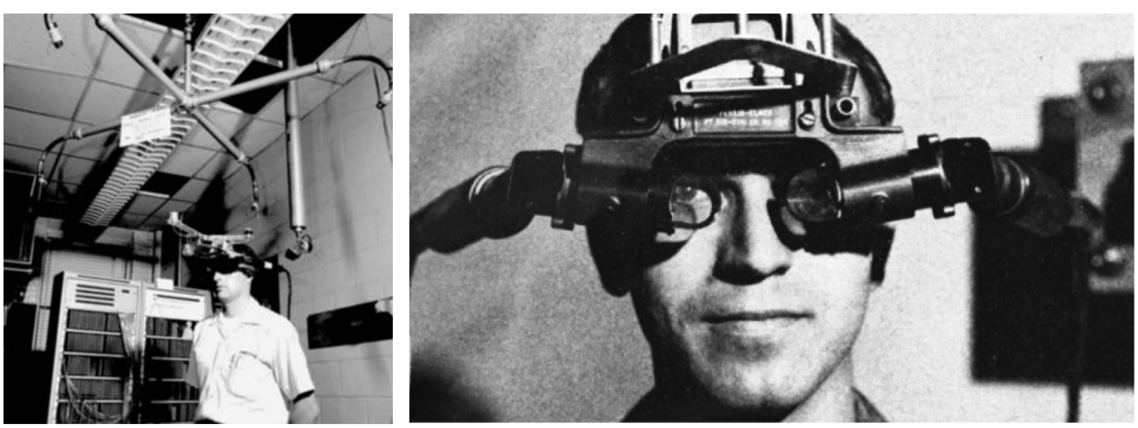
Head-mounted displays (HMDs) offer the opportunity to provide a more natural way of interacting with a VR environment (Williams et al., 1993). HMDs place duplicate two-dimensional images over each eye to create the illusion of a depth and three-dimensional environments (Ahn et al., 2014; Robinett & Rolland, 1992; Sutherland, 1968). Additionally, integrated tracking technology parallels the user's movement from the real to the virtual world. Although HMDs have recently emerged as a potentially unique display medium, the history of HMDs is complicated, with multiple failed attempts. Analyzing this history, and the failures of previous headsets, is essential in understanding why the latest generation of HMDs are distinctive from failures of the past.

Since consumer headsets were not available until the 1990's, studies on HMDs until this time would need to custom design special hardware and software. These headsets would often contain a limited set of functions, geared towards the specific experimental design or application. Additionally, these headsets were often massive with low performance and visual fidelity. Because of this, the accessibility of HMDs in both the consumer, academic, and enterprise sector was extremely limited.

Ivan Sutherland created the first HMD prototype at the University of Utah in 1968 (Sutherland, 1968). Sutherland's original headset design included two cathode ray tubes (CRTs) placed over each eye, providing a 40° field of view (FOV).

The display was too heavy to wear, so it was mounted to the ceiling and placed over the user's head (Figure 1-1). Multiple forms of tracking, including both mechanical and ultrasonic, were integrated into the headset limiting movement. The hardware, including the mechanical tracking technology, allowed the user to move within a 3x3x6 ft. volume. The size of the headset limited the up-down head tilt to 40°. At the time, no commercial computer was powerful enough to render the images for the display. Due to the lack of computational power, the VR environment was created with a simple wireframe visualization, instead of opaque objects with hidden lines. Special computing technology was set up to power the low-resolution displays at 30 Hz, which still provided latency to the end user.

Figure 1-1
Ivan Sutherland's Head Mounted Display



Through initial pilot tests, Sutherland concluded that the limited movement of the hardware minimized the space that the user could navigate, ultimately leading to the misinterpretation of shapes and scale. Although this headset was revolutionary, the limitations of computational power, display hardware, and tracking technology made the realization of HMDs unrealistic at this time.

During the 1990's, there was an explosion of commercial HMDs that came onto the market (See Figure 1-2). Although these headsets offered many improvements from previously independently developed headsets, many of the initial problems encountered by Sutherland still existed. Although display technology advanced, the lack of computational power made displaying high-resolution images (>1080x1200 per eye) without latency (>60 Hz) unattainable.

Figure 1-2
Virtuality Headset and Virtual Environment (1990's)

Additionally, limited tracking made mirroring movement from the real to the virtual environment impossible without interference. Interference and lag results in pixelated, distorted, and a low contrast image of the virtual world that drags about movements in the real world (Scarfe & Glennerster, 2015; Williams et al., 1993). Eliminating latency is critical because it has both behavioral and psychological effects, including VR sickness, headaches, diplopia, blurred vision, and sore eyes or eye strain (Scarfe & Glennerster, 2015; Williams et al., 1993; Williams et al., 1998).

Recently as 2005, researchers were forced to output to a low-resolution HMD (340x480 per eye) to meet the necessary refresh rate (Söderman, 2005). Additionally, precise mounting and adjustability of the HMD on the head/eyes were seldom possible. The lack of adjustable HMDs led to a compromise between comfort and visual presentation.

Headsets throughout the 1990's were bulky in size, heavy to wear, and not portable. Many ergonomic issues limited the use and implementation of these displays for novel applications. Additionally, the price of these headsets was extremely high and only accessible by government organizations or corporations. Because of the low-quality experiences and high price point, the initial mass-market development of HMDs throughout the 1990's ultimately failed.

Current State of Head Mounted Displays

Within the last ten years, HMD technology has become increasingly accessible to companies, researchers, and developers exploring potential applications. With the release of high-quality, low-cost commercial headsets, different industries are beginning to explore potential applications and fundamental research questions centered on HMD applications.

Until 2016, no commercial high–quality HMDs were available on the market. Beginning in 2013, companies such as Facebook's Oculus and HTC started releasing developer kits to initiate growth of HMD hardware. Due to Moore's law, the exponential growth of computational power increased to a point allowing HMDs to offer high resolution and refresh rates (90 Hz) with real-time tracking, raising the acceptability threshold of HMD experiences. (Figure 1-4; Di Gironimo & Guida, 2013; Scarfe & Glennerster, 2015). Besides computational power, the improvement in tracking technology has allowed for a more accurate translation of movement from the real to virtual world.

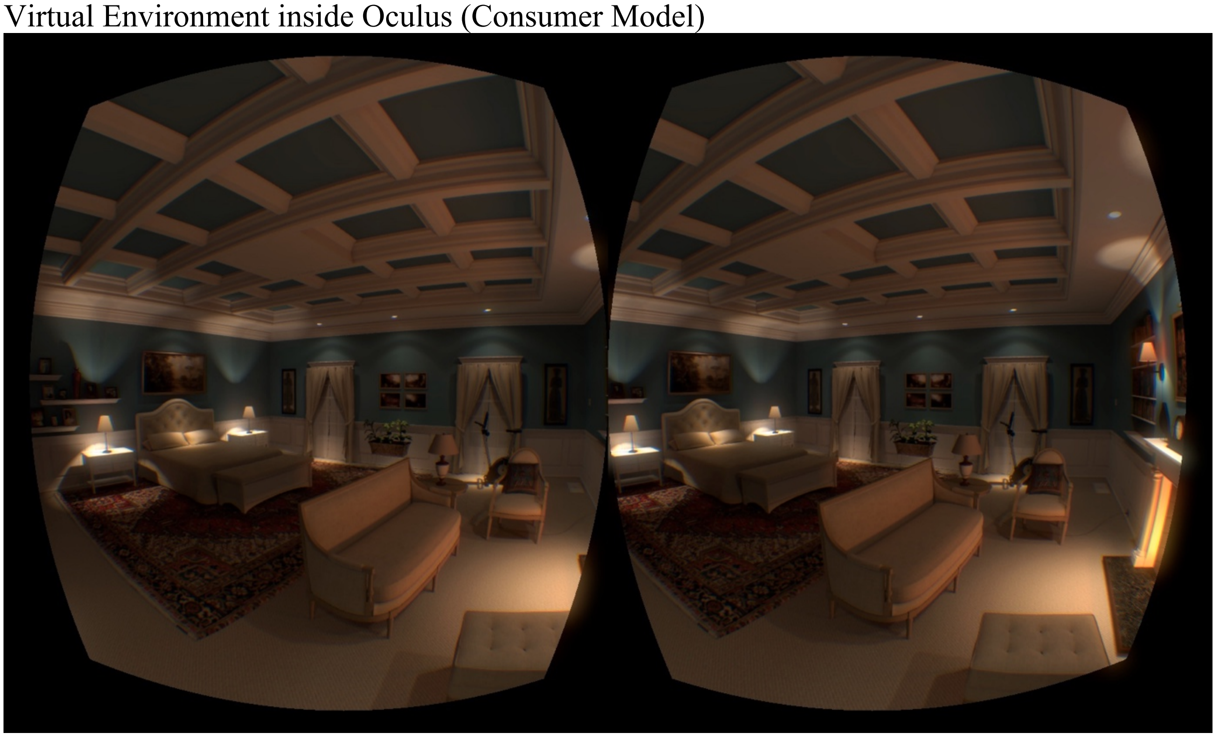
Figure 1-3
Progression of Oculus HMDs

Oculus Developer Kit 1
March 2013

Oculus Developer Kit 2
July 2014

Oculus (Consumer Model)
March 2016

Figure 1-4
Virtual Environment inside Oculus (Consumer Model)



From 2013-2016, industry leaders developed HMD developer kits until commercially released versions came onto the market (See Figure 1-3). The decrease in size, weight (<.6 kg), and improvements to the field of view (110 degrees), increased the usability of these HMDs. The headset itself directly integrates all tracking hardware, simplifying the setup process of putting on and adjusting the headset. Tracking sensors positioned on the wall allow the user to walk with minimal interference. Although these headsets currently require tethering to a computer, they still provide a comfortable experience that allows the user to move with limited interference up to a 15’ x 15’ walkable space. Additionally, the user has complete control to move their head in any direction or orientation. This freedom of movement is critical to providing a non-invasive experience applicable to both research and real-world applications.

Because of this increase in high quality, low-cost HMDs, researchers and developers no longer have to pay premium prices or create unique HMD systems. These commercially available headsets are the stepping-stone to widespread adoption in both research and real-world sectors.

Although multiple headsets have come onto the market, there is a fundamental lack of empirical research on the perception and user experience within HMDs. Understanding the perceptual differences between HMDs and more ubiquitous screen displays is critical to the success of HMD applications, specifically as a design prototyping and visualization tool.

Feasibility of Implementation- HMD vs. Screen Display

Until recently, besides hardware, the complex pipeline to create HMD experiences limited the accessibility of HMDs as a design prototyping tool. In a survey of 150 companies in 2004, data found that HMDs was the least used product representation during the design process.

This was explained by its demand for expensive investments and trained staff to be useful (Engelbrektsson & Söderman, 2004). Although VR was perceived to provide a complete understanding of a future product, many companies claimed it was unattainable at the time.

Today, the pipeline to create interactive virtual reality experiences for an HMD Vs screen display are not significantly different in terms of time, cost, and technical abilities needed. To create VR experiences, developers assemble all design assets (i.e. 3D geometry, materials, textures, and lighting) in a game engine. Through the release of HMDs onto the market, game engines such as Unreal, Unity, and Stingray, have released HMD specific templates. These templates have streamlined the process to make HMD virtual reality experiences, in particular for non-programmers and designers. HMD manufacturers, such as Oculus and HTC, are also providing direct integration with multiple game engines, providing a simple pipeline to render highly detailed virtual environments with minimal temporal lag (Scarfe & Glennerster, 2015).

Due to the increase in available hardware and software, design industries are exploring the adoption of HMDs as a design visualization and prototyping tool. It is critical to understand the distinct differences between HMDs and screen displays to validate why companies should invest in this new technology. The objective of the current study is to understand the perceptual differences between HMDs and screen displays, specifically regarding user experience process variables (i.e. spatial presence and affective reactions), and outcome variables (i.e. perceived aesthetic value).

Methods

Participants

This study included collecting original data in the DUET Lab (Design for User Experience with Technology) at Cornell University with 80 participants. To control for the possible effects of gender and design background, participants were evenly divided between men and women, as well as design and non-design backgrounds. This division resulted in a total of 20 male designers, 20 female designers, 20 male non-designers, and 20 female non-designers. Design majors included interior design, industrial design, apparel design, architecture, and landscape architecture. Subjects ranged in age from 18 to 36 years (M= 21.13 years, SE= .27), and all subjects had normal or corrected-to-normal vision with some form of prior VR experience. Recruiting was conducted through posters and social media, and participants were compensated with either a $10 gift card or extra credit in participating courses at Cornell University.

Virtual Reality Hotel Environments

Each participant toured a VR hotel environment on both the HMD and screen display. Due to potential carryover effects, both display conditions could not show the same hotel environment. Because of this, two virtual hotel rooms were used, including a luxury and economy condition. This use of two hotel conditions allows the participant to experience a different hotel environment on each display. Additionally, this provides the opportunity to analyze whether the effect size of variables was consistent across environment conditions (i.e. luxury vs. economy).

Participants were broken into four groups, allowing randomization of both the hotel environment and display medium. Twenty participants, including five male designers, five female designers, five male non-designers, and five female non-designers, were randomly assigned to one of four experimental groups. The four groups included:

Group 1: Luxury HMD; Economy Screen Display
Group 2: Economy Screen Display; Luxury Screen Display
Group 3: Economy HMD; Luxury Screen Display
Group 4: Luxury Screen Display; Economy HMD

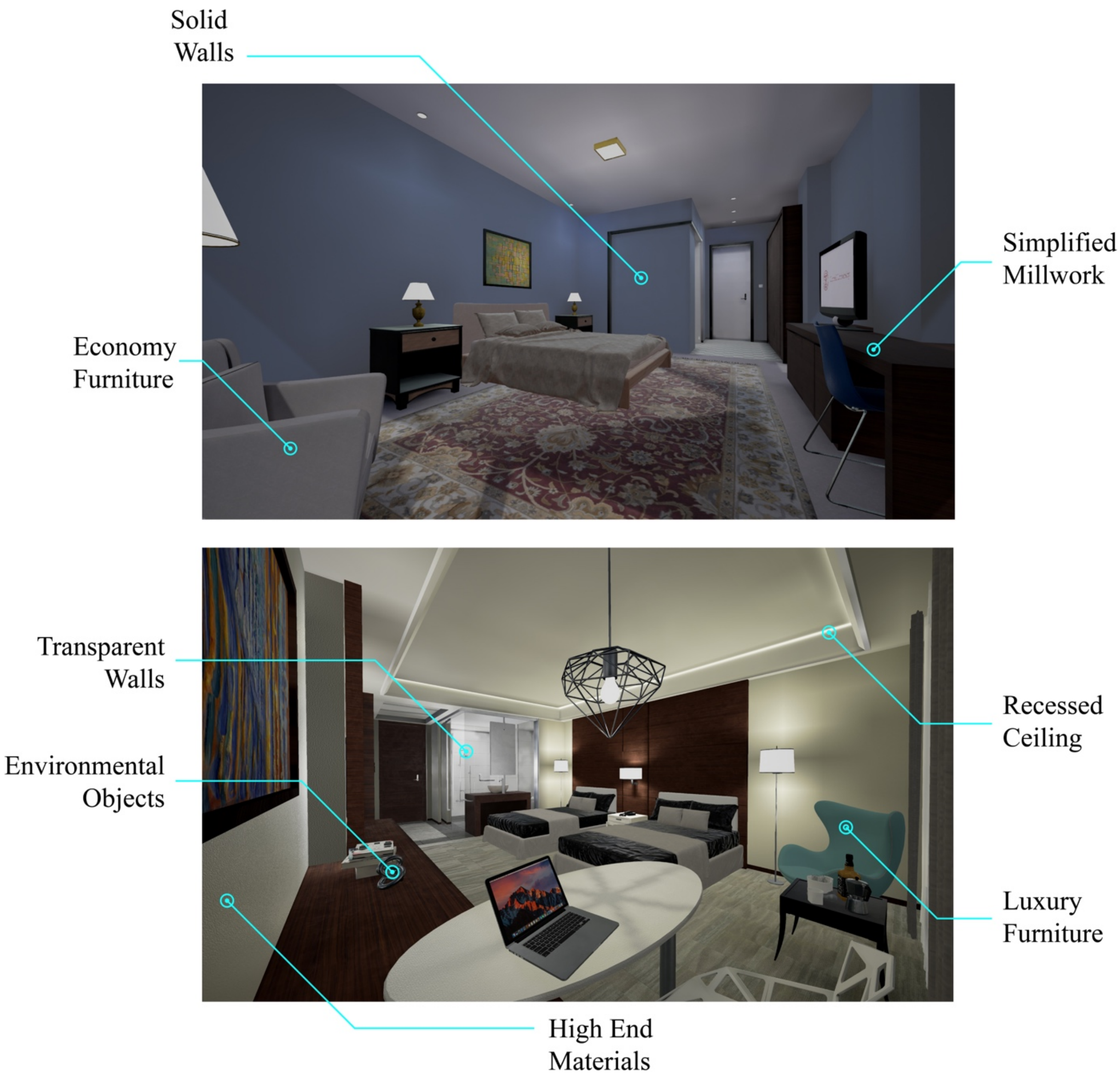
Common hotel elements guided the design of both virtual environments. Since the hotel environments were not a variable, but rather a background to test affective and cognitive perceptions, both environments created a universally perceived hotel aesthetic. Through an initial benchmarking process of collecting hotel photos, common hotel design elements were identified. Design elements include carpeting over wood flooring, warm over cool lighting, minimal patterns, and a focus on color and material.

