Eye Tracking Consumer Purchase Behavior Within Physical and Virtual Environments

David Andrew Stone
Clemson University

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

Part of the Materials Science and Engineering Commons

Recommended Citation
Stone, David Andrew, "Eye Tracking Consumer Purchase Behavior Within Physical and Virtual Environments" (2015). All Theses. 2150.
https://tigerprints.clemson.edu/all_theses/2150

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
EYE TRACKING CONSUMER PURCHASE BEHAVIOR WITHIN PHYSICAL AND VIRTUAL ENVIRONMENTS

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Masters of Packaging Science
Packaging Science

by
David Andrew Stone
May 2015

Accepted by:
Dr. Rupert Andrew Hurley, Committee Chair
Dr. Charles Tonkin
Dr. E. Jeffrey Rhodehamel
ABSTRACT

Understanding how consumers observe and make purchase decisions within a retail context is now both accessible and efficient through the process of eye tracking. Eye tracking package design aesthetics helps us understand and predict what consumers are looking at, and how likely a package might be selected. Typically, this research is conducted in an immersive retail setting where consumers can shop as they would in a normal store-shopping context. A store is stocked with products where a participant in the study shops throughout while wearing an eye tracker to gather data on what their attention fixates on within a given set of shelves. Although a physical store provides the most realistic context, a virtual store could create a more economical, cost effective, and customizable solution for measuring consumer visual attention from packaging design aesthetics.

Beginning with CUshop Consumer Experience Laboratory, a virtual store design and context was established by replicating existing fixtures in CUshop™. Using the virtual technology available at the Sonoco Institute of Packaging Design and Graphics, a digital replication of CUshop™ was created. This began by 3D modeling the store along with generating the exact content to be displayed using real time rendering software. To investigate the process of measuring consumer attention in each environment, the same study was conducted in both stores looking at shelf performance of eleven different barbecue sauce brands. Gaze data, travel time, purchase decision and presence survey scores from a modified Witmer-Singer survey helped demonstrate the feasibility of gathering valid results from a virtual store context.
Results indicated that there was not enough evidence to prove a comparison between the physical and virtual store experiments. Presence scores also did not indicate significant differences between either store environments. Analysis suggests that with a larger participant population and more immersive hardware, such as head mounted displays, eye tracking in virtual stores could be a valid process to complement studies already being conducted in real store contexts.
ACKNOWLEDGMENTS

I would like to first thank my family and friends who have been both encouraging and supportive throughout this entire process.

Thank you to my advisor, Dr. Andrew Hurley, and my committee members, Dr. Chip Tonkin and Dr. Jeff Rhodehamel for their guidance and support in helping me investigate this research topic. Also, a big thank you to Bobby Congdon, who has been a tremendous source of help in making this thesis thorough and accurate. Thank you all for your time and collaboration.

I would also like to thank the Sonoco Institute of Packaging Design and Graphics for the tools and education they provided me over the past two years.

A special thanks to Esko™ and Tobii™ for providing the technology and software needed to conduct this study from beginning to end.

And lastly, thank you to all the graduate students, present and past. I have thoroughly enjoyed working, traveling and studying with all of you. I wish you the best of luck in your future endeavors.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>..........................................................</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>..........................................................</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>..........................................................</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>..........................................................</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>..........................................................</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I.</td>
<td>INTRODUCTION</td>
<td>..........................................................</td>
</tr>
<tr>
<td>II.</td>
<td>REVIEW OF LITERATURE</td>
<td>..........................................................</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eye Tracking in Consumer Purchase Behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Virtual Environments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Immersion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Presence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Realism in Virtual Environments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screen Size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Field of View</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Navigation and Interaction</td>
</tr>
<tr>
<td>III.</td>
<td>PILOT STUDY</td>
<td>..........................................................</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methodology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Procedure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conclusion</td>
</tr>
<tr>
<td>III.</td>
<td>METHODOLOGY</td>
<td>..........................................................</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objectives</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hypothesis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stimulus</td>
</tr>
</tbody>
</table>
Table of Contents (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparatus</td>
<td>36</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>40</td>
</tr>
<tr>
<td>Procedure</td>
<td>42</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>47</td>
</tr>
<tr>
<td>IV. RESULTS AND DISCUSSION</td>
<td>49</td>
</tr>
<tr>
<td>Survey Results</td>
<td>49</td>
</tr>
<tr>
<td>Eye Tracking Metrics</td>
<td>60</td>
</tr>
<tr>
<td>Shopping Times</td>
<td>71</td>
</tr>
<tr>
<td>Shopping Data</td>
<td>75</td>
</tr>
<tr>
<td>V. CONCLUSION AND LIMITATIONS</td>
<td>78</td>
</tr>
<tr>
<td>V. RECOMMENDATIONS</td>
<td>82</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>83</td>
</tr>
<tr>
<td>A: Survey Questions</td>
<td>84</td>
</tr>
<tr>
<td>B: Demographic and Survey Results: Physical CUshop</td>
<td>89</td>
</tr>
<tr>
<td>C: Demographic and Survey Results: Virtual CUshop</td>
<td>93</td>
</tr>
<tr>
<td>D: Product Shelf Images</td>
<td>97</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>99</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pilot Study: Mean responses from the modified Witmer-Singer presence questionnaire</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>Mean responses from the modified Witmer-Singer presence questionnaire</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>Total Fixation Duration (TFD mean averages and p-values based on t-tests results.</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>TFD ranking of all eleven barbecue sauces based on Physical and Virtual CUshop average TFD times</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>Time to First Fixation (TTFF) mean averages and p-values based on t-tests results.</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>TTFF ranking of the eleven barbecue sauces based on the Physical and Virtual average TTFF times</td>
<td>66</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Heat Map and Gaze Plot</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Human-VE interaction loop (Bowman &amp; McMahan, 2007)</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Ray-casting rendering versus ray-tracing (Slater et al., 2009)</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Physical shelves compared to projected shelves</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>Test results of relations between interaction techniques and display types</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>Pilot Study: Stimuli for VCU and VR60</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Pilot Study: TTFF and TFD averages</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>CUshop Consumer Experience Laboratory plan view</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>CUshop laboratory</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>Virtual CUshop</td>
<td>34</td>
</tr>
<tr>
<td>11</td>
<td>Virtual CUshop plan view</td>
<td>35</td>
</tr>
<tr>
<td>12</td>
<td>3D modeling process for BBQ sauce models</td>
<td>36</td>
</tr>
<tr>
<td>13</td>
<td>Tobii eye tracking glasses with recording assistant</td>
<td>37</td>
</tr>
<tr>
<td>14</td>
<td>Tobii X2-60 screen mounted eye tracker</td>
<td>39</td>
</tr>
<tr>
<td>15</td>
<td>Virtual CUshop experiment setup</td>
<td>39</td>
</tr>
<tr>
<td>16</td>
<td>Keyboard and mouse used for navigation</td>
<td>40</td>
</tr>
<tr>
<td>17</td>
<td>Shopping lists</td>
<td>41</td>
</tr>
<tr>
<td>18</td>
<td>Planogram for barbecue sauce placement</td>
<td>41</td>
</tr>
<tr>
<td>19</td>
<td>Participant recording purchase decision on shopping list</td>
<td>42</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>20</td>
<td>Starting point in Virtual CUshop</td>
<td>43</td>
</tr>
<tr>
<td>21</td>
<td>Virtual CUshop participant navigating through store</td>
<td>44</td>
</tr>
<tr>
<td>22</td>
<td>Virtual CUshop training environment plan view</td>
<td>45</td>
</tr>
<tr>
<td>23</td>
<td>Walkthrough of training environment</td>
<td>46</td>
</tr>
<tr>
<td>24</td>
<td>Biological gender demographic</td>
<td>50</td>
</tr>
<tr>
<td>25</td>
<td>Physical CUshop demographics</td>
<td>51</td>
</tr>
<tr>
<td>26</td>
<td>Virtual CUshop demographics</td>
<td>51</td>
</tr>
<tr>
<td>27</td>
<td>What barbecue sauce do you commonly use? Physical Survey</td>
<td>53</td>
</tr>
<tr>
<td>28</td>
<td>What barbecue sauce do you commonly use? Virtual survey</td>
<td>53</td>
</tr>
<tr>
<td>29</td>
<td>How many hours do you spend using 3D programs and/or computer games during a month? Virtual survey</td>
<td>54</td>
</tr>
<tr>
<td>30</td>
<td>Do you have any experience with video games? Virtual survey</td>
<td>55</td>
</tr>
<tr>
<td>31</td>
<td>If so please check off the types of video games you have played? Virtual survey</td>
<td>56</td>
</tr>
<tr>
<td>32</td>
<td>What video game platforms have you had experience with? Virtual survey</td>
<td>56</td>
</tr>
<tr>
<td>33</td>
<td>Mean and standard error of presence scores</td>
<td>58</td>
</tr>
<tr>
<td>34</td>
<td>Illustration of eye tracking AOI’s through heatmap color scheme</td>
<td>61</td>
</tr>
<tr>
<td>35</td>
<td>Total Fixation Duration Averages</td>
<td>63</td>
</tr>
<tr>
<td>36</td>
<td>Physical CUshop Time to First Fixation averages</td>
<td>66</td>
</tr>
<tr>
<td>37</td>
<td>Virtual CUshop Timer to First Fixation averages</td>
<td>67</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>38</td>
<td>Noted fixations on all barbecue sauces in Physical CUshop</td>
<td>69</td>
</tr>
<tr>
<td>39</td>
<td>Noted fixations on all barbecue sauces in Virtual CUshop</td>
<td>70</td>
</tr>
<tr>
<td>40</td>
<td>Average shopping times for Physical and Virtual CUshop</td>
<td>71</td>
</tr>
<tr>
<td>41</td>
<td>Shopping times organized by age</td>
<td>72</td>
</tr>
<tr>
<td>42</td>
<td>Distribution of travel time and search time in Physical CUshop</td>
<td>74</td>
</tr>
<tr>
<td>43</td>
<td>Distribution of travel time and search time in Virtual CUshop</td>
<td>74</td>
</tr>
<tr>
<td>44</td>
<td>Purchase Decision choices between Physical and Virtual CUshop</td>
<td>75</td>
</tr>
<tr>
<td>45</td>
<td>Purchase decision based on top five performing brands</td>
<td>76</td>
</tr>
<tr>
<td>46</td>
<td>Comparison of BBQ Sauce 4 barbecue sauce in Physical and Virtual CUshop</td>
<td>77</td>
</tr>
</tbody>
</table>
INTRODUCTION

The packaging design workflow heavily relies on knowing what is appealing and engaging to the consumer. There are few methods available for successfully determining the performance of a package based on its aesthetic value. Eye tracking can provide definitive answers for how consumers shop and what products are grabbing their attention on the shelf. This means that consumers do not have to physically pick the product up off the shelf in order to decide on what they are purchasing. (Clement, 2007) Eye tracking allows us to numerically quantify and visualize the interaction of the consumer’s overt attention within the retail environment.

CUshop™, henceforth referred to as Physical CUshop, is a consumer experience laboratory used in the Sonoco Institute of Packaging and Graphics at Clemson University to conduct performance studies of products in a realistic retail context. The store layout consists of many items and features you would find in a retail environment. These include aisles with shelving, a produce section, frozen food displays and other wall decoration that help immerse the participant in a retail environment. Conducting studies in a realistic retail store allows for natural responses to packaging stimuli. This can come at a high cost without the assurance that the time spent developing and prototyping a product will yield definitive results. An alternative to these types of tests would be to generate products virtually in order to understand a packages performance before it goes through the prototypical phase. This could potentially save money, materials and time. Recent advancements in 3D rendering technology and modeling provides the capability
of generating photo realistic virtual environments that could be used for testing varieties of products and packages.

The virtual version of CUshop™, henceforth referred to as Virtual CUshop, is a non-immersive, navigable virtual environment. This means that the hardware used to conduct the study was based off of a best-case scenario, based on the technology available at the time of the experiment. The participants freely navigated the environment using a keyboard and mouse, spending as much time as needed to finish the shopping experience. Virtual CUshop was created by incorporating programs such as Rhinoceros 3D™, Google Sketchup™ and Esko Studio™ Store Visualizer™. These programs congruently generate a realistic virtual store to try and elicit the most natural responses by measuring consumer’s overt attention to packaging stimuli.

The methodology developed for comparing Physical to Virtual CUshop was based on the shelf performance evaluation of eleven brands of barbecue sauce. The study was first conducted in the Physical CUshop using eye tracking glasses to measure Time to First Fixation, Total Fixation Duration and shopping times. The same study was replicated in Virtual CUshop with identical stimuli throughout the whole store to ensure a 1:1 comparison. After completing two studies, the data was analyzed to compare the studies and determine the feasibility of eye tracking consumer purchase behavior in virtual environments.

Adoption of virtual store simulations in consumer buying and research is a natural evolution of technology. Additional validation is necessary to prove that eye tracking in virtual environments is a suitable methodology. As more brand owners and service
providers begin to adopt virtual studies as a research tool, the accuracy and effectiveness will continue to grow through future research. Studies are going to consistently demand higher levels of realism, presence and immersion within these environments, which will be the driving force for the technology to continually evolve. Meaning that the pursuit of research in virtual reality will eventually be able to compliment studies in the physical world. Meaning this could eventually be a process that would help save time and money when developing products and packaging for retail by providing faster, iterative processes.
**REVIEW OF LITERATURE**

*Eye Tracking Consumer Purchase Behavior*

Understanding how consumers observe and interact with packaging has become increasingly more pertinent and accessible through the advent of eye tracking technology. When defined, eye tracking is a process by which the movement of the eye is recorded in order to investigate and understand a cognitive process. (Duchowski, 2007) During the earliest research in the late 1800s, physiologists and psychologists used eye tracking to interpret how people read and understand written documents. As the process is conducted, the gaze data is collected from the tracker and translated into data that indicate exact eye movements caused by looking at the stimuli. Eye movements can be described by two primary characteristics, saccades and fixations. Saccades are known as the continuous movement of the eye. (Duchowski, 2007) Fixations are when the eye is relatively still in between saccades. (Duchowski, 2007) The term fixation is used relatively because the eye is never truly still causing fixations generally to last for around 200-300ms. Saccades generally last between 50 and 150ms, occurring multiple times per second. (Gofman, Moskowitz, Fyrbjork, Moskowitz, & Mets, 2009) Eye tracking has become a foundational tool used across all forms of media including billboards, web design, and magazines.

Over the past decade, eye tracking has evolved as a reliable tool in the packaging industry to help determine the overall effectiveness of a product or package in retail. Understanding what people look at with certain forms of packages can be an increasingly
useful tool for designers and brand owners alike. The packaging design workflow has become heavily influenced by what is appealing to the customer. This shows that eye tracking is a suitable method for quantifying how individuals shop and buy within a retail context. Previously, existing methods for analyzing packaging aesthetics ranged from focus groups to question based surveys that could lead to unsatisfactory or inconclusive results. (Breen, 2008; Randall, 2013) Studies show that 70% of consumer purchase decisions are made at the shelf, 85% without picking up a competitive item and 90% make a purchase decision after only examining the front of the packaging without having the product in hand. (Clement, 2007; Urbany, Dickson, & Kalapurakal, 1996). This means that a product’s visual graphics can offer cues that act as a selling stimulus for the product without the consumer even having to interact with the package.

It’s also been shown that packaging containing realistic images of a product were preferred by the consumer and that the purchasing decision is heavily influenced by the quality perceived from these images. (Underwood, 2001) Eye tracking analysis allows one to see the consumer’s purchase intentions through data analysis in addition to being quantified visually through heat maps or gaze data (Figure 1) with statistical metrics to validate differences.
Traditionally, these experiments are conducted in a retail environment similar to CUshop to simulate the atmosphere of a real shopping market or store. (Hurley, 2011)

The hardware used to conduct these experiments are a pair of eye tracking glasses that record where the user is looking in space to quantify their fixations from the shelves. These recorded fixations and eye movements are used primarily to test a packages shelf presence and quantify data gathered from the experiment as proven in several experiments. (Hurley, Galvarino, Thackston, Ouzts, & Pham, 2012; Hurley, 2011; Hutcherson, 2013; Randall, 2013; Thackston, Pham, Galvarino, & Ouzts, 2011)
Eye-tracking glasses are not the only hardware available for capturing fixation data though. Many studies use alternative hardware such as mounted trackers designed to observe consumers’ eye movements when looking at stimuli on a screen or projected on a wall. (Bazyluk, 2010; Tan, Gergle, Scupelli, & Pausch, 2003; Tonkin, Ouzts, & Duchowski, 2011; Vora et al., 2002, etc) Mounting the eye tracking hardware allows for testing of stimuli to expand beyond the physical and into graphical/virtual mediums such as computer displays, projectors, simulators, and still images. Mounted eye trackers work similarly to the glasses, except the data is captured through a camera mounted within 36 inches of the participant. In the context of packaging, this means that the experiment would be conducted completely through the computer and use stimulus that are contained within a 3D model or picture in a virtually rendered environment.

Virtual Environments

Over the past decade, companies have begun to test the viability of performing eye-tracking studies in virtual environments. (Breen, 2008) Many consumers and companies are beginning to see virtual experimentation as a means to perform rapid testing of products and graphics long before they are sent to production. In addition, other applications such as phobia therapy, military training, and entertainment have become large consumers of the technology, especially in the gaming industry. ( Bowman & McMahan, 2007) In the world of packaging and consumer sales, a virtual store could allow for a more economical, cost efficient and quick way to analyze user purchase decisions outside of a retail context. In 2008, Wal-Mart announced plans to make virtual-
store to test simulation, a key in its research practice. In addition, Kimberly-Clark and Proctor & Gamble have built extensive 3D simulation environments that allow consumers to walk and shop through store simulations. (Breen, 2008) Virtual environments can eventually be a source for major corporations to expedite their design process through quick iterations, versus conducting a full-scale study in a built setting that would require physical prototypes. However, much of the current research being conducted is investigating the relationships of virtual technology and consumer experience.