Studies have suggested that high visual fidelity and realism can lead to a more accurate translation of space and heightened presence (Söderman, 2005Witmer & Singer, 1998). Because of this, both virtual hotel environments were highly realistic. The differences between the luxury and economy hotel conditions were chosen through an additional benchmarking process and included differences in terms of furniture, fixtures, architectural detail, materials, lighting, and environmental accessories (See Figure 2-1; See Appendix Figures C1 & C3).

The luxury environment featured highly detailed FF&E (furniture, fixtures, and equipment) with complex 3D geometry, in comparison to the standard FF&E throughout the economy environment. Additionally, the luxury environment included high architectural detail such as a recessed ceiling, large windows, and detailed TV cavity. The luxury environment also featured an elevated material palate with glass walls, wood cabinets, and chrome fixtures.

Besides detail and materiality, the luxury room featured more complex ambient lighting such as recessed and task lighting. Finally, the luxury room featured many environmental accessories such as computers, bottles, books, and sculptures, while the economy room featured no such accessories (Figure 2-1; See Appendix C). Although the two hotel environments were different, they shared the same layout and scale, i.e. square feet. Although the layout of the two environments was the same, mirroring of the environments minimized potential carryover effects between the two environment conditions.

Figure 2-1
Luxury (Top) and Economy (Bottom) Virtual Hotel Environments



Procedure

When participants first arrived they filled out a consent form outlining tasks, any negative potential side effects, voluntary participation, and their agreement to be filmed (See Appendix A). Participants were required to have normal or corrected-to-normal vision with some experience in VR (on either an HMD or a screen display) to participate in the study. After completing the consent form, participants toured one of two hotel environments.

The setup process for the HMD included placing the headset on, adjusting the tightness through three Velcro straps, and adjusting the interpupillary distance (IPD). The researcher demonstrated how to put the headset on properly through an initial set-up process. Once the headset was adjusted, participants were required to stand on a position marked with an X, orienting participants in the same location. The setup process for the screen display included placing the user on fixed location, situated 2 feet away from the screen display.

While in the hotel environment, the researcher guided participants through three pre-determined perceptive locations controlled through the gaming engine. At each perspective, the participant could fully pan around, providing a 360° view, but had no ability to move through the virtual environment. On the HMD, integrated tracking technology mirrored movement of the head from the real to virtual world (See Figure 2-2). On the screen display, moving a mouse panned the view around from the fixed location (See Figure 2-3). These interactions represent the most simplistic form of interaction respectively for each display medium. Complex interactions can create learning curves that may interfere with presence (Witmer & Singer, 1998). Restricting movement minimizes biases between the two display mediums, allowing the focus of the present study to be solely on perception, rather than interaction.

After existing each virtual hotel room, subjects completed a post-environment questionnaire measuring their sense of spatial presence, affective reactions, and perceived aesthetic value. Once completed, the participant would repeat the process by viewing a second hotel environment in the second display condition. After exiting, participants would complete the same questionnaire based on their experience in the second hotel environment. After the completion of the second questionnaire, participants were allowed to ask questions about the study or related technology through a debriefing session. Once completed, participants were compensated through a $10 gift card or extra credit in participating courses at Cornell University.

Figure 2-2
Participant using HMD

Figure 2-3
Participant using Screen Display

Scales

Perceived Aesthetic Value

Preference is an outcome of perceiving things and spaces and reacting regarding their potential usefulness and supportiveness (Kaplan, 1979). Aesthetic value has a significant effect on people's preference and behavior (Cetintahra & Cubukcu, 2015). When displaying a prototype to an interdisciplinary team, different disciplines will interpret aesthetics based on their background and individual preferences.

Aesthetic preferences from groups of people, rather than specific individuals or disciplines, is critical in understanding how design can meet user requirements (Canter, 1969). Due to the impact of aesthetics on preference, and the diversity of aesthetic preferences, it is critical to understand how disciplines interpret aesthetic value across different prototyping mediums.

Human observations are the only way to measure perceived aesthetic value (Watson et al., 2001). After participants exited each virtual hotel environment, questionnaires were distributed measuring participants perceived aesthetic value. In this study, an expanded version of the visual properties questionnaire measured perceived aesthetic value (Hanyu, 1997). The original questionnaire consists of four scales including (1) presence of particular elements; (2) cognition and processes; (3) physical patterns, and (4) lighting conditions. Multiple theories of formal and symbolic perceptions of environmental aesthetics influenced the formation of these scales (Hanyu, 1997; Hanyu 2000; Kaplan & Kaplan, 1989; Ulrich, 1993).

Under each scale, various sub-scales formed to define and quantify aesthetic value. The *presence of particular elements* included naturalness, nuisance elements, vehicles, and legibility. *Cognitive process* included mystery, openness, typicality, and familiarity. *Physical patterns* included complexity and coherence, and *lighting conditions* included brightness and uniform lighting (Hanyu, 1997).

Since the original visual properties questionnaire was developed to measure exterior facades, some of the original scales and subscales did not apply to the current study. The subscales of particular elements, including naturalness, nuisance elements, and vehicles, were omitted in the present study. Naturalness, which refers to the amount of vegetation in the scene, nuisance elements, which relates to the distraction of items such as wires and poles, and vehicles, which refers to the visibility of cars, were all not applicable to indoor hotel environments. All other subscales remained due to their relevance to interior environments.

A fifth scale, materiality, was added due to its importance in creating highly realistic VR environments (Witmer and Singer, 1998). Materiality questions were self-developed, and later analyzed for content and construct validity. The final distributed questionnaire contains 22 questions, spanning ten subscales. Questions were distributed using a 5-point Likert scale ranging from “not at all” to “a great deal” (See Appendix Table B3).

Spatial Presence

If a medium provides a high immersive experience, users may respond with feelings of spatial presence or the sensation of being located within the mediated environment. Spatial presence is defined by two main characteristics including the conviction of being located in a mediated environment, and the perceived possibilities to act (Kober & Neuper, 2013; Wirth et al., 2007).

In this study, the MEC- Spatial Presence Questionnaire (MEC-SPQ) measured spatial presence (Vorderer et al., 2004). The sub-scales of the MEC-SPQ include; *process factors* (attention allocation, spatial situation model (SSM), possible actions, self-location); *action states* (higher cognitive involvement, suspension of disbelief); and *user characteristics* (domain specific interest, visual-spatial imagery).

Through previous studies, the original MEC-SPQ was refined to create highly consistent and homogeneous versions with eight, six, and four items per subscale. In the present study, the six-item version was used, totaling 48 questions evaluating spatial presence (See Appendix Figures B1 & B2). The MEC-SPQ is a tool to measure the two-level model of spatial presence (Wirth et al., 2007).

In this model, the user must first create a spatial situation model (SSM), or a cognitive mental representation of the space portrayed through the media. Spatial cues from the mediated environment, as well as relevant user characteristics such as spatial memories and cognitions, both influence the SSM (Wirth et al., 2007). Visual-spatial imagery sub-scale measures spatial memories and cognitions, while the SSM sub-scale measures mediated effects.

Both involuntary and controlled attention can affect the formation of a SSM. Controlled attention, when the user directs attention because they want to, may have an impact on a user's attention through personal bias. Domain specific interest measures this user characteristic. Involuntary attention triggers attention without requiring the user to be consciously attentive. The sub-scale attention allocation measures involuntary attention and is influenced directly by the media. Since both involuntary and controlled attention have an impact on SSM, and therefore spatial presence, the MEC-SPQ accounts for both media and user characteristics to define spatial presence.

After the formation of the SSM, spatial presence can emerge through the PERF (primary egocentric reference frame) hypothesis. If users have built an SSM of the mediated environment, both the reference frame of the mediated and the real world is available to them. The PERF hypothesis states that the spatial environment represented in the media is the primary reference frame, and users position themselves and potential actions within the mediated space (Wirth et al., 2007). Witmer and Singer state that the degree of presence can range from the physical world to the mental world of the virtual environment (Witmer and Singer, 1998). If the PERF hypothesis is true, the user prioritizes the mental world over the physical world. Low-immersive media formats such as books can provide the opportunity to form a SSM, but users cannot achieve the PERF hypothesis because the mediated environments cannot transport them from the real to mediated world beyond imagination. The sub-scales spatial presence: possible actions and higher cognitive involvement measure the PERF.

An additional process that affects the PERF hypothesis is the suspension of disbelief (SoD). SoD is when a user does not pay attention to real-world stimuli and internal cognitions that distract from the mediated environment (Wirth et al., 2007). Suspending one's disbelief can strengthen the PERF hypothesis by removing factors that may contradict.

The two-level model of spatial presence was chosen for this study because it incorporates earlier theoretical frameworks of spatial presence into a more coherent conceptualization while connecting to established constructs from psychology and communication (Wirth et al., 2007). Additionally, this theory of spatial presence is applicable to the exposure of different media formats, i.e. both HMD and screen displays. The MEC-SPQ is validated across media formats including text, film, and virtual environments (Vorderer et al., 2004).

Affective Reactions

Understanding user experience (UX) within HMDs is critical to the adoption of HMDs in both consumer and enterprise applications. Due to low-quality HMD hardware, previous studies have documented the negative impact of HMDs on UX, leading to headaches, diplopia, blurred vision, and eye strain (Williams et al., 1993; Williams et al., 1998). With updated HMD hardware, researches can conduct evaluations while minimizing the negative impact on UX from studies of the past. A critical component of understanding UX is analyzing affective reactions, i.e. emotions. Interactions with environmental scenes elicit emotional responses that can be objectively measured and quantified. Measuring and understanding these reactions are necessary because, like perceived aesthetic value, affective reactions are critical to the perception and preference of objects and environments.

In this study, two measurements, including the semantic differential and self-assessment manikin (SAM), measure affective reactions. Russel and Mehrabian developed the semantic differential, which measures a set of basic emotional responses that are independent of the media involved (Russel & Mehrabian, 1974). This makes it a particularly appropriate measurement in environmental studies. Studies using the semantic differential have shown that three basic emotional reactions including pleasure, arousal, and dominance (PAD) characterize human emotions across a diverse range of stimuli. The theory of using PAD to measure affective reactions is based on the original work conducted by Osgood (Osgood, 1952). These three sub-scales account for significant variance among different stimuli, validating their importance in defining affective reactions (Bradley & Lang, 1994).

The sub-scales of affective reactions, pleasure, arousal, and dominance, are bipolar, meaning they can range from one extreme to another. For example, pleasure can range from extreme displeasure to extreme pleasure. Russel and Mehrabian developed an initial set of 28 adjective bipolar word pairings on intuition and represented a tentative set of descriptors for PAD (Russel & Mehrabian, 1974). Through initial studies, the 28 adjective pairings were reduced to a final set of 18, six respectively for pleasure, arousal, and dominance. Each bipolar pair features an adjective on each end of a 9-point scale, with five representing a neutral emotion (See Appendix Figure B4). Odd numbered questions included reverse scoring to reduce acquiescent and extreme response bias.

Besides the semantic differential, the self-assessment manikin (SAM) is a non-verbal pictorial assessment technique that directly measures pleasure, arousal, and dominance, and was used as an additional measure of affective reactions. (Bradley & Lang, 1994; Lang, 1980).

SAM includes three scale questions, one for pleasure, arousal, and dominance. Each scale includes 5 figure images, highlighting the bipolar nature of PAD. The scale for pleasure ranges from smiling, animated figure to a frowning, unhappy figure. The scale for arousal varies from an excited, wide-eyed figure to a relaxed, sleepy figure, while the scale for dominance is illustrated by changing the size of the figure relative to the background box (See Appendix Figure B5; Bradley & Lang, 1994). SAM has been used to measure emotions from a range of stimuli including images, sounds, and advertisements.

Display Mediums

Both the HTC Vive and the LG 65UF9500 represent the best commercially available HMD and screen display on the market at the time of this experiment. No HMD offers a higher resolution, refresh rate, field of view, or trackable area than the HTC Vive. Respectively, the LG 65UF9500 includes a 65” UHD 4K resolution, the standard for high-quality screen displays. These technologies represent what industry leaders are using to develop and display VR design visualizations and prototypes in 2016.

The FOV for the environments is 110° for both the screen and HMD conditions. This FOV provides a comfortable yet realistic translation of space in both display conditions. The field of view in the HMD is defaulted to 110° by the manufacturer (HTC). The height of the perspective for the screen condition environments is 5’10”, while the viewpoint height in the HMD is determined through the tracking technology and the location of the user’s head.

HTC Vive HMD

This experiment used the HTC Vive HMD. The Vive represents a high-end commercial VR HMD, overcoming limitations of previous empirical studies including improved resolution, refresh rate, FOV, usability, and adjustability.

The specifications of the Vive include a 2160 x 1200 (1080 x 1200 per eye) resolution at a 90Hz refresh rate with an 110° field of view (FOV). The vive provides the ability to walk within a 15’ x 15’ space, with complete control over head orientation. The Vive is non-invasive and is features two forms of customizability. The first is the ability to adjust the tightness of the headset with three adjustable straps, located at the top and each side of the head. This adaptable hardware allows the user to adjust the overall fit of the headset. The second form of customizability is the ability to adjust the interpupillary distance (IPD). Users can adjust the IPD through a knob located on the side of the headset.

LG Screen Display

The LG 65UF9500 was used as the screen display and contains the standard features for a high-end screen display including an Ultra-HD 4K (3840 x 2160) resolution with a 65” display. The FOV of a TV depends on the TV width, distance away from the viewer, and TV ratio. To match the 110° of the HTV, the participant sat 1.7 feet away from the TV display.

Hypotheses

Since interdisciplinary team members use prototyping as a design and communication tool, it is critical to understand how different displays affect user experience (UX) and aesthetic preferences. Additionally, UX can be examined as a process variable between display medium and perceived aesthetic value. The use of two hotel environments, including a luxury and economy condition, provides the opportunity to analyze the consistency of display medium across variables. Additionally, to understand the consistency of display on variables across disciplines, the interaction between display medium and background, i.e. designer vs. non-designer, is analyzed.