Researchers are investigating other applications of virtual environments in the context of online shopping and alternate realities within video games. Entrepreneurs have started to see the potential of virtual environments for e-commerce and have begun setting up stores within platforms like Second Life to sell virtual shoes, food, residences and furniture. With over seven million “inhabitants”, Second Life (SL) is really another world within itself. (Hassouneh, 2008) Brands like Nike and Reebok built stores within SL to allow customers to create custom shoes to be used both inside and outside of the virtual environment. Reebok was able to sell 27,000 pairs of digital shoes within their first ten weeks of sales. (Hassouneh, 2008) These congruencies between the real and virtual worlds show how the two can intertwine to positively influence shopping behaviors. It has been shown that there is a “perceived enjoyment” proven to be a significant contributor to predicting how people will interact when purchasing goods in the virtual world. (Eder, 2008) That could help show how and why people shop in virtual environments, in addition to actively engaging the user, which would result in the
potential for businesses to yield better sales results when doing commerce in virtual environments.

There are many variables to consider when designing a virtual environment for consumer shopping habits in regards to packaging. Ninety-five percent of purchase decisions are dictated through one’s unconscious thought, meaning that emotion is the primary source of decision making for the consumer when shopping. (Lindstrom, 2010) Clements said that the majority of the decisions are made without actually touching the product, meaning, “what you see is what you get.” (Clement, 2007) Therefore, to efficiently create a virtual environment that causes consumers to shop naturally, certain variables of the experience should be heavily regarded such as immersion, presence and realism.

*Immersion*

A shopping context is not only associated with content imbedded but also with how the environment is physically set up. Virtual environments may or may not be fully immersive since it is based on an objective level of sensory, but according to Slater et al, immersion can be broken down into four key terms:

- Inclusive
- Extensive
- Surrounding
- Vivid (Slater & Lotto, 2009)
The term immersion should be used simply for how and what the technology is delivering, whether it is the field of view, frame rate, stereoscopic headset, head tracking and so forth. (Slater & Lotto, 2009) In order for one to be truly immersed in an environment, there needs to be a threshold of technology required to generate this type of experience. If the environment is inclusive, the content and the hardware used must occupy the user’s peripheral both within the context of the testing space and the display type used, essentially shutting reality out. Extensive requires further integrating the human body and its sensory modalities in to the experiment. (Slater & Lotto, 2009) This could be achieved through a form of motion sensors, navigators or joysticks that allow the user to pick up objects in an environment. Surrounding is concerned with the participant’s field of view and their ability to look around using their total range of vision without interruption from the display. Head mounted displays that incorporate tracking would provide the maximum field of view with the ability of naturally looking around with your head. And lastly, Vivid pertains to the clearness; resolution and fidelity that strive to convince a user that they are actually looking at the real world. These four key terms indicate a deeper understanding but still don’t help illustrate how immersion fits into VR systematically. Figure 2 visualizes how immersion fits into the cycle of an experiment process. (Bowman & McMahan, 2007)
Figure 2: Human-VE interaction loop. (Bowman & McMahan, 2007)

The cycle begins with the model or environment, which is then processed through the computer system and rendered from the display. The user is then shown a visual representation of that data which illustrates an image that causes them to feel some level of immersion. This is technically where immersion begins and ends. If the display device is not capable of immersing the user in the environment, then there will be no recordable accounts of it. Bowman argues that interaction is completely separate from immersion and does not help raise the levels of immersion in the environment. (Bowman & McMahan, 2007) One could argue that interaction with the environment has much to do with the hardware and setup, thus naturally integrates itself with the immersive qualities of the system. (Kjeldskov, 2001; Lapointe & Vinson, 2002)
Presence

Presence is how we understand the way users psychologically perceive virtual contexts. This leads to a sense of feeling like they’re in the environment. This is not to be confused with one feeling immersed in an environment. Immersion and presence are understood to be completely different. However, they both can result in a more intuitive and realistic experience.

Slater defines presence as a subjective psychological response to which computer displays are capable of delivering the illusion of reality to the sense of the human participant. (Slater & Lotto, 2009) Furthermore, several researchers describe presence simply as a state of consciousness, a sense of “being there”, a sensation primarily experienced in the virtual technology. (K. Lee, 2004; D. S. Tan, Gergle, Scupelli, & Pausch, 2006) Therefore, it seems in order to cause the user to feel enveloped in the virtual context, there must be concomitant levels of both immersion and presence. Schroeder agrees that presence is the sense of ‘being there’ but that it is partly to do with the technology and the participants state of mind. (Schroeder, 2006) Several studies have looked at how presence can be quantified and applied to using virtual environments. Witmer and Singers’ study developed a questionnaire based on four subjects to quantify presence; Involvement, immersion, sensory fidelity and interface quality. Mean scores from these questions were based on a 1-7 Likert Scale and were designed to help draw conclusions of the levels of presence in a virtual environment. Results from the survey showed that fully immersed and present observers feel as if they are actually interacting directly, not indirectly, with the given environment. (Witmer, 1998) However, there tends
to be difficulties with proving the effectiveness of a presence questionnaire when contrasting more than one type of environment. Tonkin’s experiment comparing projected shelves to physical shelves yielded no significant differences between environments. This is partially attributed to the difficulty of comparing physical to virtual but also because there may have not been enough participants to show a difference. (Charles Tonkin, 2011) Lee’s study comparing the effects of visual realism on mixed reality simulation was also unable to draw any conclusions about presence because of the difficulty involved in comparing presence between different environments (real vs. virtual). (C. Lee, Rincon, & Meyer, 2013) It remains difficult to say what actually causes one to be present in the virtual world. Vora et al conducted an experiment where participants were asked to look around a virtually replicated fuselage of an airplane to detect mechanical errors that existed within the plane through a stereoscopic and a high definition display. The results from the study showed that the subjects experienced greater levels of involvement in the search task when using the stereoscopic display, which caused their experiences to feel as natural as real world ones. (Vora et al., 2002) It was also noted that the environments lacked depth due to only the 2D imagery available for the study that limited the environments ability to have a virtual reality effect on the participant. These studies indicate that in order to have a participant feel immersed and present in the virtual environment it takes a combination factors such as the right hardware, visual fidelity, field of view, interactivity and an environment with a high level of realism.
Realism in Virtual Environments

Realism influences the amount of presence in the virtual world, giving that sense of “being there” that fosters natural reactions when navigating virtual environments. Realism can be quantified in terms of both size and graphical quality. (Breen 2008) An ideal virtual environment would incorporate the three varieties of realism according to Ferwerda:

- Physical realism
- Photo-realism
- Functional realism (Ferwerda, 2003)

Though, creating and rendering a space that incorporates all of these factors of realism is unlikely due to the number of variables involved, such as limitations with technology and hardware in addition to several others. Physical realism, visual stimulation in the scene, can have many limitations because of the difficulties in creating 3D models that aren’t too complex to render at a consistent, smooth rate in virtual environments. Photo-realism, visual response in the scene, requires the correct hardware to be available for rendering and producing texture maps that have the photo quality and texture needed to be convincing enough. In a study testing how visual realism elicits realistic responses in VR, participants were exposed to two different rendered environments. Each version had one of two types of rendering techniques: ray-tracing; with shadows and reflections and ray-casting; had no shadows or reflections. (Slater, Khanna, Mortensen, & Yu, 2009) The experiment measured realism in addition to stress levels by exposing the participant to
the edge of a cliff. (Figure 3) The ray-tracing environment showed much higher levels of presence and stress than the ray-casting due to the level of detail rendered in the environment.

Figure 3: Ray-casting rendering (left) versus ray-tracing (right) (Slater et al., 2009)

This outcome was primarily attributed to the participant being able to see their own reflection and shadows around them leading to more immersive behavior. Slater et al indicates that reflections and shadows of a person’s body significantly add to subjective presence and appropriate physiological response. (Slater & Lotto, 2009) Functional Realism, visual information in the scene, pertains to not just the aesthetics of the environment but the quality of how it looks and feels when navigating throughout. Scott Young of Perception Research Services argues that “larger, full store contexts” will yield better realism results due to the level of detail achievable. (Breen, 2008) Programs like Esko Store Visualizer™ provide a pre-rendered and pre-sized context to build virtual shopping environments. This limits how large an experiment’s environment can become, but allows freedom and control when designing a virtual store that already has set
lighting and textures. Therefore, to properly immerse the user and engage natural responses, it is ideal to incorporate a high level of realism (shadows, reflections, textures, scale) to ensure that experiments conducted in virtual environments can illicit the most pertinent and valid data from realistic responses.

*Screen Size*

Virtual environments are heavily weighted on the context itself but also depend on how that information is effectively relayed through the type of display. Displays can be immersive, three-dimensional, motion tracked, big or small. The common issue is that there is a physical and psychological disconnection that exists between the monitor and the user when viewing the virtual world. Many of these disconnections can be caused from errors in set-up between the monitor and the user, such as lack of resolution, size and FOV. This disconnection can begin to be minimized with the addition of more immersive hardware like 3D glasses, head-mounted displays (HMD), large high definition screens, and projectors. Determining which display applies best to what experiment can be difficult and has been a common research topic in virtual reality.