The following hypotheses guided the analyses:

Hypothesis 1: A virtual environment viewed on a HMD will provide a higher sense of spatial presence and affective reactions compared to the same virtual environment viewed on a screen display

Hypothesis 2: A virtual environment viewed on a HMD will provide superior perceived aesthetic value compared to the same virtual environment viewed on a screen display.

Hypothesis 3: Process variables, spatial presence, and affective reactions, will also have an impact on perceived aesthetic value.

Hypothesis 4: The effect of display type on process and outcome variables will be consistent across environmental conditions.

Hypothesis 5: The effect of display type on process and outcome variables will be consistent across different majors (i.e. Designer vs. Non-Designer)

Statistical Methods

Data Screening

Before analysis, scales were examined using JMP for the accuracy of data entry, missing values, reversed values, the fit between their distributions, and the assumptions of multivariate analysis. 80 subjects participated in the study. Because participants completed the post-environment questionnaire after touring two virtual hotel environment, spatial presence, affective reactions, and aesthetic value were recorded twice, totaling 160 collections. Using JMP, the data was scanned for multivariate outliers by looking for values that exceed the Mahalanobis critical value of 20.515 (chi-square, $p<.0001$). Multivariate outliers were not detected, and all cases remained for analysis. Table 3-1 presents characteristics of this sample.

Table 3-1
Sample Characteristics

Demographic Variables			N=80
Gender	*Female*		40
	Male		40
Age	*Under 21*		48
	21~25		26
	26~30		4
	31~35		2
Major	*Architecture or Design***		
		Female	20
		Male	20
	Non-Architecure or Design		
		Female	20
		Male	20

***Included architecture, interior design, industrial design, and landscape architecture majors*

Scale Refinement- Reliability and Validity

Reliability refers to the stability of a measure over time (test-retest) and the internal consistency of measures (Nunnally 1978). Internal consistency of the scales was assessed using JMP to calculate the Cronbach's Alpha of each sub-scale. Internal consistency reliabilities vary from a minimum of 0 to a high of 1.0. These scores represent the proportion of the variance in the respondent's scores attributed to true differences on the scale (DeVellis 1991). DeVellis recommends an alpha below .60 as unacceptable; .60-.65 undesirable; .65-.70 minimally acceptable; .70-.80 respectable; .80-.90 very good; and if much above .90 excellent and potentially provides an opportunity for the researcher to shorten the scale.

To refine the original scales (See Appendix B), individual-to-total reliability was examined for each sub-scale. If any individual question reduced the total reliability (Cronbach Alpha) substantially, that question was removed from the scale. Five subscales, including two process and three outcome variables, were revised. Eight questions were deleted from the administered survey. A summary of the omitted questions is in Table 3-2 & Table 3-3. By averaging sub-scales together, final scales were established, and re-assessed for overall scale reliability with no additional questions removed.

Table 3-2 shows the results of the reliability tests performed for all the scales, with multiple sub-scales, used in this study. The reliability coefficient (alpha) for 16 variables exceeded the respectable level of .70 (Cronbach 1951; Nunally 1978). The remaining two variables, dominance, and coherence were in the minimally acceptable level. Although these scores were lower than the rest, they are still considered acceptable (DeVellis 1991; Moss et al. 1998). Besides the sub-scale mystery, all scales used in this study were internally consistent and reliable measures of the associated variables.

Table 3-2

Scale Refinement

Scale and Source	Sub- Scale	Number of Cases	Original Number of Items	Number of Items Kept	Questions Omitted	Cronbach Alpha
Spatial Presence	Attention	160	6	6	NA	0.88
MEC-SPQ	Spatial Situation Model	160	6	6	NA	0.85
	Self Location	160	6	6	NA	0.95
	Possible Actions	160	6	6	NA	0.93
	Higher Cognitive Invovement	160	6	6	NA	0.77
	Suspension of Disbelief	160	6	5	4	0.87
	Visual Spatial Imagery	160	6	6	NA	0.75
	Domain Specififc Interest	160	6	6	NA	0.92
Affective Reactions	Pleasure	160	6	6	NA	0.9
Semantic Differential	Arousal	160	6	6	NA	0.78
	Dominance	160	6	5	4	0.69
Aesthetic Value	Complexity	160	3	3	NA	0.79
Visual Properties	Legibility	160	5	5	NA	0.79
Questionnaire	Coherence	160	3	3	NA	0.68
	Mystery	160	3	2	13	0.55
	Openness	160	2	2	NA	0.73
	Typicallity	160	1	1	NA	NA
	Brightness	160	1	1	NA	NA
	Uniform Lighting	160	1	1	NA	NA
	Familiarity	160	1	1	NA	NA
	Materiality	160	2	2	NA	0.83

Table 3-3

Scale Refinement- Omitted Questions

Sub-Scale	Question Number	Question
Object Imagery	11	In school, I had no problems with geometry.
Object Imagery	12	I find it difficult to imagine how a three-dimensional geometric figure would exactly look like when rotated.
Object Imagery	14	I can easily imagine and mentally rotate three-dimensional geometric figures.
Spatial Imagery	4	I can close my eyes and easily picture a scene that I have experienced.
Spatial Imagery	7	Sometimes my images are so vivid and persistent that it is difficult to ignore them.
Spatial Imagery	11	I enjoy picture with bright colors and unusual shapes like the ones in modern art.
Suspension of Disbelief	4	I took a critical viewpoint of the virtual hotel room.
Dominance	4	Awed - Important (Word Pairing)
Mystery	13	To what degree does the environment hide either positive or negative encounters that may lie ahead?

While reliability focuses on the stability and consistency of a measure, validity focuses on how well a scale measures some theoretical construct. Since a scale can be reliable but not valid, i.e. measuring some other theoretical construct, the validity of the scale was measured.

The initial scale development process considered content validity, which refers to the representativeness or sampling adequacy of the content. Questions in the instrument were based on previous studies exploring similar subject areas, refining content validity (Hanyu, 1997; Lang, 1980; Russel and Mehrabian, 1974; Vorderer et al., 2004).

Construct validity refers to the extent to which the scale measures a theoretical construct. It also addresses the issue of whether the scale demonstrates expected patterns between subscales of the same construct. Spatial presence, affective reactions, and the aesthetic value contained multiple sub-scales. The sub-scales of spatial presence included attention, spatial situation model, self-location, possible actions, higher cognitive involvement, suspension of disbelief, visual-spatial imagery, and domain specific interest. Correlation matrix and Cronbach alpha were used to understand whether these sub-scales should remain unidimensional or averaged together to represent a single scale.

Table 3-4

Intercorrelation Matrix for spatial presence sub-scale

	Attention	*SSM*	*Self Location*	*Possible Actions*	*Higher Cognitive Involvement*	*Suspension of Disbelief*	*Visual Spatial Imagery*
Attention	1						
SSM	0.67***	1					
Self Location	0.74***	0.56***	1				
Possible Actions	0.69***	0.55***	0.87***	1			
Higher Cognitive Involvement	0.53***	0.53***	0.57***	0.59***	1		
Suspension of Disbelief	0.03	0.03	0	-0.05	-0.35***	1	
Visual Spatial Imagery	0.39***	0.61***	0.35***	0.31***	0.47***	-0.03	1
Domain Specific Interest	0.28***	0.28***	0.4***	0.44***	0.31***	-0.02	0.23***

****p<.001*

Correlation analysis for the eight sub-scales of spatial analysis was constructed. As seen in Table 3-4, correlations for all subscales, except suspension of disbelief, were highly correlated. Because of this, the seven highly correlated subscales were averaged to represent spatial presence, with the grand average of 48 questions. Suspension of disbelief remained unidimensional. In addition to spatial presence, the sub-scales of affective reactions were analyzed using a correlation matrix as seen in Table 3-5.

Table 3-5
Intercorrelation Matrix for affective reactions sub-variable

	Pleasure	*Arousal*	*Dominance*	*Pleasure-SAM*	*Arousal-SAM*
Pleasure	1				
Arousal	0.49***	1			
Dominance	0.26***	0.32***	1		
Pleasure- SAM	0.6***	0.51***	0.16**	1	
Arousal- SAM	0.14	0.3***	0.15	0.14	1
Dominance-SAM	0.08	0.14	0.04	0.15	0.18**

****p<.001, **p<.05*

Affective Reactions has three sub-scales including pleasure, arousal, and dominance. These subscales were measured using two scales, one visual and one verbal. The semantic differential includes six bipolar Likert word pairings per subscale (18 pairings in total; Russel and Mehrabian, 1974). The second scale, the self-assessment manikin (SAM), uses a visual scale of emotions also categorized by pleasure, arousal, and dominance. (Lang, 1980).

As seen in Table 3-5, pleasure and arousal are correlated across the verbal and visual scales. Because of this, a pleasure and arousal score was developed by adding the verbal and visual scores together. Dominance remained unidimensional, only using the semantic differential scale, discarding the SAM Dominance Scale. Besides spatial presence and affective reactions, the sub-scales of aesthetic value were analyzed using a correlation matrix as seen in Table 3-6.

Table 3-6

Intercorrelation Matrix for affective reactions sub-scales

	Complexity	*Legibility*	*Coherence*	*Mystery*	*Openness*	*Typicallity*	*Brightness*	*Uniform Lighting*	*Familiarity*
Complexity	1								
Legibility	0.12	1							
Coherence	0.53***	0.51***	1						
Mystery	0.51***	0.34***	0.42***	1					
Openness	0.3***	0.51***	0.56***	0.49***	1				
Typicallity	0.37***	0.39***	0.56***	0.34***	0.34***	1			
Brightness	0.31***	0.47***	0.52***	0.34***	0.36***	0.35***	1		
Uniform Lighting	0.31***	0.38***	0.43***	0.21***	0.36***	0.29***	0.5***	1	
Familiarity	0.24***	0.64***	0.53***	0.34***	0.54***	0.48***	0.35***	0.39***	1
Materiality	0.46***	0.29***	0.6***	0.3***	0.33***	0.52***	0.41***	0.37***	0.42***

****p<.001*

Aesthetic Value has ten sub-scales including complexity, legibility, coherence, mystery, openness, typicality, brightness, uniform lighting, familiarity, and materiality. The questions for all subscales, except materiality, were taken from the Visual Environment Properties Questionnaire (Hanyu, 1997). Materiality questions were self-developed through a focus group with ten undergraduate design students.

As seen in Table 3-6, all sub-scales of aesthetic value were significantly correlated. Because of this, an overall aesthetic value score was developed by averaging 21 questions together.

Distribution of Variables

The distribution of all process and outcome scale residuals were analyzed by their goodness of fit to a normal distribution. Since this study used Likert scales to measure process and outcome variables, a normal distribution was required to perform a paired t-test and linear analysis. If the variable were not normally distributed, a non-parametric analysis would need to be conducted.

Table 3-7
Distribution of Scales

Scale	HMD, n=80	Screen Display, n=80	Goodness of Fit, n=160
	M (SD)	M (SD)	Prob<W
Spatial Presence	5.46	4.27	0.51
Suspension of Disbelief	4.45	4.55	0.08
Asthetic Value	3.84	3.5	0.45
Pleasure	10.71	8.97	0.39
Arousal	7.72	6.37	.029*
Dominance	4.96	4.81	0.053

***p<.001, *p<.05*

The goodness of fit is measured with a Shapiro-Wilk Test. If the probability of W is significant, this indicates that the variable is not normally distributed. As seen in Table 3-7, all scales except arousal are normally distributed, indicating that these variables can be examined using paired t-test and linear analysis.

Environment Manipulation Check

Since the hotel environments created for this experiment represent a luxury and economy environment, a manipulation check was conducted to ensure that the two environments represent these conditions. First, hotel environment vs. aesthetic value was analyzed. From this analysis, the luxury environment had an average score of 3.86, while the economy room had a 3.49. A simple linear model of environment type on the aesthetic value indicates that the luxury room was perceived with higher quality, validating that the luxury environment had a significantly higher perceived aesthetic value (p=0.001; Figure 3-1).

To further address environment representativeness between the two rooms, participants were asked to rate how much they would pay per night for the virtual hotel room. Participants rated the luxury environment with an average score of 6.4 ($275) and the economy room with a 4.19 ($168; Figure 3-2). The higher perceived economic and aesthetic value validates that the luxury environment represented a luxury hotel environment, while the economy environment represented an average hotel environment.

Figure 3-1

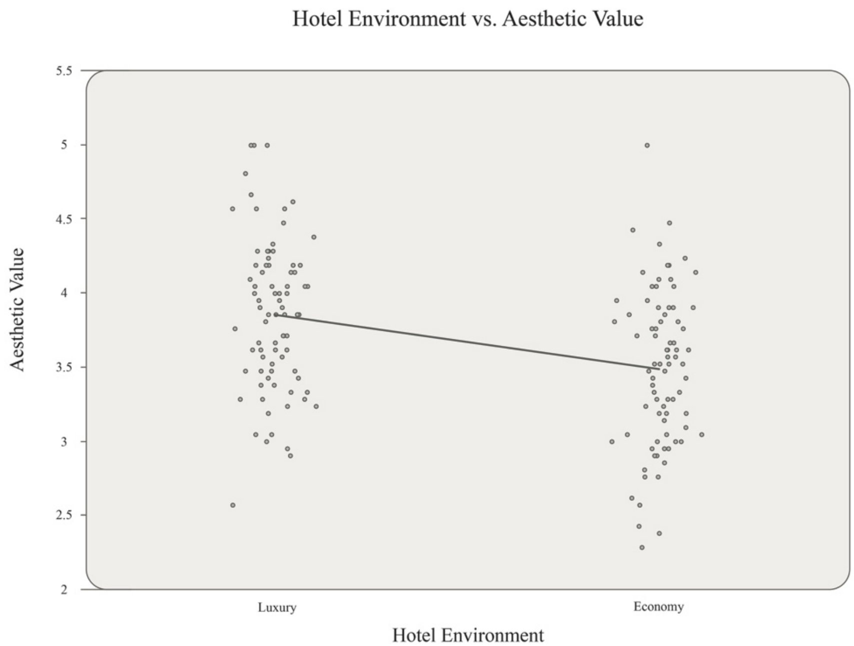
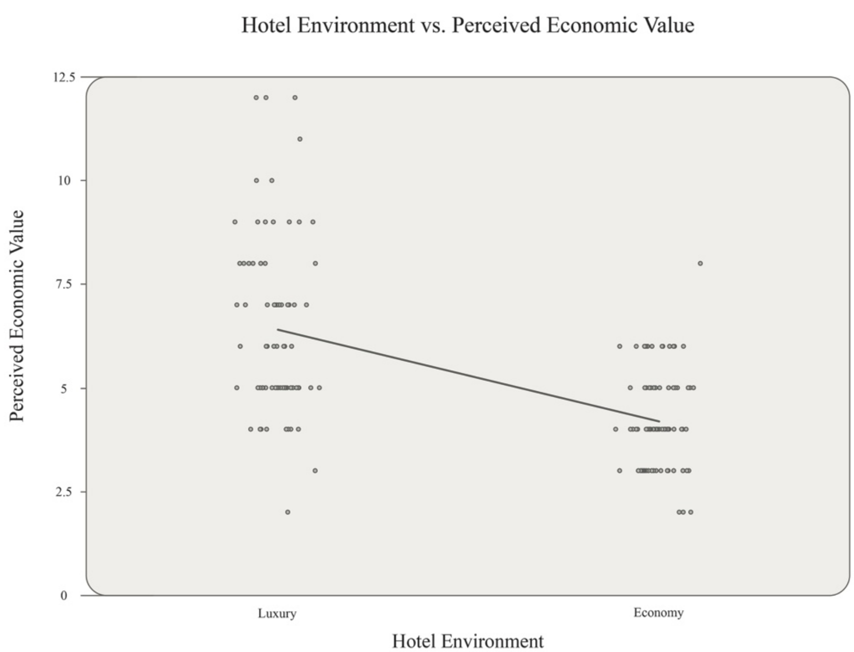


Figure 3-2



Statistical Analysis

Overview of Analysis for Hypotheses Tests

The following hypotheses guided the statistical analysis.