Patrick et al conducted a performance study on spatial knowledge of three viewing conditions: head-mounted display, large projection screen and desktop monitor. Results indicated that the size of a desktop monitor caused participants to perform significantly worse in the study due to lack of range of viewing and scale.(Patrick et al., 2000) There was no significant difference found between head-mounted display and the projection screen though, which indicates that projection screens could be an inexpensive substitute
for the high cost of head-mounted displays. (Patrick et al., 2000) Even though it is expected that the increase in peripheral vision of a head-mounted display allow for participants to freely look around an environment, it shows that larger displays can indeed give a participant’s comparable spatial knowledge. (Patrick et al., 2000) Tonkin’s study with projected image shelves showed that consumers completed their search task significantly faster when immersed in a real world setting than with the projected shelves. (Figure 4) Several conclusions were drawn about why the projected shelves performed worse, two of which were that the fidelity of the screen was too low for participants to read the content and many participants seemed discouraged to move around the virtual shelf as they would in a store aisle. (Charles Tonkin, 2011)

![Figure 4: Physical shelves(left) compared to virtual(right)](Charles Tonkin, 2011)

After reviewing where the concentrations of fixations were on the projected image, it was noted that many participants focused on the center of the screen, causing a ‘center bias’ when looking at the shelf. (Charles Tonkin, 2011) This shows how resolution of a display and unfamiliarity of a projected shelf can heavily influence the
participant’s ability to naturally react to the environment. It also indicates that one should take much care and consideration before conducting an experiment with large displays and projectors to ensure the resolution is fitting to the task. Desney proved that users performed 26% better on spatial task orientation done on a large display. (D. Tan, 2004) He also indicated that larger screen size resulted in a 10-26% increase in effectiveness of the screen for user interaction. (D. S. Tan et al., 2006) However, this does not show that display size results are dependent of immersive and presence qualities (such as interaction and mental aids) within virtual environments. (D. S. Tan et al., 2006) Park’s study indicated that screen size was only apparent for portions of his study with motion conditions, and that still video resulted in no effect caused by screen size on a user’s subjective levels of presence. (Park, 2001) This can be explained by the effect that large displays have on a user’s FOV, which engages their peripheral vision to allow for a wider range of motion to look around in the virtual environment.

Field of View

Several studies have shown that wider FOV’s allow for more immersive experiences from visual displays. (Czerwinski, Tan, & Robertson, 2002; Lapointe & Vinson, 2002; Park, 2001; Tversky, 2001) People typically go to see movies in theatres because they want to feel immersed in the movie screen due to watching on a larger screen. This thought process could be applied very simply to testing participants in virtual environments: having a wider and larger screen essentially increases the scale of the visual aid, which should naturally cause higher levels of immersion through extended
FOV. Evidence has proven that the larger the screen, the better and more immersive the FOV becomes. A study focusing on navigation and FOV by Lapointe found that “there is a perceived increase of immersion with a larger real FOV.” (Lapointe & Vinson, 2002) However, a larger FOV may affect individuals differently when navigating virtual environments. For instance, Czerwinski et al sought to prove that women and men perceive virtual worlds differently based on FOV. (Czerwinski et al., 2002) They first proved that a wide screen does actually increase spatial understanding amongst both genders during the experiment. The second experiment tested a large display with a wider FOV than previously (75 degrees) to see if females responded any differently than the males. Results indicated that women did indeed perform better in navigation, speed and performance remarkably better than the male’s performance. (Czerwinski et al., 2002) Proving that women perform better in virtual environments with a wider FOV. Their findings indicated that a 1:1 ratio of display FOV to geometric FOV optimizes the benefits of widening FOV in virtual environments. (Czerwinski et al., 2002) In contrast, larger screens have been shown to cause negative effects in certain contexts, as ambient light created from screens can be distracting to experiments conducted in closed spaces where not much room lighting is provided. (Park, 2001) This is not to discourage large screens in experimenting environments; it simply shows that care must be taken to show not only how the virtual environment is perceived, but also how the physical context is received when participating. An alternative to using a large display would be incorporating a head mounted display as a substitute. Although expensive, head mounted displays provide the ideal amount of peripheral vision needed
to maximize FOV, leading to increased levels of immersion and ease of navigation with the ability to look around while moving throughout the environment. Bourk suggests that higher FOV is subjectively easier to navigate than lower FOV’s and can contribute to a subjective feeling of comfort. (Bourk et al., 2007) This indicates the fragility of how the experiment set-up can be in virtual worlds in regards to navigating and visual framework to create a user-friendly, yet intuitive experience.

Navigation and Interaction

Movement is one of the key factors involved in allowing a participant to feel presence or ‘being there’ in the virtual world. While also one of the most challenging aspects, it can make a major difference in the success or failure of any experiment focused around virtual environments. Movement and navigation is essentially the act of interacting with the environment. Kjeldskov defines interaction as a very broad description of computer use, and thus breaks it down into three components:

- **Orienting** oneself in virtual reality addresses the need for being able to look around in a virtual environment to develop a sense of presence.
- **Moving** in virtual reality addresses the need for being able to move around in a virtual space.
- **Acting** in virtual reality covers both the tasks of selection/picking, moving, rotating and transforming objects in the virtual environment as well as control on a system level. (Kjeldskov, 2001)
Therefore, when a participant interacts with a virtual environment there should be specific parameters as to how that set-up and experience is designed to maximize one or all three of these interaction components. Lapointe indicates that a good virtual interface must be simple to use, easy to learn, provide sustained performance and allow for fast yet accurate navigation. (Lapointe & Vinson, 2002) Unlike the space around the body, the space of navigation is generally conceived of in two dimensions. (Tversky, 2001) Meaning that when one moves through space, they are generally moving as imagined from how you would visualize if you were to look down on yourself from above while walking around an environment. This alternate perception of space creates a strong argument for simplified ways to navigate in the virtual world. While a mouse and keyboard are the most common and familiar navigation devices amongst the general population, many argue that a joystick and devices with 3 or more degrees of freedom can provide simplistic ways to navigate around virtual environments. (Kjeldskov, 2001; Lapointe & Vinson, 2002) However, it can be argued that more complex navigation devices can have a higher learning curve and cause users to fumble with the controls, rather than use more native devices commonly associated with computers.
Figure 5: Test Results of relations between interaction techniques and display types.

(Kjeldskov, 2001)

Figure 5 breaks down participants interactions in Kjeldskov’s study using two virtual contexts; a partial immersive and fully immersive environment, to understand how different navigation and interaction devices perform against each other. Based on the overall performance of the devices tested, it appears that a combination of head tracking, position tracking and a space mouse would be a highly engaging setup for a participant to act and feel as if they are in reality within the experiment. Although a testing setup with this equipment could be quite costly and cumbersome in addition to requiring much
tailoring to work effectively with the designed environment. It seems to better simulate the shopping experience; consumers should be allowed to wander up and down aisles, as they would in reality. (Charles Tonkin, 2011). A device that can closely emulate the act of physically walking or moving around would help provide the most immersive experience. Ruddle argues that full physical movement plays a vital role in navigational search and that only moderate detail is necessary. (Ruddle & Lessels, 2006) Therefore, it would be beneficial to consider whether it’s actually necessary to have a navigation framework that is so intuitive. Not all virtual studies need to use HMD’s and position tracking to achieve optimal results.(Kjeldskov, 2001) As technology in navigation evolves, devices like the Omni Treadmill and Leap motion will allow people to navigate and interact in virtual worlds in a fixed location to make them feel like they are actually physically walking around. The intended outcome is for the participant to ‘be there’ in the study, which can be achieved through many combinations of hardware, displays, and navigation devices.
PILOT STUDY

Objective

The objective of the pilot study was to develop an experiment to learn how visual realism in the virtual CUshop affected participant’s presence scores and eye tracking data. No results were found to be statistically significant, however certain aspects of the experiment indicated the need for more detailed investigation and evolution of the methodology. The following section is an overview of the pilot study containing the methodology, results and conclusions. Many areas of this experiment overlap with the final virtual CUshop and will be further detailed in the next section of the main experiment.

Methodology

The experiment was conducted in 2013 at Pack Expo Las Vegas. The study was conducted over the course of 3 days and tested a total of 126 participants. 13 of the total participants were eliminated since their data was unusable. Each day, the video stimuli changed between the three pre-recorded walkthroughs. Two were video recordings of a walkthrough of virtual CUshop at different FOV’s and the third was a video walkthrough of the physical CUshop laboratory at Clemson University. To capture the video walkthrough of CUshop lab for the experiment, a Canon Rebel T1i was mounted to a tripod with wheels and pushed along a predetermined path throughout the lab. A pilot study with three graduate students found that the mean time spent in front of each shelf searching for the product was approximately eight seconds. Therefore, the camera path
paused at the two target shelves for eight seconds each to allow adequate search time to locate the product. This video stimulus will be referred to as \textit{VCU (Video of CUshop)}.

After creating the virtual environment in various 3D programs, we established two fields of view with the scene camera at 60mm (VR60) and 90mm (VR90). This compensated for the notion that peripheral vision is used to widen the search beyond a camera's limited field of view. (D. S. Tan et al., 2006) A video sequence was generated using camera position presets placed in the scene and timed to sync up with the video recorded in the CUshop consumer lab. When played, the sequence moved and rotated around to simulate the participant walking through the virtual grocery store. The videos were exported as high-resolution .avi files at a frame rate of 24. Stimuli for the experiment were based on comparing the effectiveness of Raisin Bran cereal and 94% fat free popcorn in comparison to its competitors. (Figure 6)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Stimuli for physical CUshop (left) and virtual CUshop(right)}
\end{figure}

\textbf{Procedure}

The participant was placed in front of a 24” Tobii T60XL monitor and calibrated before the start of the experiment. Next, they were instructed to watch a video navigating through a version of CUshop (VCU, VR60, VR90) looking for popcorn and raisin bran
cereal. The pre-recorded video moved and panned around the store to each shelf location for 8 seconds, where each target product existed. Once the video concluded the participants were given a modified Witmer Singer presence survey.