> *Hypothesis 1: A virtual environment viewed on a HMD will provide a higher sense of spatial presence and affective reactions compared to the same virtual environment viewed on a screen display*
>
> *Hypothesis 2: A virtual environment viewed on a HMD will provide superior perceived aesthetic value compared to the same virtual environment viewed on a screen display.*
>
> *Hypothesis 3: Process variables, spatial presence, and affective reactions, will also have an impact on perceived aesthetic value.*
>
> *Hypothesis 4: The effect of display type on process and outcome variables will be consistent across environmental conditions.*
>
> *Hypothesis 5: The effect of display type on process and outcome variables will be consistent across different majors (i.e. Designer vs. Non-Designer)*

This study documents the use of a stereoscopic HMD potentially distinctive from a monoscopic screen display regarding perceptual outcomes and UX processes. Perceptual outcomes were measured by perceived aesthetic value, while UX processes were measured by spatial presence and affective reactions.

To address hypotheses 1 and 2, the main effects of display were examined across all process and outcome variables, using a paired t-test. To address hypothesis 3, the impact of multiple variables on perceived aesthetic value was analyzed through linear modeling. To address hypothesis 4, potential differences across the two environmental conditions, luxury vs. economy, were examined through the interaction of display and environment on process and outcome variables. Finally, to address hypothesis 5, potential differences across design background, designer vs. non-designer, were examined through the interaction of display and background on process and outcome variables.

Hypotheses 1 and 2

Hypothesis 1: A virtual environment viewed on a HMD will provide a higher sense of spatial presence and affective reactions compared to the same virtual environment viewed on a screen display

Hypothesis 2: A virtual environment viewed on a HMD will provide superior perceived aesthetic value compared to the same virtual environment viewed on a screen display.

Table 4-1
Paired T-Test of Display Medium on Process and Outcome Variables

Scale	HMD, n=80	Screen Display, n=80	Goodness of Fit, n=160	Test Statistic	
	M (SD)	M (SD)	Prob<W	T or Z	p value
Spatial Presence	5.46	4.27	0.51	11.08	.0001**
Suspension of Disbelief	4.45	4.55	0.08	-0.51	0.61
Asthetic Value	3.84	3.5	0.45	4.09	.0001**
Pleasure	10.71	8.97	0.39	7.31	.0001**
Arousal	7.72	6.37	.029*	4.81	.0001**
Dominance	4.96	4.81	0.053	0.81	0.42

***p<.001, *p<.05*

The impact of the display medium, HMD vs. screen display, was analyzed for all process and outcome variables using a paired t-test (when the scale was normally distributed according to Shapiro-Wilks Test), or a non-parametric Wilcoxon/ Kruskal-Wallis Test (when not normally distributed; Table 4-1). Since arousal was not normally distributed, a non-parametric Wilcoxon/ Kruskal-Wallis Test was used to analyze the effect of display medium on arousal, producing a Z instead of T statistic.

The scales with a significant p-value indicate that display medium had a significant impact on the scale. This analysis suggests that differences in the display medium significantly affected spatial presence, aesthetic value, pleasure, and arousal.

In this model, the four assumptions of linear modeling were met. Since the independent variable (display medium) was categorical, it met the assumption that the model was linear. Because of the experimental design, each participant answered the questionnaire twice, once after each hotel room. This was done because of practicality reasons, so each participant could experience both display devices while in the lab. Because of this, observations were not independent, i.e. each participant took the same questionnaire twice. Although the observations were not independent, this can be addressed by adding participant number as a random predictor variable. A low percentage of variance in the data across all models was due to participant ID, producing a value lower than .01. The residual of each model was analyzed. To analyze homoscedasticity, each residual was plotted against the predicted value. All of the paired t-tests met the requirement of homoscedasticity.

Besides paired t-tests and the Wilcoxon/ Kruskal-Wallis Test, linear modeling was used to further analyze the relationship between display medium and process and dependent variables. While a t-test allows us to see if differences in a particular variable between groups are due to chance, or if the groups indicate differences based on individual differences, linear modeling allows us to test a pattern of variables to predict the influence of a treatment.

For this study, a dichotomous dummy variable represented the different display mediums (HMD vs. Screen Display). HMD technology positively influenced process and outcome variables. Between the 80 subjects who experienced the HMD and the 80 subjects who experienced the screen display, all process and outcome variables were analyzed.

As seen in Figures 4-1 through 4-6, HMD has a positive influence on process variables: spatial presence, pleasure, and arousal. Additionally, as seen in Figures 4-7 through 4-8, HMD had a positive influence on outcome variable: aesthetic value. The R^2 of these variables ranged from a high of .44 (spatial presence) to a low of .1 (aesthetic value; Table 4-2). Figures 4-1 through 4-8 visually illustrate the participant's spatial presence, aesthetic value, affective reactions, and relative subscale scores, across both the HMD and screen display conditions. The line connects the mean of each score.

Table 4-2

Simple Linear Model- Predictor Variable: Display Medium

	Predictor Variable: Display Type				
Outcome Variables	b	β	t	p	R^2
Spatial Presence (Display --> Spatial Presence)	4.87	0.6	11.08	.0001**	0.44
Suspension of Disbelief (SoD) (Display --> SoD)	4.51	-0.0525	0.51	0.61	0.0017
Aesthetic Value (Display --> Asethetic Value)	3.67	0.17	4.09	0.001**	0.1
Pleasure (Display --> Pleasure)	9.84	0.87	7.31	.0001**	0.25
Arousal (Display --> Arousal)	7.04	0.68	4.79	.0001**	0.13
Dominance (Display --> Dominance)	4.89	0.07	0.81	0.42	0.005