**Results**

Each video-based stimuli was evaluated versus the physical store by means of modified Witmer-Singer presence questionnaire scores and eye-tracking performance metrics of time to first fixation (TTFF) and total fixation duration (TFD).

**Presence Questionnaire**

The modified Witmer-Singer survey used consisted of fourteen questions each ranked on a 7-point Likert scale. (Table 1) Questions were categorized as Involvement, Immersion, Sensory Fidelity or Interface Quality. Four of the questions were negative in nature, so scores were transformed to be consistent with the other questions, in which a high score indicated a positive response. An ANOVA for mean total score (all 14 questions) showed no significance. However, an ANOVA for the means of each of the question subcategories revealed significance for Involvement (p<0.02) and Sensory Fidelity (p<0.01), but not for Immersion or Interface Quality.
Table 1: Mean responses for the modified Witmer-Singer presence questionnaire using a 7-point Likert scale with 7 indicating the highest level of agreement and 1 indicating lowest level of agreement. Statements 5, 6, 12, and 13 are negative in nature and values were transformed when calculating Group Means for score consistency.

### Eye-Tracking Metrics

Eye-tracking metrics were calculated in Tobii Studio 3.1 using the Tobii fixation filter, and calculating gaze point using the Average method for stimuli recorded on the Tobii XL60 monitor (VCU, VR60 and VR90). Data was exported for each of the two search tasks (raisin cereal and 94% fat free popcorn) for all three video stimuli and the physical store. An ANOVA showed no significance for TTFF for either search task, but significance for TFD on the cereal (p<0.01) and the popcorn (p<0.01). For the popcorn, the VR90 had the lowest TFD. For cereal, the physical store had the lowest TFD.
Given the fixed amount of time in each of the video stimuli, it is worth noting the number of participants for each that did not locate the popcorn packages (VCU: 5.8%, VR60: 10%, and VR90: 18.2%). However, this data proved statistically insignificant. (Figure 7)

![Graphs showing mean time to first fixation and total fixation duration for popcorn and cereal across different stimuli environments.](image)

**Figure 7: Mean Time to First Fixation and Total Fixation Duration for each target item across the four stimuli environments.**

**Conclusion**

We have analyzed various methods of conducting monitor-based search tasks for packaging on a shelf versus the same search task in a physical lab environment. An ideal virtual environment would score similarly on qualitative measures of participant
presence, and would gather similar data on important eye-tracking metrics such as TTFF and TFD. Although mean scores for presence were not significantly different, two of the four subcategories in our questionnaire pointed to shortcomings in the video stimuli, presumably related to lack of navigational control. Total Fixation Duration was the only eye-tracking metric examined that showed significance, although results between target products varied in which stimuli they performed poorest. We concluded that none of the three video-based stimuli presented would yield accurate results for packaging evaluation. However, results suggest that the level of detail and rendering available in the store visualization software represent packaging on a store shelf with a sufficient degree of accuracy. Lastly, presence scores from the sensory fidelity category indicated that the ability to navigate as needed to analyze stimuli up close and from different angles could increase presence in the virtual store.
METHODOLOGY

Objectives

The purpose of this study consists of two parts; analysis of a bottle label performance study conducted in CUshop consumer experience lab followed by a 1:1 comparison of a replicated study in the virtual CUshop. Primary objectives are to obtain data relating to consumer behavior, focusing on TTFF(time to first fixation), TFD(total fixation duration), shopping time and purchase decisions. Through quantitative analysis using eye tracking metrics, the study will help gain further research in understanding the methodology needed for eye tracking consumer purchase habits and overt attention within virtual environments.

Hypothesis

The Null Hypothesis, $H_0$: Attributes collected from eye tracking such as TTFF, TFD, Purchase Decision and Shopping Times would be similar between Physical and Virtual CUshop.

The Alternative Hypothesis, $H_A$: Attributes collected from eye tracking such as TTFF, TFD, Purchase Decision and Shopping Times would not be similar between Physical and Virtual CUshop.
Participants

The physical store study in CUshop consisted of 34 participants 60% female and 40% male (Appendix B) over a 1-day period. Participants were given a $20 gift card as incentive for being a part of the study. Each participant had to meet three requirements:

- Use a grill at least once a month during grilling season.
- Buy BBQ sauce at least once every 2-3 months.
- Must at least share in-house grocery shopping.

The Virtual store experiment used 49 total participants with a similar distribution of sex: 50% Female and 50% Male. To encourage a diverse demographic of age and gender, a $10 dollar incentive was given out to those that participated. Due to poor tracking percentages, 5 of the 48 participants were eliminated from the eye tracking data, leaving 44 to be processed post-experiment. Additionally, participants who took part in the physical CUshop experiment were eliminated from being able to participate in the virtual store study in order to keep the purpose of the study undisclosed. This required removing previous participants from the emailing list and gathering a new pool for the second half of the study in the virtual store. Each participant was given unique number that correlated with his or her experiment and survey data.

Stimulus

Physical CUshop

The CUshop consumer lab is equipped with 3 aisles, each with shelving units spanning 4’ in length and 6’ in height. (Figure 8) Each shelf is approximately 16” in
depth with gloss black sheet metal plating and dark grey pegboards as a backdrop for the packages. The aisles are set at 7’ in width to allow for maximum circulation for the shopper. Also included in the CUshop is a section for fruit and frozen foods that spans the outer edges of the store. (Figure 9) Fluorescent lighting is used to mimic that of a typical grocery store including the level of lighting to provide sufficient light to view the products effectively. The floor of the CUshop has a semi-gloss brown marble texture that provides slight reflections of both the shelves and products displayed within the store.

Figure 8: CUshop Consumer Experience Laboratory plan view.
Virtual CUshop

The Virtual CUshop was measured and replicated to the exact specifications of the CUshop consumer experience lab. (Figure 11) This required extensive attention to detail when 3D modeling each component of the store. The store walls, ductwork, and doors were created in Rhinoceros 3D and exported into Sketchup™ to distinguish layers for texture mapping. The shelves, frozen food cabinets and fresh fruit containers were designed and replicated within Sketchup™ then exported as collada files (.dae). Collada files are a universal file type for importing and exporting 3D models between various programs. Each model has a custom UV map, which is applied to a texture map in order to give realistic texture to the model (wood paneling, marble floor, wall colors etc) as well as reflections and shadows. After completion of the texture mapping, the store was equipped with all of the aesthetics in the room including; light switches, electrical outlets,
wall art, branding, exit signs, base boards, etc. Upon importing the models into Studio Store Visualizer, they were then placed precisely as they exist within the physical CUshop, displaying an array of objects within the shopping environment. (Figure 10)

Figure 10: Virtual CUshop modeled with Esko Studio Store Visualizer.
Stimuli for the virtual CUshop were generated based on comparing the effectiveness of eleven different brands of barbecue on the shelf. Products were measured and modeled in 3D using Solidworks™. (Figure 12: Step 1) Followed by importing the model into Studio Visualizer for labels where the area of the label is “painted” on the model and prepared for transitioning to Adobe Illustrator to add graphics. (Figure 12: Step 2) The labels from the products are then scanned and cropped to place into Adobe Illustrator to visualize placement on the model created in SolidWorks™.
Figure 12: 3D modeling process for BBQ Sauce models.

The 3D model with graphics is then exported as a Collada archive file (.zae), and then imported into Esko Studio Store Visualizer™ for placement on the shelves. (Figure 12: Step 3) After importing adjustments are made to the bottle to match colors, reflectivity and gloss as needed for each model. (Figure 12: Step 4)

Apparatus

Two methods were used for capturing the eye tracking data for the Physical and Virtual CUshop:

Physical CUshop

Eye tracking metrics were measured using Tobii eye tracking glasses (Figure 13). The glasses have a monocular lens with a recording rate of 30 Hz [Tobii]. The eye tracking glasses are paired with IR markers to help delineate which product on the shelf
they are looking at. This is done using a plane created in Tobii Studio software called an AOA (Area of Analysis). An AOA gives feedback on specific locations on the shelf to show a fixation within an exact location. The glasses then relay that data to the recording assistant (Figure 13) which stores both the calibration and gaze data from the glasses on an SD card. Additionally to the data gathered, the glasses also record a real time video that delineate the fixation points on the shelf and can be referenced when analyzing data from the experiment. Each participant of the study must be calibrated through a 9-point vertical plane that is marked through the use of an IR marker and the accuracy is displayed on the recording assistant.

Figure 13: Tobii Eye Tracking glasses with Recording Assistant
Virtual CUshop

Data collection for the virtual store was conducted using a Tobii X2-60 (Figure 14) mounted eye tracker that was mounted onto a camera tripod in front of a 46” high-definition TV. (Figure 15) A larger screen was chosen to expand the FOV for the user to allow for a wider range of motion for eye movement. (Bourk et al., 2007) All measurements were recorded at 60 Hz sampling rate with a processing latency of <35 ms. [Tobii] The calibration process for the mounted eye tracker works similarly to the glasses in that the user is asked to look at 9 dots on the screen while refraining from moving their head as much as possible. The dot moves around the inner edges and middle of the screen, allowing the eye tracker to map the parameters of the screen in which the eye tracker will be recording. A Logitech webcam was mounted to the opposite side of the eye tracker to capture the screen recording for post processing the eye tracking data. This recorded the participant as they shopped throughout the virtual store. The webcam was chosen because the Tobii Studio screen recording function was not compatible with Store Visualizer™. Navigation for the virtual environment was conducted using a keyboard and mouse setup. The mouse controls where the camera is looking on the screen while the arrow keys from the keyboard propel the user forward and backward in the virtual environment. (Figure 16)
Figure 14: Tobii X2-60 screen mounted eye tracker

Figure 15: Virtual CUshop experiment setup
Experimental Design

The location of the experiment was in either the physical CUshop or the virtual CUshop. Participants could only take part in the virtual or the physical study in order to negate the possibility of knowing the target product beforehand or being familiar with the environment. Both studies used an identical planogram (Figure 18) that was developed to effectively arrange the barbecue sauces as they would appear in retail. Each participant was given 1 of 4 randomly arranged shopping lists with barbecue sauce, hot cocoa, sharp cheddar cheese, cookies and hair conditioner as items to purchase. (Figure 17)
Figure 17: Shopping lists (Product shelf images in Appendix D)

Figure 18: Planogram for barbecue sauce placement
Procedure

Physical CUshop

The physical CUshop experiment began with the participant entering and was assigned a participant ID number followed by a short survey gathering demographical data. They are then guided to place the Tobii glasses on their heads and fasten the security strap to ensure accurate eye tracking calibration. Each participant was run through a quick calibration using the 3pt x 3pt grid for calibrating the eye tracking glasses. Once the calibration was completed, the participant was given the shopping list and instructed to mark the item of their choosing on the checklist. (Figure 19) Each target item was paired with a reference number that was listed below the product and correlates with the target product. After completing the shopping list the participant exited the store and ended the eye-tracking portion of the experiment. Upon removal of the glasses, they were then led to a room to fill out a post survey questionnaire. Total shopping times were measured from the starting point (Figure 20) of the experiment where the participant entered the store and time stopped when they crossed the same point.

Figure 19: Participant recording purchase decision on shopping list
Virtual Store

The virtual store experiment begins by bringing the participant into a closed room where the TV and eye tracking hardware are set up. (Figure 21) A closed room was important due to the need for controlling lighting for the mounted eye tracker to ensure a minimal amount of glare and distraction from surroundings. (Park, 2001) After asking them to stand comfortably in front of the TV, the eye tracker was adjusted according to the users height and distance from the screen. Before beginning the calibration process, the participant was guided through a simple training environment to familiarize them with the navigation controls. (Figure 22) The training environment was tailored specifically to the navigation controls of this experiment to ensure participants felt...
comfortable using the devices. During the training they were asked to navigate the maze-like turns until they get to a small set of shelves with generic product at the end. (Figure 23) Once they arrived at the shelves, instructions were given to find different products on each shelf to indicate that they were now familiar with the controls. Once confirmed, the main part of the experiment began.

*Figure 21: Virtual CUpshop participant navigating through store*
Figure 22: Virtual CUshop training environment plan view with camera angles from Figure 24
Figure 23: Walkthrough of training environment

Once they feel familiar using the keyboard and mouse, the experiment begins.

The researcher prompts them to keep their head as still as possible while running the eye
calibration process. A red dot appears on the screen and moves to all areas of the screen to calibrate the participant’s eyes. A shopping list is given to the participant to write down his or her choices while shopping to compare the purchase decision data from the physical CUshop study. Once they are finished shopping, the participant is instructed to exit the virtual store, which ends the eye-tracking portion of the experiment. A post-survey questionnaire partly consisting of a modified Witmer-Singer survey asks questions pertaining to presence in the environment. The second half of the survey asks about video game and computer experience. This survey acquired general feedback to better understand each participant’s experience and if they had thoughts on what could have been better. Once the survey is completed, the participant exited the room and the study is finished.

**Statistical Analysis**

Demographical information was collected through Survey Monkey, an online survey platform, and exported into Excel for analysis and graphing. Eye tracking metrics investigated in this study were *Time to First Fixation* and *Total Fixation Duration*. These two metrics can be described as total amount of time (seconds) it takes someone to first look at a package (TTFF) and total duration time looked at the target package (TFD). The raw eye tracking data was collected and exported using Tobii studio software. In order to gather data on targeted products in the virtual environment, AOI’s (Areas of Interest) were drawn over each product to calculate TFD and TTFF (Figure 32). Since the fixations were captured using a web camera framed around the screen for playback of the
virtual experiment, this required hand coding each participant’s shopping video which meant adjusting each of the eleven barbecue sauce AOI’s frame by frame to ensure all eye tracking data was captured. Additionally, shopping times were recorded by re-watching each participant’s video, which required marking his or her entry and exit time to take the difference and acquire a total. After analyzing the data exported from Tobii studio, there was inconsistency among the TTFF times for the virtual environment because of how the AOI’s had to be hand coded frame by frame. This required going back to each participants recording and figuring out when the first AOI became active and subtracting that value from the TTFF values for each barbecue sauce to offset the times sufficiently for viable comparison. Once the data were comparable, both the physical and virtual CUshop TTFF/TFD mean values were compared by plotting their average and standard error values to begin to see what kind of data trends would be expected from eleven t-tests using SAS Plus. The data were then reorganized to run pair-wise t-tests on each of the eleven barbecue sauces based on a comparison of physical (A) vs. virtual (B). Results were plotted and evaluated based on the p-values gathered for each comparison. A t-test was also conducted on the physical and virtual shopping times to gather further evidence on their comparison. A p-value less than or equal to 0.05 was considered to be significant at a 95% confidence interval.
Results and Discussion

The concluding of this experiment yielded a total of 83 participants between the Physical CUshop study and Virtual CUshop study. Post-reviewing of participant eye tracking videos from each study revealed that five participants had poor eye tracking percentages. This meant that no fixations were reported from the data, resulting in removing those five from the virtual study. After the removal of 5 participants, 44 remained from the virtual store and 34 from the physical store experiment. Additional demographic data can be accessed in Appendix B.

Eleven t-tests were conducted between each barbecue sauce in the physical and virtual CUshop to represent the following significant findings. A 95% confidence interval (<.05 alpha) was used to report a difference. A combination of Excel and SAS Plus were used to calculate the T-test findings. Findings were calculated and visualized using a combination of Excel, SAS Plus and Survey Monkey.

Survey Results

Each participant was given a pre-survey and post-survey, gathering information regarding demographics and other questions related to the study. All users were given a unique participant number to ensure confidentiality. The following information outlines the results of these questions:
Pre-Survey:

- Biological gender demographic:
  
  o The Physical CUshop study resulted in 71% female and 29% male.
  
  o The Virtual CUshop study resulted in 50% male and 50% female.

![Pie charts showing gender distribution for Physical and Virtual CUshop](image)

*Figure 24: Gender between Physical and Virtual CUshop*

- Age Distribution:
  
  o **Physical CUshop:**
    
    - Of the 34 participants surveyed, 7% were 18-20, 31% were 21-29, 20% were 30-39, 16% were 40-49, 20% were 50-59, 4% were 60-64, and 1% were 65 or older.
Virtual CUshop:
- Of the 44 participants surveyed, in the virtual CUshop 20% were 18-20, 46% were 21-29, 17% were 30-39, 4% were 40-49, 9% were 50-59 4% were 60-64 and 0% were 65 or older.
• Shopping Results:
  
  o Are you the primary shopper for your household?
    
    ▪ Physical CUshop: 65% answered yes, 7% answered no, and 28% answered sometimes.
    
    ▪ Virtual CUshop: 62% answered yes, 22% answered no, and 16% answered sometimes.
  
  o How often do you shop for barbecue sauce?
    
    ▪ Physical CUshop: 2.6% answered once a week or more, 29.4% answered once every 2-3 weeks, 42.8% answered once a month, 20.5% answered once every 2-3 months and 4.4% answered less than every 3 months.
    
    ▪ Virtual CUshop: 0% answered once a week or more, 18% answered once every 2-3 weeks, 28% answered once a month, 24% answered once every 2-3 months and 30% answered less than every 3 months.
Post-Survey:

Q2 Which brands of barbecue sauce do you commonly use?

Figure 27: Physical CUshop Survey Response, What barbecue sauce do you commonly use?

Q2 Which brands of barbecue sauce do you commonly use?

Answered: 48 Skipped: 0

Figure 28: Virtual CUshop Survey Response, What barbecue sauce do you commonly use?
Figure 28 and 29 represent the data gathered from the question, “What barbecue sauce do you commonly use?” Rankings for both stores were similarly distributed with BBQ Sauce 1 being chosen most, followed by BBQ Sauce 2, BBQ Sauce 3 and BBQ Sauce 4.