****p<.001, **p<.001, *p<.05**

Figure 4-1

Semantic Differential by Display Medium

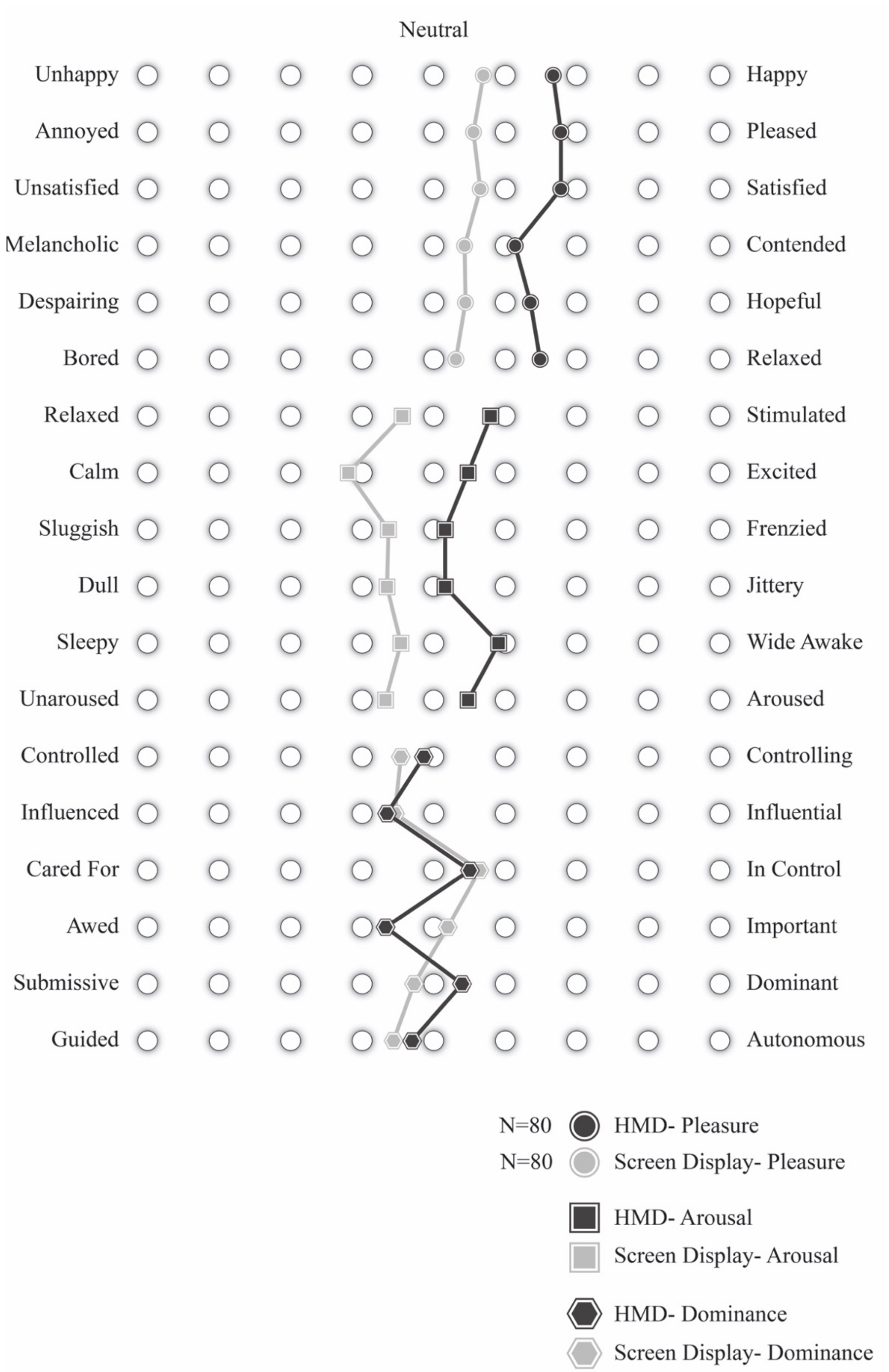


Figure 4-2

Self-Assessment Manikin by Display Medium

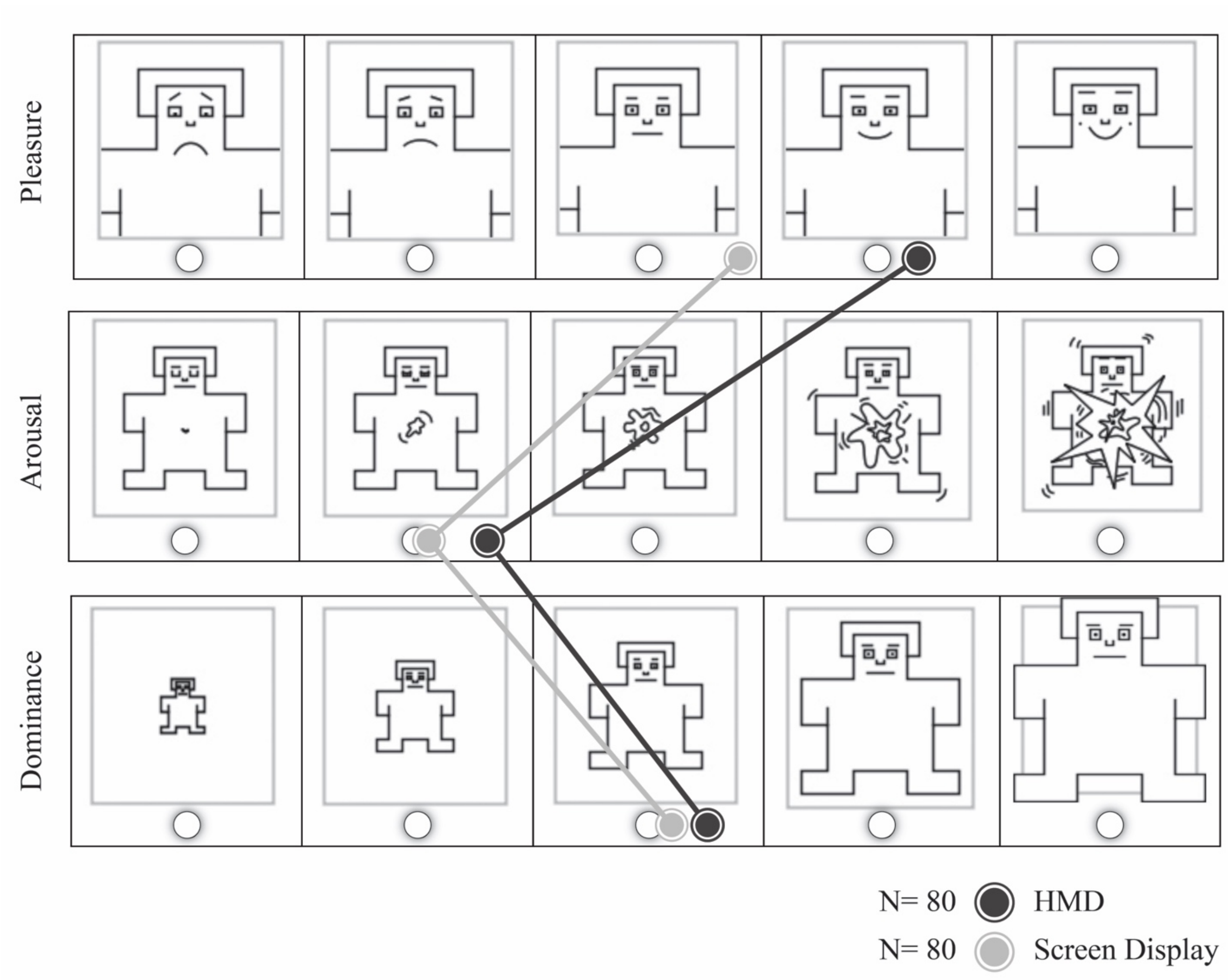


Figure 4-3

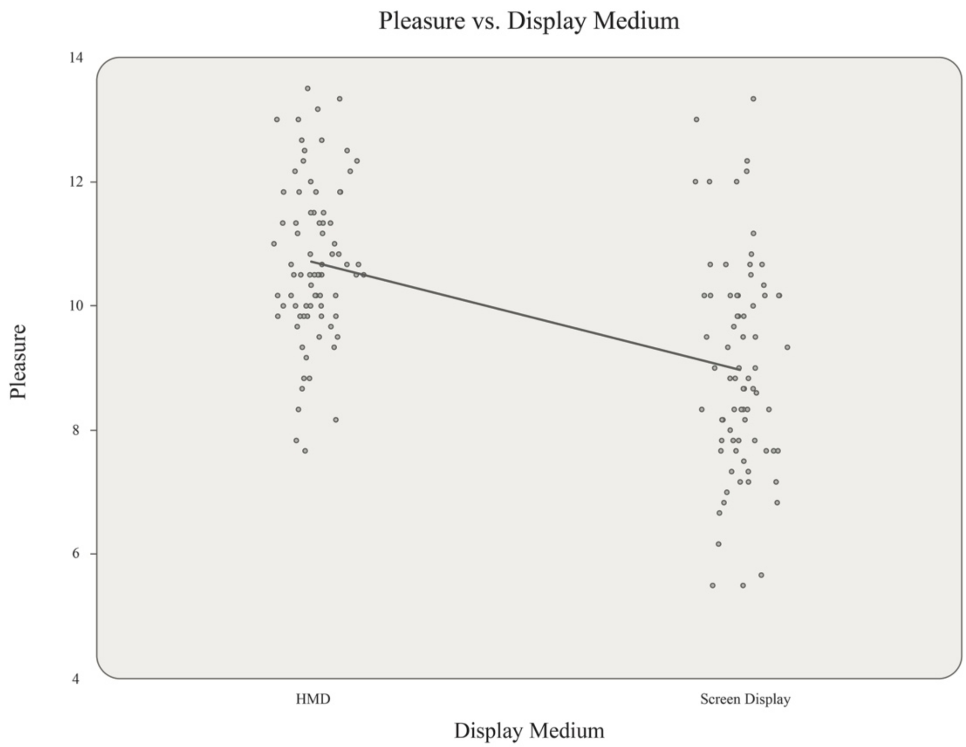


Figure 4-4

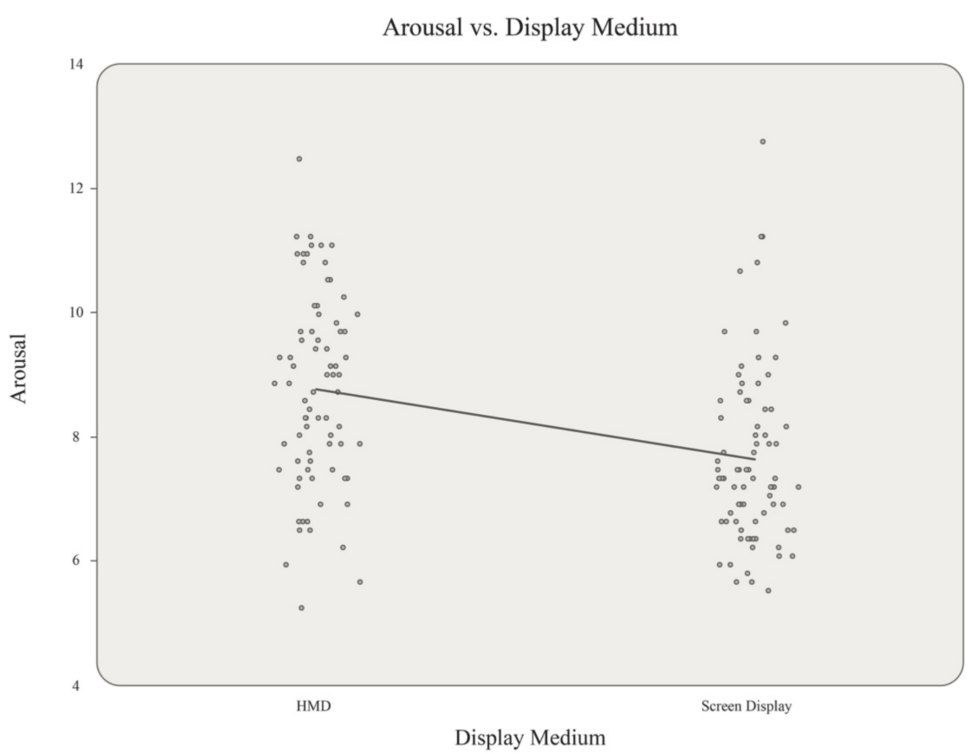


Figure 4-5

Spatial Presence- Subscale Score by Display Medium

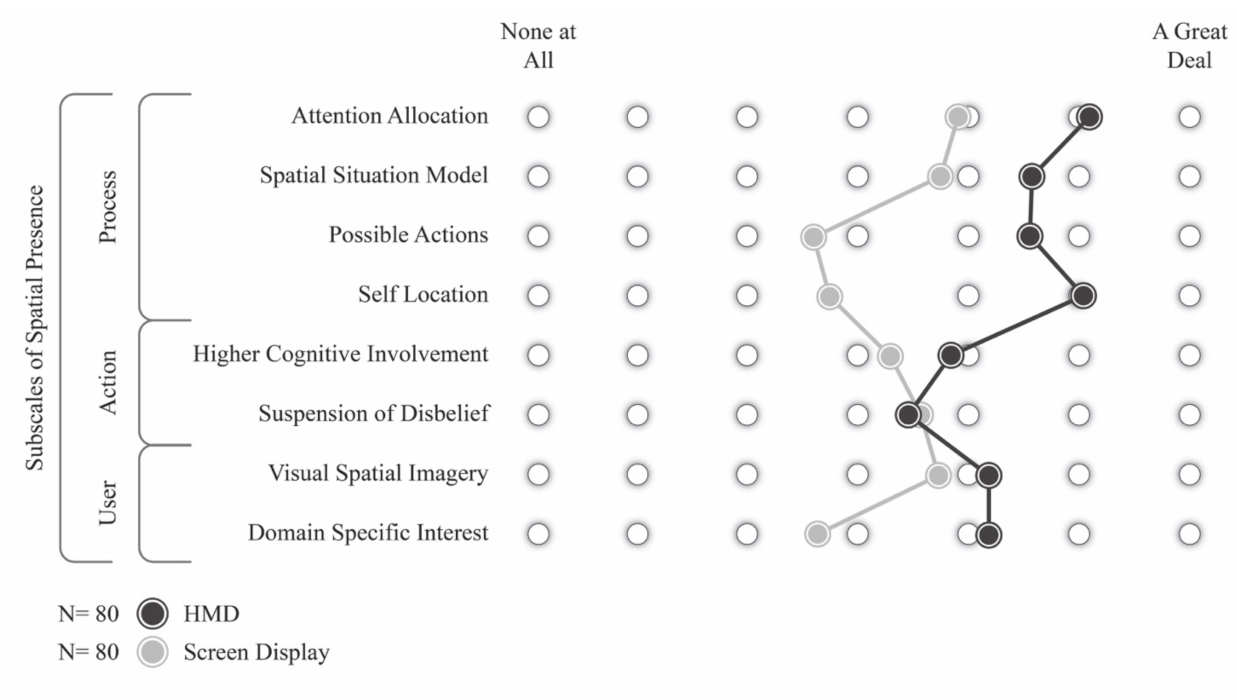


Figure 4-6

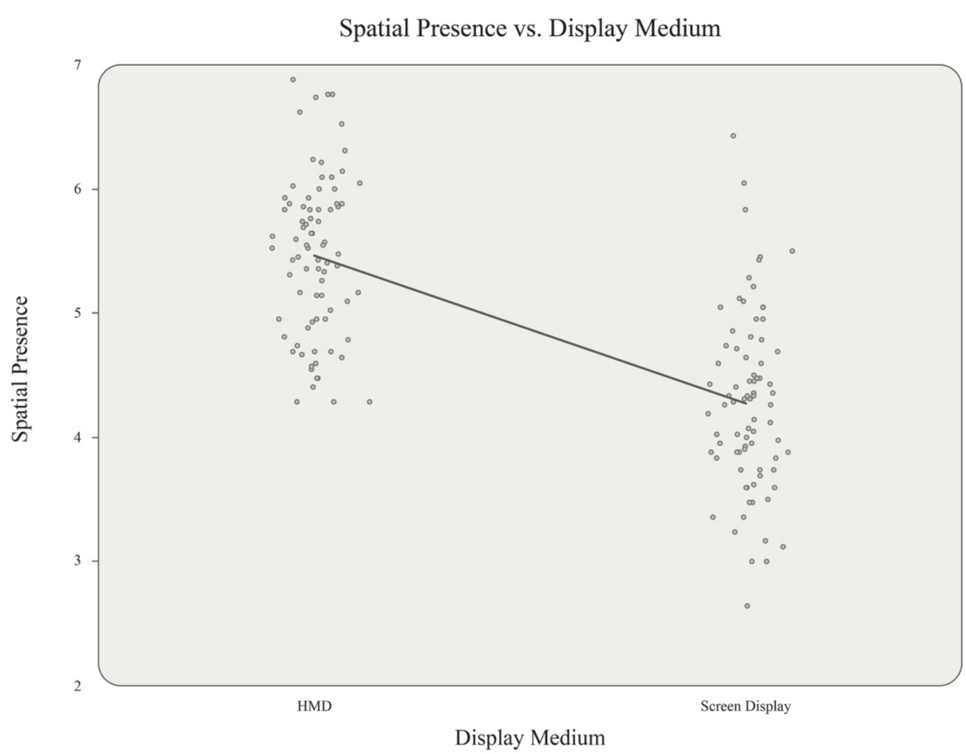


Figure 4-7

Perceived Aesthetic Value- Subscale Score by Display Medium

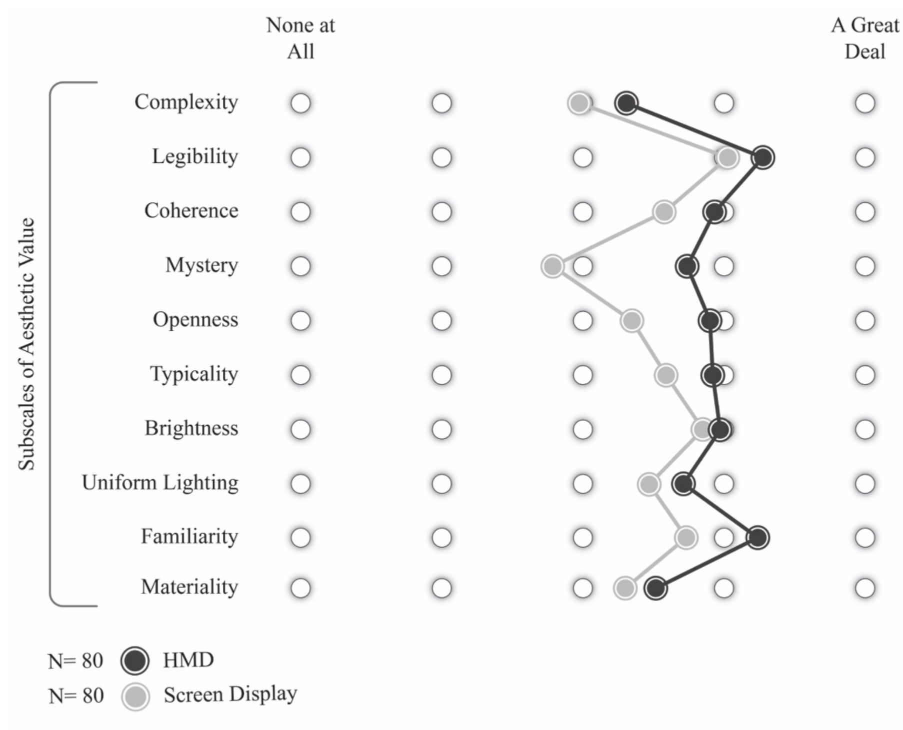
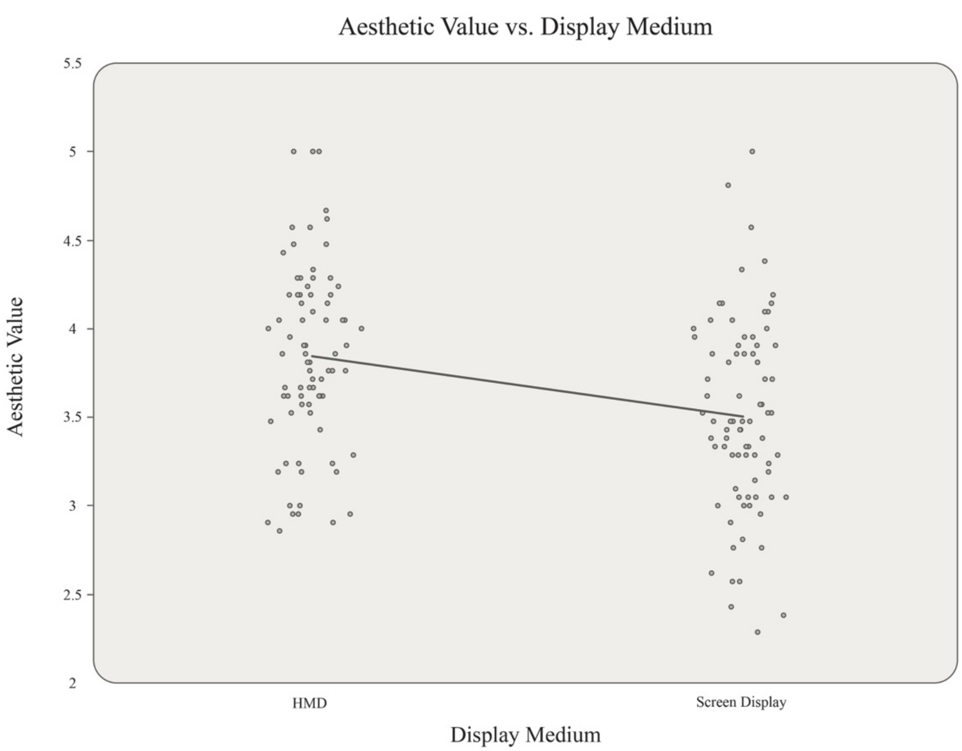


Figure 4-8



Hypothesis 3

Hypothesis 3: Process variables, spatial presence, and affective reactions, will also have an impact on perceived aesthetic value.

After the initial linear models had been analyzed for hypotheses 1 and 2, the impact of multiple scales on aesthetic value was analyzed through new linear models. Based on our original assumptions, there should be an impact of both display medium, and process variables (i.e. spatial presence and affective reactions) on perceived aesthetic value. To address hypothesis 3, a linear model with aesthetic value as the outcome variable, and the predictor variables of display medium (two-level category), pleasure and spatial presence (continuous scales), random effect of person ID, and display order to check if people tended to like the first thing shown to them.

Table 4-3
Linear Modeling- Outcome Variable: Aesthetic Value

Outcome Variable: Aesthetic Value				
Overall R^2=.46				
Predictor Variables	Estimate	Standard Error	T-Ratio	P
Spatial Presence	0.3	0.06	5.26	0.0001***
Pleasure	0.11	0.03	4.32	0.0001***
Arousal	0.0096	0.02	0.46	0.64
Display Medium	-0.12	0.04	-2.68	.0082**
Display Medium Order	-0.004	0.03	-0.11	0.91

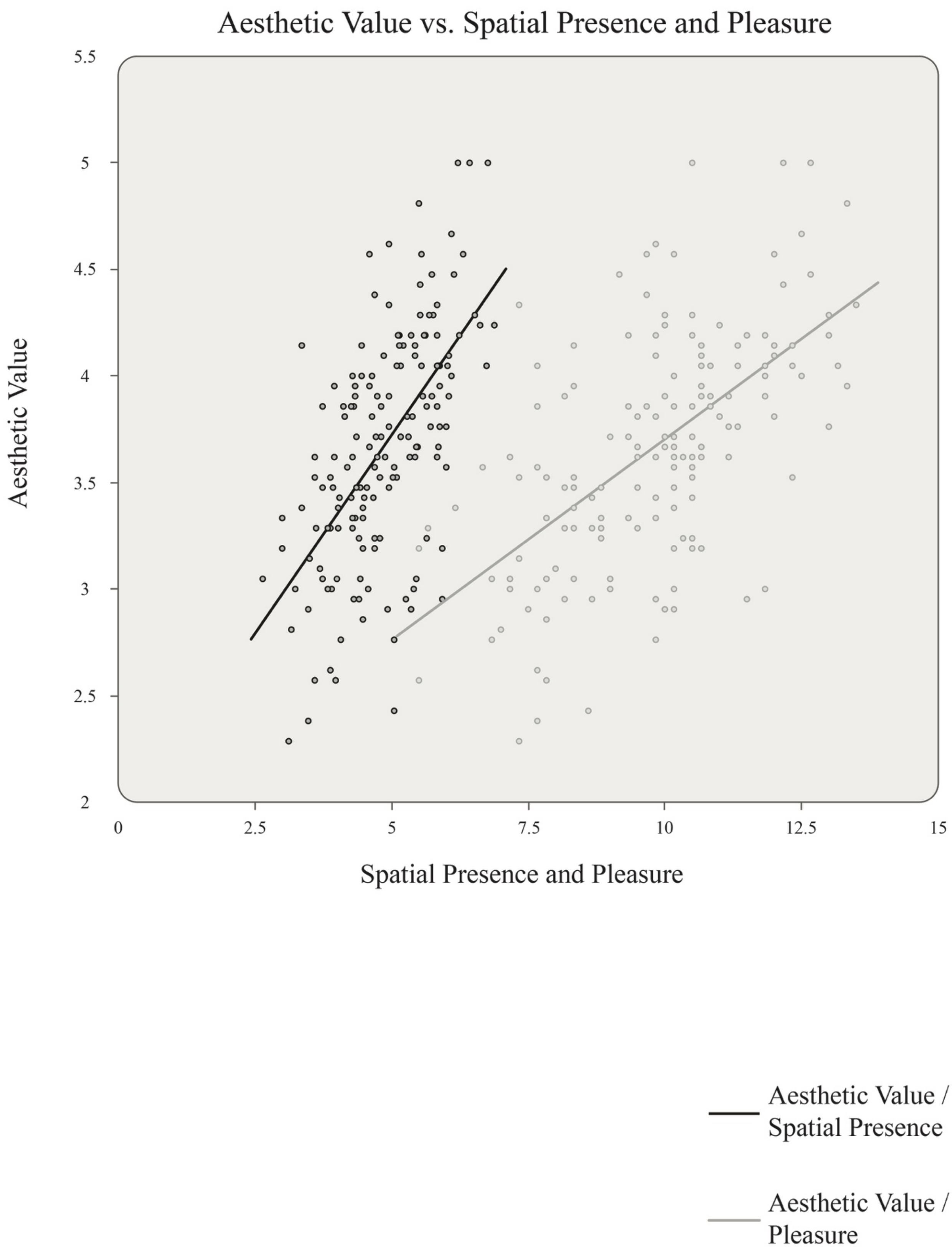
****p<.001, **p<.01, *p<.05*

Table 4-3 shows the impact of multiple variables (i.e. display, pleasure, spatial presence, and display order) on aesthetic value. In this model, spatial presence, pleasure, and display medium had a significance level under .01. This suggests that these predictor variables had a significant impact on aesthetic value. The only non-significant scale was arousal and the order of the display medium. Order of display medium was included to make sure that the order had no impact on the outcome of the data. The overall R^2 was .46, indicating that approximately 46% of the change in aesthetic value can be explained by the predictor variables. These findings validate Hypotheses 3 that the process variables, i.e. affective reactions, and spatial presence have a significant impact on perceived aesthetic value.