Virtual CUshop Questions

To understand if users abilities to navigate and shop virtually were affected by video games, a series of questions were added to measure video game experience and general computer competency. Below are the results:

- How many hours do you spend using 3D programs and/or computer games during a month?

![Bar chart showing the distribution of hours spent on 3D programs and computer games.]

**Figure 29: Virtual CUshop Survey Response, How many hours do you spend using 3D programs and/or computer games during a month?**
Figure 29 shows how familiar participants were with navigating 3D environments on a computer. Of the participants, 53% used 3D programs or computer games 0-4 hours, 16% spent 5-9 hours, 10% spent 10-14 hours, 6% spent 15-20 hours and 14% spent 20+ hours.

- Do you have experience with video games?

![Bar Chart: Question 5 (Q5) - Do you have any experience with video games?](image)

Figure 30: Virtual CUshop Survey Response, Do you have any experience with video games?

Figure 30 shows that 80% of participants have experience with video games while 20% say they have no experience with video games. Two additional survey questions were asked specifying what genres of games and consoles were use to understand if certain genres affected users ability. Of the 49 participants in virtual study, 41 said they played video games requiring them to respond to the following two questions:

- If so, please check off the types of video games you have played?
Figure 31: Virtual CUshop Survey Response, If so, please check off the types of video games you have played?

- What video game platforms have you had experience with?

Figure 32: Virtual CUshop Survey Response, What video game platforms have you had experience with?

Presence Questionnaire

Concluding the post-survey, each participant was asked to answer five chosen questions from a modified Witmer-Singer survey ranked on a 7-point Likert scale. (Table
2) The four categories noted were Involvement, Immersion, Sensory Fidelity and Interface Quality. Mean totals showed no significant similarities between Virtual and Physical CUshop. A graph plotted of mean, standard deviation and standard error show slight significance between Immersion, Sensory Fidelity and Interface Quality. (Figure 33) These findings are consistent with previous studies that found it difficult to compare physical and virtual environments based on levels of presence. (C. Lee et al., 2013; Slater & Lotto, 2009)

Further analysis of the presence scores was considered from a virtual CUshop pilot study conducted at Pack Expo Las Vegas. The scores from this virtual study compared to the Las Vegas pilot study helps understand whether the addition of navigation controls increases a users level of presence. Scores from the Virtual CUshop study indicated a higher level of presence than those recorded from the Las Vegas study. (Table 2)

<table>
<thead>
<tr>
<th>Question</th>
<th>Physical CUshop</th>
<th>Virtual CUshop</th>
<th>Vegas Virtual CUshop</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Involvement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My interactions with the shelving environment seemed natural.</td>
<td>5.81</td>
<td>5.15</td>
<td>4.89</td>
</tr>
<tr>
<td><strong>Immersion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All my senses were completely engaged.</td>
<td>5.13</td>
<td>4.73</td>
<td>4.15</td>
</tr>
<tr>
<td><strong>Sensory Fidelity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My experience with the shelving system seemed consistent with my real-world experience.</td>
<td>5.66</td>
<td>5.29</td>
<td>4.75</td>
</tr>
<tr>
<td><strong>Interface Quality</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Visual display quality interfered or distracted me from completing my task.</em></td>
<td>4.44</td>
<td>4.56</td>
<td>4.00</td>
</tr>
<tr>
<td>I was able to concentrate on the search task and not on devices use to perform the task.</td>
<td>6.25</td>
<td>5.73</td>
<td>5.20</td>
</tr>
<tr>
<td>Group Means(Means of Means)</td>
<td>5.35</td>
<td>5.15</td>
<td>4.60</td>
</tr>
</tbody>
</table>

*Table 2: Mean responses from the modified Witmer-Singer presence questionnaire using a 7-point Likert scale with 7 indicating highest level of agreement and 1 indicating lowest level of agreement. Statement 4 is negative in nature and values were transformed when calculating means for score consistency.*
Discussion of Survey Responses

Demographic results showed a significant difference between the percentage of males and females that participated within both studies, which could cause a comparison between the physical and virtual environments to be difficult. This difference was further analyzed to determine whether or not males and females shop BBQ sauce differently. A comparison of participants BBQ sauce purchase decisions (the BBQ sauce they indicated they would have bought) was conducted between the physical study and the virtual study, isolating both genders. Results from graphing this data indicated that there was not enough evidence based on purchase decision to show that males and females shop
barbecue sauce differently. This evidence doesn’t necessarily prove that males and females in fact shop similarly or differently though. It means that not enough subjects participated, thus a definitive calculation cannot be computed at this time.

Initially, it was expected that the larger percent of younger (18-29) participants (Figure 25 and 26) in the virtual study would have performed better due to being more familiar with computer use than the older demographic (30 +). However, when asked “How many hours do you spend using 3D programs and/or computer games during a month?” 53% indicated that they only use 3D programs 0-4 hours per month. (Figure 29) In contrast, the results from the survey question regarding video game usage (Figure 30) indicate that 80% of participants played video games. This shows that video game usage may have been a factor in indicating that people who use video games are better at navigating virtual environments. However, this could partly be due to the fact that 66% of the participants were under the age of 30. To further understand how video game usage affected the study, those participants who do play games were asked to classify what genres they played.(Figure 31) First person shooters were the most commonly played genre (68%) which is most relatable to the camera positioning used in the virtual study, indicating that a vast majority of participants were familiar with navigating from that perspective.

Presence data also indicates differences in how navigation affected the study. Scores were compared from the Virtual CUshop study where the participant was able to freely navigate the environment to the Vegas CUshop VR Pilot study where the user watched a pre-recorded video. Presence scores were higher in the Virtual CUshop study
indicating that having the ability to freely navigate a virtual store can improve users' ability to shop naturally in a grocery store context.

**Eye Tracking Metrics**

Eye-tracking measurements were calculated in Tobii Studio using the Tobii fixation filter between both studies. The fixation filter calculates gaze point using the Average method for stimuli recorded with the Tobii glasses (physical store) and the Tobii X2-60 mounted eye tracker (virtual store). Data was exported for each of the 11 AOI’s (Areas of Interest) created from each search task. (Figure 34) Below is a graphical representation of Total Fixation Duration (TFD) performance in the physical and virtual CUnshop experiment to illustrate how eye tracking metrics were gathered using AOIs.
Figure 34 illustrates through the graphical color values of a heat map, how eye tracking metrics were quantified using AOI’s. Barbecue sauces are ranked based on TFD performance in the Virtual CUshop.

Total Fixation Duration and Time to First Fixation were investigated to determine if there were any significant differences amongst the performance of each of the eleven
individuals barbecue sauces in the virtual environment. These metrics are calculated and results are discussed below.

**TFD**

Total Fixation duration (TFD) defines the total amount of time, in seconds, that a participant fixates on a given Area of Interest (AOI). Eleven t-test’s were conducted on each barbecue sauce from the physical and virtual CUshop experiments.

- **Physical and Virtual CUshop TFD**

<table>
<thead>
<tr>
<th>BBQ Sauce</th>
<th>Physical CUshop</th>
<th>Virtual CUshop</th>
<th>P-Value</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBQ Sauce 1</td>
<td>2.86</td>
<td>4.22</td>
<td>0.1077</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 4</td>
<td>1.47</td>
<td>1.51</td>
<td>0.9931</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 7</td>
<td>1.39</td>
<td>0.90</td>
<td>0.2984</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 6</td>
<td>1.15</td>
<td>1.25</td>
<td>0.8280</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 8</td>
<td>1.03</td>
<td>2.43</td>
<td>0.0452</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 5</td>
<td>0.76</td>
<td>0.71</td>
<td>0.8457</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 9</td>
<td>0.60</td>
<td>0.70</td>
<td>0.6942</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 10</td>
<td>0.52</td>
<td>0.81</td>
<td>0.2531</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 11</td>
<td>0.36</td>
<td>0.82</td>
<td>0.0203</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 3</td>
<td>0.35</td>
<td>0.72</td>
<td>0.0096</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 2</td>
<td>0.33</td>
<td>0.39</td>
<td>0.5047</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Total Fixation Duration mean averages (in seconds) and p-values based on t-test’s. Green color indicates no difference and red indicates a difference in p-value.
<table>
<thead>
<tr>
<th>TFD Ranking</th>
<th>Real</th>
<th>Virtual</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBQ Sauce 1</td>
<td>BBQ Sauce 1</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 4</td>
<td>BBQ Sauce 8</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 7</td>
<td>BBQ Sauce 4</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 6</td>
<td>BBQ Sauce 6</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 8</td>
<td>BBQ Sauce 7</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 5</td>
<td>BBQ Sauce 11</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 9</td>
<td>BBQ Sauce 10</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 10</td>
<td>BBQ Sauce 3</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 11</td>
<td>BBQ Sauce 5</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 3</td>
<td>BBQ Sauce 9</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 2</td>
<td>BBQ Sauce 2</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: TFD Ranking of all eleven barbecue sauces based on the Physical and Virtual CUshop average TFD times (seconds)

Figure 35: Physical vs. Virtual CUshop TFD Averages plotted with standard error.
Based on the p-values calculated from each TFD barbecue sauce comparison, all but three (BBQ Sauce 8, BBQ Sauce 11 and BBQ Sauce 3) showed no significant difference between the real and virtual environments (p-value>.05) (Table 2). Differences are graphically represented in Table 3 with red indicating significant difference and green indicating no difference. Therefore, there is not enough evidence to determine a difference among the real and virtual experiments based on TFD metrics.