Additionally, all assumptions for linear modeling were met in this model. Since the independent variable (display medium) was categorical, it met the assumption that the model was linear. The continuous predictors, i.e. spatial presence, and pleasure, were plotted against aesthetic value and produced a linear relationship (Figure 4-9).

Like the paired t-test analysis, random effect of participant number was used to address the non-independence of the research design. A very low percentage (0.001) of the variance in the model was due to participant ID. Additionally, the residual was plotted against the predicted value to address homoscedasticity. The distribution of the residual also proved to be normal (Prob $|w|<.27$). Therefore, this model met all assumptions for linear modeling. Figure 4-9 visually illustrates spatial presence and pleasure against aesthetic value. As both spatial presence and pleasure increases, so does aesthetic value.

Figure 4-9


Aesthetic Value vs. Spatial Presence and Pleasure
Aesthetic Value
5.5
5
4.5
4
3.5
3
2.5
2
0
2.5
5
7.5
10
12.5
15
Spatial Presence and Pleasure
Aesthetic Value / Spatial Presence
Aesthetic Value / Pleasure

Hypothesis 4

Hypothesis 4: The effect of display type on process and outcome variables will be consistent across environmental conditions.

During the experiment, each participant experienced a virtual hotel room on both the HMD and screen display and used a self-report questionnaire to measure the aesthetic value of both virtual hotel environments. Because the participant measured aesthetics, showing the same environment twice could produce carryover effects. Because of this, two environments were needed to test the perception of aesthetic value from the same participant in both the HMD and screen display. Because two environments were needed, one was designed to be a high-end luxury room while the other was an average economy room. By having two different environments, we could then analyze whether the impact and effect size of the process and outcome variables were consistent across both environmental conditions.

To test this hypothesis, two main analyses were conducted. First, multiple t-tests were conducted examining the effect of display medium on process and outcome variables by the environment (Economy vs. luxury).

Table 4-4

T-Test of Display Medium on Process and Outcome Variables by Environment

Luxury	*HMD, n=40*	*Screen Display, n=40*		
Variable	*M (SD)*	*M (SD)*	*t-test (t -ratio)*	*p value*
Spatial Presence	5.60 (61)	4.36 (.82)	7.71	.0001***
Suspension of Disbelief	4.39 (1.25)	4.51 (1.23)	-0.45	0.65
Asthetic Value	4.02 (.48)	3.68 (.49)	3.07	.0001***
Pleasure	11.04 (1.17)	9.56 (1.75)	4.46	.0001***
Arousal	7.86 (1.79)	6.61 (2.0)	2.91	.0047**
Dominance	4.89 (1.19)	4.99 (1.26)	-0.38	0.7
Economy	*HMD, n=40*	*Screen Display, n=40*		
Variable	*M (SD)*	*M (SD)*	*t-test (t -ratio)*	*p value*
Spatial Presence	5.33 (.65)	4.18 (.60)	8.16	.0001***
Suspension of Disbelief	4.52 (1.46)	4.60 (1.27)	-0.28	0.78
Asthetic Value	3.66 (.48)	3.31 (.53)	3.06	0.003**
Pleasure	10.38 (1.31)	8.38 (1.46)	6.44	.0001***
Arousal	7.58 (1.92)	6.12 (1.33)	3.96	.0002***
Dominance	5.04 (.93)	4.65 (1.10)	1.74	0.087

****p<.001, **p<.01, *p<.05*

As seen in Table 4-4, display medium impacted the same variables in both environments. This indicates that the impact of display medium on process and outcome variables was consistent across both environments. Although this suggests that display medium had a bearing on the same variables, it does not validate whether the effect size was consistent across both environments.

To test whether the effect size of on the different process and outcome variables was consistent across both environmental conditions, we used a linear model including a random effect of person ID, main effects of display and environment, and an interaction term for display medium and environment. By examining the interaction effect, we could then analyze whether the effect size of each variable was the same across the different environmental conditions.

Table 4-5

Interaction of Display Medium and Environment on Spatial Presence

Least Squares Means	*HMD*	*Screen Display*	*Difference between Environments*
Economy	5.33	4.18	*1.15*
Luxury	5.6	4.36	*1.24*
Difference between Displays	*0.27*	*0.18*	

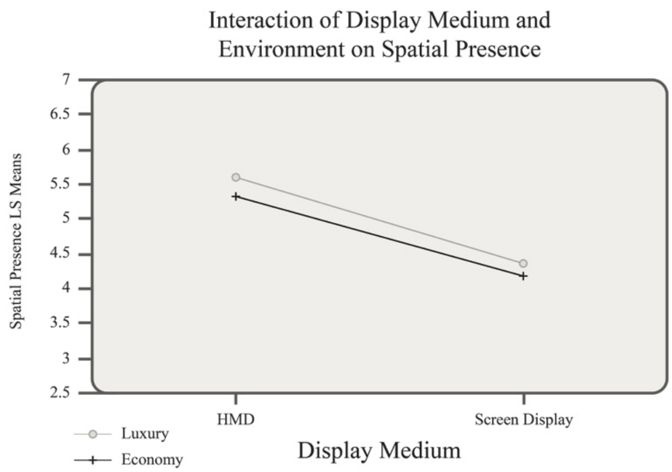


Table 4-6

Interaction of Display Medium and Environment on Aesthetic Value

Least Squares Means	*HMD*	*Screen Display*	*Difference between Environments*
Economy	3.66	3.31	*0.35*
Luxury	4.02	3.69	*0.33*
Difference between Displays	*0.36*	*0.38*	

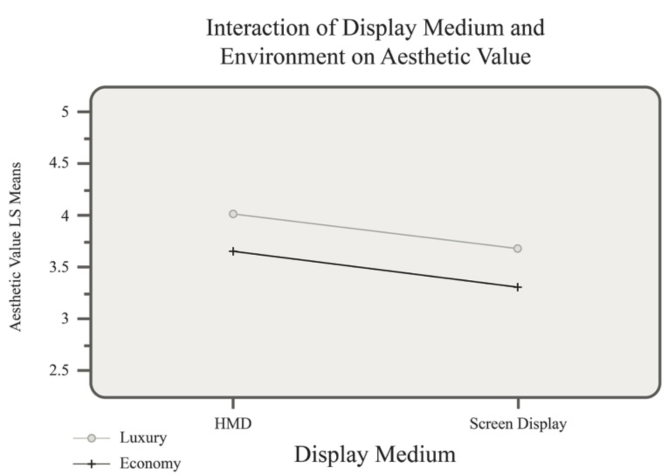


Table 4-7
Interaction of Display Medium and Environment on Pleasure

Least Squares Means	*HMD*	*Screen Display*	*Difference between Environments*
Economy	10.38	8.38	*2*
Luxury	11.04	9.56	*1.48*
Difference between Displays	*0.66*	*1.18*	

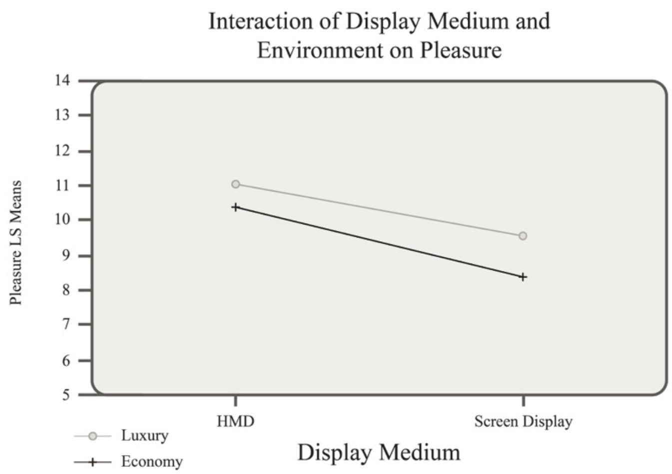
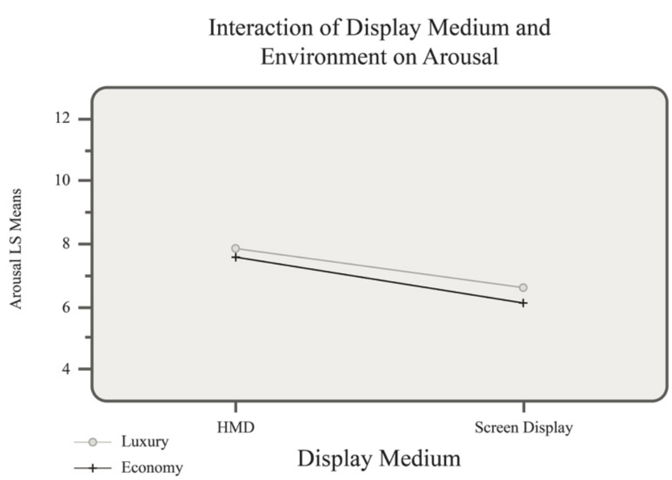


Table 4-8
Interaction of Display Medium and Environment on Arousal

Least Squares Means	*HMD*	*Screen Display*	*Difference between Environments*
Economy	7.58	6.12	*1.46*
Luxury	7.86	6.61	*1.25*
Difference between Displays	*0.28*	*0.49*	



As seen in Table 4-5 through 4-8, the effect size of the display was consistent across both the economy and luxury hotel environments for the outcome variables of aesthetic value (interaction term $p=0.94$), spatial presence (interaction term $p=0.66$), pleasure (interaction term $p=0.23$), and arousal (interaction term $p=0.70$). As seen in Table 4-7, there was a more severe increase in pleasure in the economy room vs. the luxury room. This indicates that the headset had a less dramatic effect on pleasure in the luxury hotel room. Although there was a difference, it was not substantial. Overall, this validates that effect of display type on process and outcome variables was consistent across environmental conditions.

Hypothesis 5

Hypothesis 5: The effect of display type on process and outcome variables will be consistent across different majors (i.e. Designer vs. Non-Designer)

Participants were evenly divided among design and non-design majors. To test whether the effect size of on the different process and outcome variables was consistent across designers and non-designers, we used a linear model including a random effect of person ID, main effects of display and major, and an interaction term for display and major. By examining the interaction effect, we could then analyze whether the effect size of each variable was the same across designers vs. non-designers.

Table 4-9
Interaction of Display Medium and Design Background on Spatial Presence

Least Squares Means	*HMD*	*Screen Display*	*Difference between Designer/Non-Designer*
Designer	5.46	4.25	*1.21*
Non-Designer	5.46	4.3	*1.16*
Difference between Displays	*0*	*0.05*	

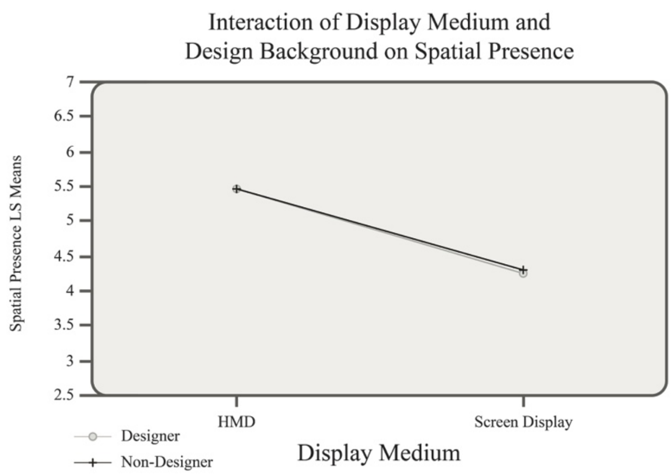


Table 4-10
Interaction of Display Medium and Design Background on Aesthetic Value

Least Squares Means	*HMD*	*Screen Display*	*Difference between Designer/Non-Designer*
Designer	3.85	3.47	*0.38*
Non-Designer	3.83	3.54	*0.29*
Difference between Displays	*-0.02*	*0.07*	

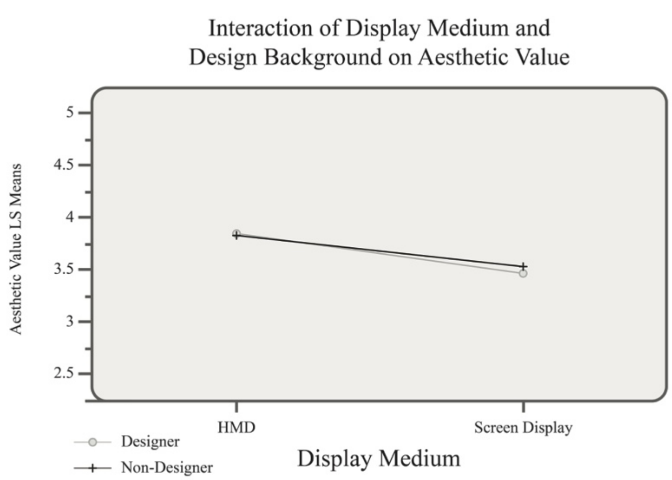


Table 4-11
Interaction of Display Medium and Design Background on Pleasure

Least Squares Means	*HMD*	*Screen Display*	*Difference between Designer/Non-Designer*
Designer	10.52	8.92	*1.6*
Non-Designer	10.9	9.02	*1.88*
Difference between Displays	*0.38*	*0.1*	

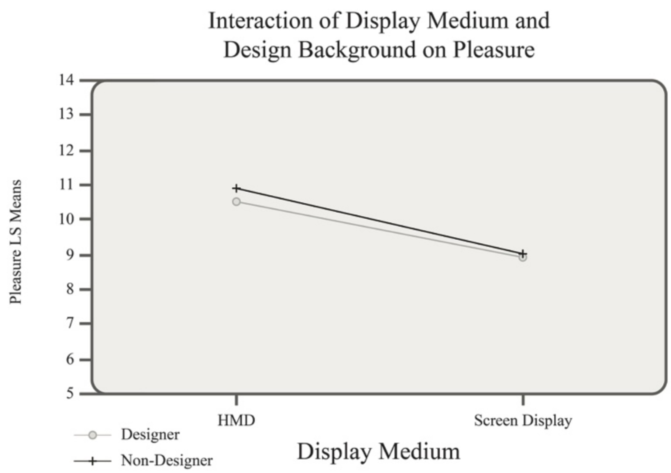
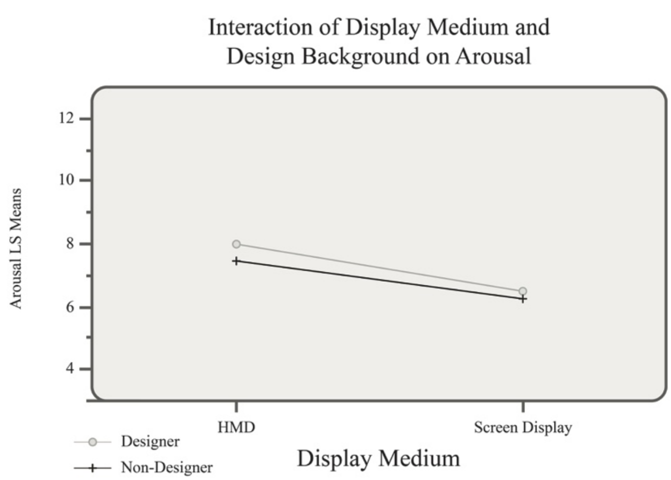


Table 4-12
Interaction of Display Medium and Design Background on Arousal

Least Squares Means	*HMD*	*Screen Display*	*Difference between Designer/Non-Designer*
Designer	7.99	6.49	*1.5*
Non-Designer	7.45	6.25	*1.2*
Difference between Displays	*-0.54*	*-0.24*	



As seen in Figure 4-9 through 4-12, design background had no main effect or interaction effect on spatial presence (interaction term p=0.82), aesthetic value (interaction term p=0.62), pleasure (interaction term p=0.56), or arousal (interaction term p=0.60). As seen in Figure 4-11, non-designers had a more dramatic increase in pleasure between the two display conditions than designers. Additionally, as seen in 4-12, designers had a more substantial increase in arousal between the two display conditions. Although pleasure and arousal were not completely consistent between designers and non-designers, the difference is marginal. Overall, this validates that effect of display type on process and outcome variables was consistent across designers and non-designers.

Results

The statistics section of this experiment includes statistical methods and hypothesis testing. Statistical methods first included screening the data for any outliers. Next, each process and outcome scale were screened for content validity or internal consistency. By looking at the item-to-total Cronbach Alpha, unreliable questions were removed from the original questionnaire. Next, each scale was screened for construct validity, or how well some questionnaire measures some theoretical construct. By analyzing correlation matrixes of different scales, some subscales remained on their own, while other subscales were averaged to represent a larger variable. For example, seven of the eight original subscales of spatial presence were averaged together, while one of the subscales remained as an independent scale.

After statistical methods had been complete, analysis of the data was conducted to validate five key hypotheses. This analysis was completed through multiple t-tests and linear modeling. The findings of the analysis suggested there was a significant impact of display on spatial presence, pleasure, arousal, and aesthetic value. Additionally, there was a significant impact of spatial presence and pleasure on aesthetic value. Linear model assumptions were addressed including the linear relationship of variables, independent observations, homoscedastic observations, and normally distributed variables. There were no significant interaction effects between the environment and display medium on the various process and outcome variables. Finally, the interaction between design background and display medium was analyzed. Findings validate that the effect of display medium on outcome and process variables was consistent across design and non-design backgrounds.

Discussion

This study verifies that display medium, i.e. HMD vs screen display, can impact the perception of virtual reality environments. As HMDs become increasingly accessible, it is critical to understand the distinct advantages that HMDs can provide over more commonly used screen displays. These differences are particularly applicable to design industries hoping to use HMDs as prototyping, communication, and visualization tools between different stakeholders.

Spatial presence, pleasure, arousal, and aesthetic value were significantly higher while using the HMD, suggesting that experience and preference for virtual environments is greater while using HMDs over screen displays. This preference validates why designers and real-estate firms should use HMDs when displaying VR environments. Additionally, the effect of display on process and outcome variables were consistent across both environment conditions. This consistency validates that designers can use HMDs from simplistic to elaborate virtual environments. If multidisciplinary teams hope to use HMDs as a visualization and communication tool, it is critical that different disciplines perceive virtual environments the same. By examining the interaction of display and discipline across all process and outcome variables, the consistency across disciplines was validated. This consistency of perception across disciplines indicates that multidisciplinary teams can use HMDs as communication, prototyping, and visualization tools.

Advances in technology will always progress visual fidelity and realism of VR experiences. This advancement in quality will exponentially increase as hardware and rendering pipelines become more streamlined toward HMD experiences. As advances in technology progress, and realism and interaction possibilities expand, it is critical to conduct additional empirical studies with updated technology.

Besides updated technology, there are few empirical studies centered on the usability of HMDs as a design tool. Exploration of studies around interaction, usability, and human factors are needed for HMDs to become a universal design tool. By addressing basic usability and perception questions through studies similar to this, designers can begin to explore integrating HMDs as a design tool, ultimately transforming the design and prototyping process.

Appendix A- Consent Form

Figure A1 includes the consent form distributed and signed by all participants before completing the experiment.

Figure A1
Consent Form

We are asking you to participate in a research study. In this form, you will find all the necessary information about the study. If you have any further questions, please ask to the researcher in charge.

Project Title: Analyzing the impact of display mediums (Virtual Reality HMD VS TV Display) within virtual hotel environments.
Principal Investigator: Ethan Harris Arnowitz; Design & Environmental Analysis
Email: eha38@cornell.edu **Phone**: (516) 965-0930

What the study is about
The purpose of this research is to study how humans perceive virtual environments on different display mediums. Particularly, we are interested in your perceptions and opinions of two virtual hotel rooms.

What we will ask you to do
In this session, you enter two virtual environments. One of these environments will be viewed on a large screen TV display, while the other will be on a virtual reality head mounted display. While in the virtual environments, you will answer a series of distance questions. A camera may record you while in the environment. This is necessary for analysis and will be destroyed after analysis is concluded. After exiting each virtual environment, you will be asked to fill out a questionnaire about your experience.

Risks and discomforts
We do not anticipated risks beyond those encountered in day-to-day life. Possible side effects may include motion sickness, nausea, or eyestrain. If you are experiencing any of these side-effects, please feel free to stop the experiment and get the attention of the researcher. If these side-effects persist after a short-time (i.e. longer than 2-hours), you should seek medical attention.

Exclusion Criteria
- No serious visual disability (Legally blind or Blind)
- Not pregnant
- No recent concussions
- No seizure disorders
- No history of fainting or seizures
- No visual impairment or condition that makes you prone to dizziness or disorientation

Benefits/Payments
There are no direct-benefits from participating in this study, but you will be compensated with a $10 Amazon gift-card at the completion of the study.

Privacy/Confidentiality
We anticipate that the videos of your participation and your results will be private and used only for the purpose of the study. Only the researcher will have access to this material. Once the study is completed all the files will be saved indefinitely on the personal hard-drive of the researcher for future reference. Consent forms will be

locked in a secure drawer within the DUET lab at Cornell University. In order to demonstrate our study procedures, we would like your permission to show photos or videos of you in publications and presentations. These images would show you wearing the headset or using the TV display, and therefore the probability of you being identified will be low.

Do you allow the researcher to take pictures and recordings of you during the study?
Yes [ ] No [ ]

Do you allow the researcher to add a picture of you in scientific publications?
Yes [ ] No [ ]

Do you allow the researcher to add pictures/video of you in media (newspapers, journals or public events), in order to illustrate the research and methodology? Yes [ ] No [ ]

Taking part is voluntary
Your participation is voluntary. You may refuse to participate before the study begins, discontinue at any time, or skip any questions/procedures that may make you feel uncomfortable.

If you have questions
The main researcher conducting this study is Ethan Harris Arnowitz, a MSc Student at Cornell University. Please ask any questions you have now. If you have questions later, you may contact Ethan Arnowitz at eha38@cornell.edu or at 516-965-0930. If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Institutional Review Board (IRB) for Human Participants at 607-255-5138 or access their website at http://www.irb.cornell.edu. You may also report your concerns or complaints anonymously through Ethics point online at www.hotline.cornell.edu or by calling toll free at 1-866-293-3077. Ethics point is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured. You will be given a copy of this form to keep for your records.

Statement of Consent
I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature________________________________ Date____________________________________

Your Name (printed)____

Signature of person obtaining consent Date____

Printed name of person obtaining consent____

This consent form will be kept and secured by the researcher for at least five years beyond the end of the study in the DUET Lab at Cornell University.

Appendix B- Post Environment Questionnaire

Appendix B includes the post-environment questionnaire distributed immediately after each participant exited both virtual hotel environments. The questionnaire was digitally distributed through Qualtrics on a desktop computer. This questionnaire measures spatial presence, affective reactions, and perceived aesthetic value.

Table B1 includes questions for all sub-scales of spatial presence except domain specific interest. Spatial presence questions were adapted from the MEC Spatial Presence Questionnaire (Vorderer et al., 2004), using a 7-point Likert scale. Qualtrics randomized all questions were. Questions 1-6 measure attention allocation, questions 7-12 measure spatial situation model (SSM), questions 13-18 measure spatial presence, questions 19-24 measure spatial presence: possible actions (SPPA), questions 25-30 measure higher cognitive involvement, questions 31-36 measure suspension of disbelief, and questions 37-42 measure visual spatial imagery.

Table B1
Spatial Presence Questionnaire Part 1

		Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly agree
1	I devoted my whole attention to the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
2	I concentrated on the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
3	My attention was claimed by the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
4	The virtual hotel room captured my senses.	❍	❍	❍	❍	❍	❍	❍
5	I dedicated myself completely to the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍

6	My perception focused on the virtual hotel room almost automatically.	❍	❍	❍	❍	❍	❍	❍
7	I was able to imagine the arrangement of the spaces presented in the virtual hotel room very well.	❍	❍	❍	❍	❍	❍	❍
8	I had a precise idea of the spatial surroundings presented in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
9	I was able to make a good estimate of the size of the presented space.	❍	❍	❍	❍	❍	❍	❍
10	I was able to make a good estimate of how far apart things were from each other.	❍	❍	❍	❍	❍	❍	❍
11	Even now, I still have a concrete mental image of the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
12	Even now, I could still find my way around the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
13	I felt like I was a part of the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
14	I felt like I was actually there in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
15	I felt like the objects in the virtual hotel room surrounded me.	❍	❍	❍	❍	❍	❍	❍
16	It was as though my true location had shifted into the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍

17	I felt as though I was physically present in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
18	It seemed as though I actually took part in the action of the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
19	I had the impression that I could act in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
20	I had the impression that I could be active in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
21	I felt like I could move around among the objects in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
22	The objects in the presentation gave me the feeling that I could do things with them.	❍	❍	❍	❍	❍	❍	❍
23	It seemed to me that I could have some effect on things in the presentation, as I do in real life.	❍	❍	❍	❍	❍	❍	❍
24	It seemed to me that I could do whatever I wanted in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
25	I thought most about things having to do with the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
26	I thought intensely about the meaning of the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
27	I thoroughly considered what the things in the virtual hotel room had to do with one another.	❍	❍	❍	❍	❍	❍	❍

28	The virtual hotel room activated my thinking.	❍	❍	❍	❍	❍	❍	❍
29	I thought about whether the virtual hotel room could be of use to me.	❍	❍	❍	❍	❍	❍	❍
30	I thought about just how much I know about the things in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
31	I concentrated on whether there were any inconsistencies in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
32	I didn't really pay any attention to the existence of error or inconsistencies in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
33	I directed my attention to possible error or contradictions in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
34	I took a critical viewpoint of the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
35	It was important for me to check whether inconsistencies were present in the virtual hotel room.	❍	❍	❍	❍	❍	❍	❍
36	It was not important for me whether the virtual hotel room contained errors or contradictions.	❍	❍	❍	❍	❍	❍	❍
37	When someone shows me a blueprint, I am able to imagine the space easily.	❍	❍	❍	❍	❍	❍	❍
38	It's easy for me to negotiate a space in my mind without actually being there.	❍	❍	❍	❍	❍	❍	❍

39	When I read a text, I can usually easily imagine the arrangement of the objects described.	❍	❍	❍	❍	❍	❍	❍
40	When someone gives me directions to a place, I can picture the route as though I were watching a film.	❍	❍	❍ 2	❍	❍	❍	❍
41	When someone describes a space to me, it's usually very easy for me to imagine clearly.	❍	❍	❍	❍	❍	❍	❍
42	When a picture shows only part of the space, I can clearly imagine the rest of the space.	❍	❍	❍	❍	❍	❍	❍

Table B2 includes questions from the final sub-scale of spatial presence, domain specific interest. These questions were not randomized and presented as a single block separate from the other spatial presence sub-scales, because of the items wording. B2 shows the distributed questions for the head mounted display condition. The screen condition questionnaire replaced all instances of "head-mounted display," with "screen display."

Table B2

Spatial Presence Questionnaire Part 2 (Domain Specific Interest)

		Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
1	I am generally interested in the topic of head mounted displays.	❍	❍	❍	❍	❍	❍	❍
2	The head mounted displays corresponded very well with what I normally prefer.	❍	❍	❍	❍	❍	❍	❍
3	I have felt a strong affinity to head mounted displays for a long time.	❍	❍	❍	❍	❍	❍	❍
4	There was already a fondness in me for the topic of head mounted displays before I was exposed to it today.	❍	❍	❍	❍	❍	❍	❍
5	Things like a head mounted display have often attracted my attention in the past.	❍	❍	❍	❍	❍	❍	❍
6	I just love to think about the topic of head mounted displays.	❍	❍	❍	❍	❍	❍	❍

Table B3 includes questions for all sub-scales of aesthetic value. Aesthetic value questions were adapted from the Visual Properties Questionnaire (Hanyu, 1997), except materiality which were self-developed through a focus group. The Visual Properties Questionnaire uses a 5 point Likert scale from “none at all” to “a great deal.” Qualtrics randomized all questions. Questions 1-3 measure complexity, questions 4-8 measure legibility, questions 9-11 measure coherence, questions 12-14 measure mystery, questions 15-16 measure openness, question 17 measures typicality, question 18 measures brightness, question 19 measures uniform lighting, question 20 measures familiarity, and questions 21-22 measure materiality.

Table B3
Aesthetic Value Questionnaire

		None at all	A little	A moderate amount	A lot	A great deal
1	How much is going on the virtual hotel room?	❍	❍	❍	❍	❍
2	How much is there to look at?	❍	❍	❍	❍	❍
3	How much does the scene contain many elements?	❍	❍	❍	❍	❍
4	How easily would it be to find your way around the virtual hotel room depicted?	❍	❍	❍	❍	❍
5	How easily could you orient yourself within the virtual hotel room?	❍	❍	❍	❍	❍
6	How easily would it be to figure out where you are at any given moment?	❍	❍	❍	❍	❍
7	How easily would it be to find out your way back to any given point in the virtual hotel room?	❍	❍	❍	❍	❍
8	How easily could you draw a floor plan of the virtual hotel room?	❍	❍	❍	❍	❍

9	How well does the scene hang together?	❍	❍	❍	❍	❍
10	How easy is it to organize and structure the scene?	❍	❍	❍	❍	❍
11	How much did the virtual hotel room feel complete?	❍	❍	❍	❍	❍
12	How much does the virtual hotel room promise more to be seen if you could walk deeper into it?	❍	❍	❍	❍	❍
13	To what degree does the virtual hotel room hide either positive or negative encounters that may lie ahead?	❍	❍	❍	❍	❍
14	If given more time, I would you like to explore the virtual hotel room more.	❍	❍	❍	❍	❍
15	How much feeling of spaciousness or depth does the scene convey?	❍	❍	❍	❍	❍
16	How deep and wide can you see into the environment from your point?	❍	❍	❍	❍	❍
17	The virtual hotel room is representative of a real hotel room	❍	❍	❍	❍	❍
18	To what extent does the virtual hotel room have bright, clear lighting?	❍	❍	❍	❍	❍
19	To what extent does the scene have uniform lighting?	❍	❍	❍	❍	❍
20	How well do you feel like you know the environment after the presentation?	❍	❍	❍	❍	❍
21	To what extend do the materials in the virtual hotel room look like materials in the real world?	❍	❍	❍	❍	❍
22	The materials in the environment matched what I expect in reality	❍	❍	❍	❍	❍

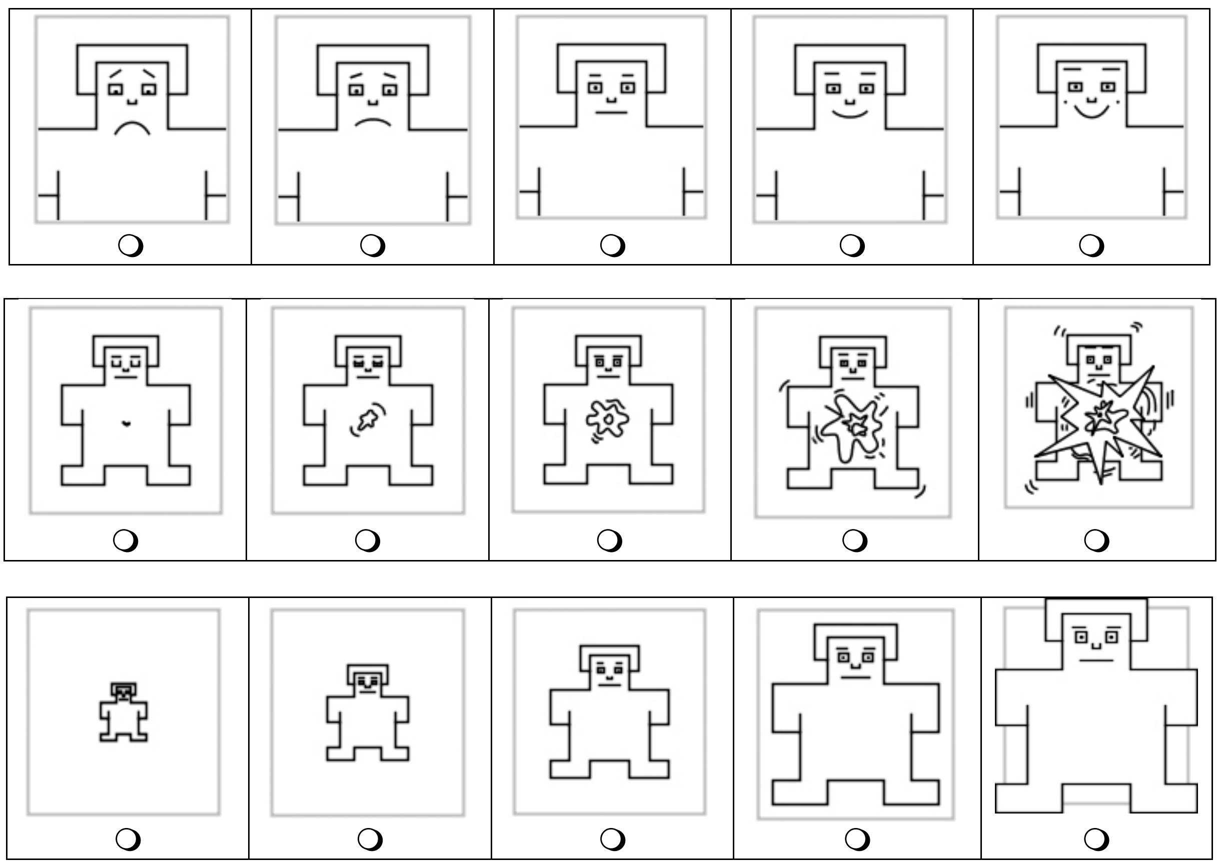
Table B4 includes the questions for the semantic differential, a measure of affective reactions (Russel & Mehrabian, 1974). The semantic differential includes 18 bipolar word pairings from 1-9, with a score of 5 marked as neutral. Qualtrics randomized all questions. Additionally, odd questions had a negative bipolar choice on the left, while even questions had a negative bipolar choice on the right. Scores were reversed before analysis. Questions 1-6 measure pleasure, questions 7-12 measure arousal, and questions 13-18 measure dominance.

Table B4
Semantic Differential

						Neutral					
1	Unhappy	❍	❍	❍	❍	❍	❍	❍	❍	❍	Happy
2	Pleased	❍	❍	❍	❍	❍	❍	❍	❍	❍	Annoyed
3	Unsatisfied	❍	❍	❍	❍	❍	❍	❍	❍	❍	Satisfied
4	Contended	❍	❍	❍	❍	❍	❍	❍	❍	❍	Melancholic
5	Despairing	❍	❍	❍	❍	❍	❍	❍	❍	❍	Hopeful
6	Relaxed	❍	❍	❍	❍	❍	❍	❍	❍	❍	Bored
7	Relaxed	❍	❍	❍	❍	❍	❍	❍	❍	❍	Stimulated
8	Excited	❍	❍	❍	❍	❍	❍	❍	❍	❍	Calm
9	Sluggish	❍	❍	❍	❍	❍	❍	❍	❍	❍	Frenzied
10	Jittery	❍	❍	❍	❍	❍	❍	❍	❍	❍	Dull
11	Sleepy	❍	❍	❍	❍	❍	❍	❍	❍	❍	Wide Awake
12	Aroused	❍	❍	❍	❍	❍	❍	❍	❍	❍	Unaroused
13	Controlled	❍	❍	❍	❍	❍	❍	❍	❍	❍	Controlling
14	Influential	❍	❍	❍	❍	❍	❍	❍	❍	❍	Influenced
15	Cared For	❍	❍	❍	❍	❍	❍	❍	❍	❍	In Control
16	Important	❍	❍	❍	❍	❍	❍	❍	❍	❍	Awed
17	Submissive	❍	❍	❍	❍	❍	❍	❍	❍	❍	Dominant
18	Autonomous	❍	❍	❍	❍	❍	❍	❍	❍	❍	Guided

Table B5 includes the questions for the self-assessment manikin (SAM), a visual measure of affective reactions including three pictorial scales (Lang, 1980). The first row represents pleasure; the second row represents arousal, and the final row represents dominance.

Table B5
Self-Assessment Manikin



Appendix C- Virtual Hotel Environments

Appendix C includes pictures of the two virtual hotel environments, in addition to benchmarking photos used during the development process.

Figure C1
Luxury Hotel Environment- Benchmarking Photos

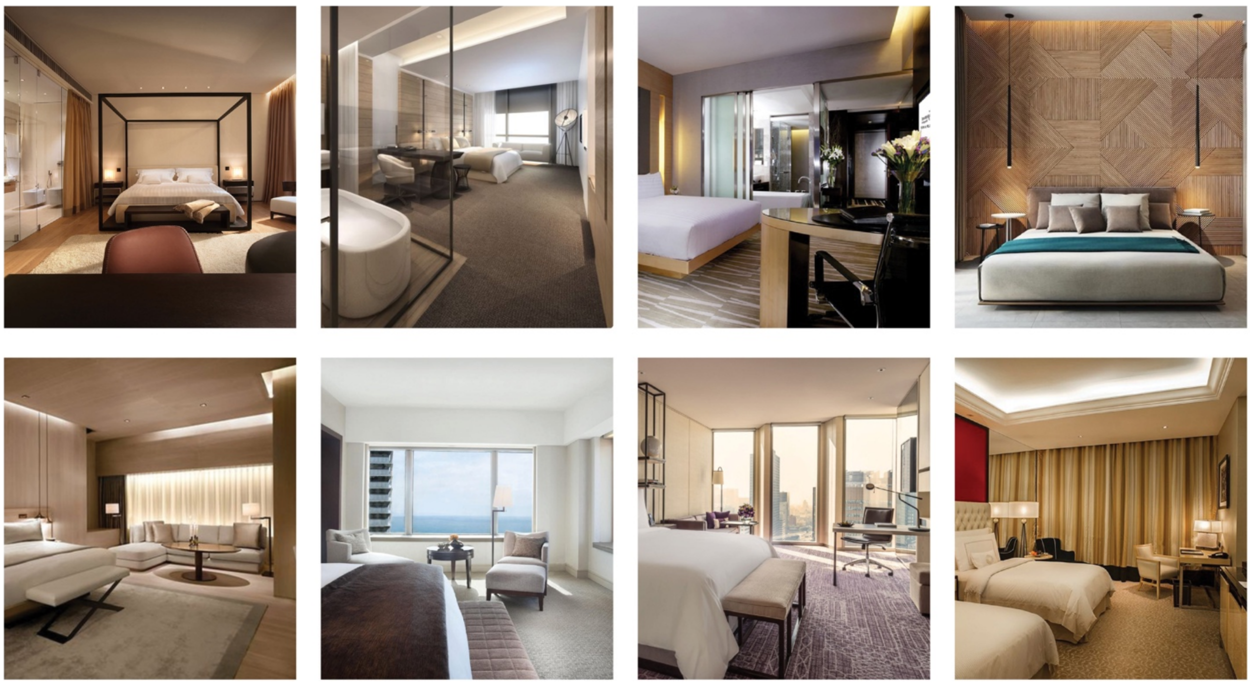


Figure C2
Virtual Luxury Hotel Environment


INTERIORS

Figure C3

Economy Hotel Environment- Benchmarking Photos

Figure C4
Economy Hotel Environment

References

Ahn, D., Seo, Y., Kim, M., Kwon, J. H., Jung, Y., Ahn, J., & Lee, D. (2014). The effects of actual human size display and stereoscopic presentation on users' sense of being together with and of psychological immersion in a virtual character. *Cyberpsychology, Behavior, and Social Networking*, *17*(7), 483-487.Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49-59.

Browne, G. J., & Rogich, M. B. (2001). An empirical investigation of user requirements elicitation: Comparing the effectiveness of prompting techniques. *Journal of Management Information Systems*, *17*(4), 223-249.

Canter, D. (1969). An intergroup comparison of connotative dimensions in architecture. *Environment and behavior*, *1*(1), 37.

Cetintahra, G. E., & Cubukcu, E. (2015). The influence of environmental aesthetics on economic value of housing: an empirical research on virtual environments. *Journal of Housing and the Built Environment*, *30*(2), 331-340.

Cooper, R. G., & Kleinschmidt, E. J. (1986). An investigation into the new product process: steps, deficiencies, and impact. *Journal of product innovation management*, *3*(2), 71-85.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Dahan, E., & Srinivasan, V. (2000). The predictive power of internet-based product concept testing using visual depiction and animation. *Journal of Product Innovation Management*, *17*(2), 99-109.

DeVellis, R. F. (1991). Guidelines in scale development. *Scale Development: Theory and Applications. Newbury Park, Calif: Sage*, *5191*.

Di Gironimo, G., Guida, M., & Lanzotti, A. (2013). Improving design validation of playground equipment in virtual reality. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *7*(3), 191-201.

Droz, Dan. "Prototyping: A key to managing product development." *journal of Business Strategy* 13.3 (1992): 34-38.

Engelbrektsson, P., & Söderman, M. (2004). The use and perception of methods and product representations in product development: a survey of Swedish industry. *Journal of engineering design*, *15*(2), 141-154.

Hanyu, K. (1997). Visual properties and affective appraisals in residential areas after dark. *Journal of Environmental Psychology*, *17*(4), 301-315.

Hanyu, K. (2000). Visual properties and affective appraisals in residential areas in daylight. *Journal of Environmental Psychology*, *20*(3), 273-284.

Hayes, R. H., & Wheelwright, S. C. (1979). Link manufacturing process and product life cycles. *Harvard business review*, *57*(1), 133-140.

Holbrook, M. B., & Moore, W. L. (1981). Feature interactions in consumer judgments of verbal versus pictorial presentations. *Journal of Consumer Research*, *8*(1), 103-113.

Kaplan, S. (1979). Perception and landscape: conceptions and misconceptions.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. CUP Archive.

Kaulio, M. A. (1998). Customer, consumer and user involvement in product development: A framework and a review of selected methods. *Total Quality Management*, *9*(1), 141-149.

Kober, S. E., & Neuper, C. (2013). Personality and presence in virtual reality: does their relationship depend on the used presence measure?. *International Journal of Human-Computer Interaction*, *29*(1), 13-25.

Koskinen, K. U. (2012). Knowledge integration in systems integrator type project-based companies: a systemic view. *International Journal of Managing Projects in Business*, *5*(2), 285-299.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications.

Mon-Williams, M., Warm, J. P., & Rushton, S. (1993). Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. *Ophthalmic and Physiological Optics*, *13*(4), 387-391.

Mon-Williams, M., Plooy, A., Burgess-Limerick, R., & Wann, J. (1998). Gaze angle: a possible mechanism of visual stress in virtual reality headsets. *Ergonomics*, *41*(3), 280-285.

Nunnally, J. (1978). Psychometric methods.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological bulletin*, *49*(3), 197.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.

Robinett, W., & Rolland, J. P. (1992). A computational model for the stereoscopic optics of a head-mounted display. *Presence: Teleoperators & Virtual Environments*, *1*(1), 45-62.

Scarfe, P., & Glennerster, A. (2015). Using high-fidelity virtual reality to study perception in freely moving observers. *Journal of vision*, *15*(9), 3-3.

Schrage, M. (1993). The culture (s) of prototyping. *Design Management Journal (Former Series)*, *4*(1), 55-65.

Slater, M., Linakis, V., Usoh, M., Kooper, R., & Street, G. (1996, July). Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. In *ACM virtual reality software and technology (VRST)* (pp. 163-172). New York, NY: ACM Press.

Söderman, M. (2005). Virtual reality in product evaluations with potential customers: An exploratory study comparing virtual reality with conventional product representations. *Journal of engineering design*, *16*(3), 311-328.

Srinivasan, V., Lovejoy, W. S., & Beach, D. (1997). Integrated product design for marketability and manufacturing. *Journal of Marketing Research*, 154-163.

Stevens, G. A., & Burley, J. (1997). 3,000 raw ideas= 1 commercial success!. *Research-Technology Management*, *40*(3), 16-27.

Sutherland, I. E. (1968, December). A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I* (pp. 757-764). ACM.

Ulrich, R. S. (1993). Biophilia, biophobia, and natural landscapes. *The biophilia hypothesis*, *7*.

Vorderer, P., Wirth, W., Saari, T., Gouveia, F. R., Biocca, F., Jäncke, F., ... & Laarni, J. (2003). Constructing Presence: Towards a two-level model of the formation of Spatial Presence. *Unpublished report to the European Community, Project Presence: MEC (IST-2001-37661). Hannover, Munich, Helsinki, Porto, Zurich*.

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., ... & Klimmt, C. (2004). MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application. *Report to the European Community, Project Presence: MEC (IST-2001-37661)*, *3*.

Watson, B., Friedman, A., & McGaffey, A. (2001, August). Measuring and predicting visual fidelity. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (pp. 213-220). ACM.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., ... & Biocca, F. (2007). A process model of the formation of spatial presence experiences. *Media psychology*, *9*(3), 493-525.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, *7*(3), 225-240.

Wheelwright, S. C., & Clark, K. B. (1994). Accelerating the design-build-test cycle for effective product development. *International Marketing Review*, *11*(1), 32-46.