**TFD Discussion**

Total Fixation Duration results indicate that the virtual environment yielded comparable results and showed similarities to the physical environment. All but three of the barbecue sauces performed equally in both environments. BBQ Sauce 8 Barbecue sauce reported a p-value of 0.0452, which is very close to the alpha value of .05 used to determine significance. The average mean comparisons graphed above show that BBQ Sauce 8 mean value indicate a significant difference in TFD. (Figure 35) Table 3 shows that BBQ Sauce 8 remained in the top five for both Physical and Virtual which shows consistency between both environments. Even though there were few differences between the real and virtual, it is worth noting the slight variations between the performance ranking shown in Table 4. The top five barbecue sauces remained consistent between physical and virtual, but order of ranking changed between the two environments.
**TTFF**

Time to First Fixation (TTFF) indicates the amount of time it takes a person to first fixate on an Area of Interest (AOI). Eleven t-tests were conducted comparing physical vs. virtual barbecue sauces to determine if there were any significant differences. Below are the results:

- **Physical and Virtual CUshop TTFF**

<table>
<thead>
<tr>
<th></th>
<th>Physical CUshop</th>
<th>Virtual CUshop</th>
<th>P-Value</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBQ Sauce 1</td>
<td>2.02</td>
<td>3.56</td>
<td>0.1539</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 4</td>
<td>3.69</td>
<td>4.85</td>
<td>0.3496</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 2</td>
<td>4.48</td>
<td>7.72</td>
<td>0.0997</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 5</td>
<td>4.98</td>
<td>4.68</td>
<td>0.8887</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 6</td>
<td>5.96</td>
<td>9.88</td>
<td>0.1091</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 7</td>
<td>6.33</td>
<td>9.68</td>
<td>0.1137</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 8</td>
<td>6.59</td>
<td>7.34</td>
<td>0.7442</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 9</td>
<td>6.64</td>
<td>7.89</td>
<td>0.4311</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 10</td>
<td>7.44</td>
<td>10.06</td>
<td>0.3200</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 3</td>
<td>8.23</td>
<td>7.78</td>
<td>0.8107</td>
<td></td>
</tr>
<tr>
<td>BBQ Sauce 11</td>
<td>9.06</td>
<td>7.99</td>
<td>0.6799</td>
<td></td>
</tr>
</tbody>
</table>

*Table 5: Time to First Fixation (TTFF) mean averages and p-values based on t-test’s. Green color indicates no difference and red indicates a difference in p-value.*
<table>
<thead>
<tr>
<th>TTFF Ranking</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Real BBQ Sauce 1</td>
<td>Virtual BBQ Sauce 1</td>
</tr>
<tr>
<td>Real BBQ Sauce 4</td>
<td>Virtual BBQ Sauce 5</td>
</tr>
<tr>
<td>Real BBQ Sauce 2</td>
<td>Virtual BBQ Sauce 4</td>
</tr>
<tr>
<td>Real BBQ Sauce 5</td>
<td>Virtual BBQ Sauce 8</td>
</tr>
<tr>
<td>Real BBQ Sauce 6</td>
<td>Virtual BBQ Sauce 2</td>
</tr>
<tr>
<td>Real BBQ Sauce 7</td>
<td>Virtual BBQ Sauce 3</td>
</tr>
<tr>
<td>Real BBQ Sauce 8</td>
<td>Virtual BBQ Sauce 9</td>
</tr>
<tr>
<td>Real BBQ Sauce 9</td>
<td>Virtual BBQ Sauce 11</td>
</tr>
<tr>
<td>Real BBQ Sauce 10</td>
<td>Virtual BBQ Sauce 7</td>
</tr>
<tr>
<td>Real BBQ Sauce 3</td>
<td>Virtual BBQ Sauce 6</td>
</tr>
<tr>
<td>Real BBQ Sauce 11</td>
<td>Virtual BBQ Sauce 10</td>
</tr>
</tbody>
</table>

Table 6: TTFF Ranking of the eleven-barbecue sauces based on the Physical and Virtual CUshop average TTFF times in seconds.

Figure 36: Physical CUshop TTFF Averages plotted and ranked with standard error
Based on the p-values calculated from eleven t-tests between the physical and virtual CUshop TTFF values, there is not enough evidence to report a difference in the two environments. (Table 5) An ordered ranking was conducted to compare performance rankings based on quickest TTFF average (Table 6). This indicates that within the TTFF data, participants in the physical CUshop did not look at the barbecue sauces in the same order as they did in the virtual.

TTFF Discussion

Figure 36 and 37 break down how the rankings of physical and virtual compared based on average times. Connected white dots indicate that the marked sauces TTFF times were not significantly different based on standard error. Results show that although...
they ranked differently, it was not significant enough to say one actually scored higher than another in TTFF scores. In the physical store, BBQ Sauce 1 clearly had the lowest TTFF. The graph indicates that BBQ Sauce 5 (2nd place) and BBQ Sauce 4 (3rd place) were nearly the same based on their mean values. Furthermore, BBQ Sauce 8 (4th place) and BBQ Sauce 2 (5th place) are very close to identical as well. Figure 37 illustrates that the virtual store TTFF rankings are almost indistinguishable, indicating that a ranking of the sauces between virtual and physical are so close that it’s not possible to say which significantly performed better.

As mentioned in *Statistical Analysis* section of the Methodology, the physical store TTFF values were calculated automatically in the software while the values for Virtual CUshop had to be offset by a certain interval of time based on each participant. This required going back in to each participants recording and subtracting the time it took for the first AOI to activate from the TTFF value exported from Tobii Studio. This inconsistency makes it difficult to conclude that the values are indeed useful being that there was no exact way to calculate the interval necessary to offset each value.

*Noted Eye Tracking Fixations*

A Noted analysis for the eye tracking experiment analyzes TTFF and TFD fixations to understand how they vary based on the amount of actual participants that did fixate and did not fixate on the barbecue sauce. Since not every participant fixates on each of the eleven barbecue sauces, their data has to be removed from the t-test, subsequently changing the N value for many of the tests. The noted values help show
transparency within the statistical analysis to understand how the fixation’s varied. Figure 38 and 39 show the noted value comparisons for physical CUshop and virtual CUshop based on gender. The physical environment garnered a higher number of fixations for the female population and a significantly lower number for males. Figure 38 shows that males in the virtual store consistently fixated more often on the eleven barbecue sauces than women.

![Noted Physical CUshop Fixation Percentages](image)

*Figure 38: Noted fixations on all barbecue sauces in Physical CUshop based on gender.*
**Noted Virtual CUshop Fixation Percentages**

![Noted Virtual CUshop Fixation Percentages](image)

*Figure 39: Noted fixations on all barbecue sauces in Virtual CUshop based on gender*

**Noted Fixations Discussion**

Based on the data collected, it shows that the difference in demographics may have caused variations in the percentage of males and females that fixated on each barbecue sauce. Figure 38 shows that within the top 5 sauces, women tended to fixate slightly higher percentages than the men. Although, it is difficult to determine a significant difference in the percentage of fixations made between the two genders in the physical store. Comparing the male and female values based on performance ranking shows that the top 5 between genders were consistent with the one exception, BBQ Sauce 11. Figure 39 indicates the opposite result between male and female. Both genders in the virtual environment generally had a high percentage of fixations across all eleven-barbecue sauces. This could be attributed to the fact that participants took more time to
shop in the virtual CUshop which allowed them to spend more time scanning the shelves and making their purchase decision.

**Shopping Times**

Total shopping times between physical and virtual CUshop were recorded by re-watching each participants recording. Times were calculated by taking the difference in time of when the participant enters and exits the main door of the store. A pairwise t-test was conducted to determine whether or not there was any significance between shopping times when both environments were compared. A p-value of less than 0.05 was considered to be significantly different. (p-value<.05)

![Physical Vs. Virtual Shopping Times](image)

*Figure 40: Average shopping times with standard error between physical and virtual CUshop*
Figure 40 illustrates average shopping times with standard error bars to indicate whether or not the environments were similar. Averages revealed a 28-second difference (14%) between the total times shown above, indicating that it takes people longer to shop in the virtual CUshop. Additionally, the error bars shown in figure 40 indicate a very close comparison of the total shopping times. A pairwise t-test comparing the physical and virtual total shopping times resulted in a p-value of .0661. A p-value >.05 indicates that there is not enough evidence to show a difference between total shopping times within either environment.
Shopping Time Discussion

Filtering shopping times by age shows how different age groups total-times varied based on the demographics recorded. Figure 41 shows that people shop fairly similarly among all age groups in the physical store. In contrast, the virtual store times steadily rise based on age with ages 18-20 having an average time of 192.2 seconds and ages 60-64 having an average time of 372 seconds. This difference may not be as large based on standard error though. If the 60-64 age group were to have had more participants, the times might have adjusted to be similar to the 18-30 age group. Regardless, it is clear that participants generally take longer to shop in the virtual store.

Shopping times were further broken down to understand where the difference in physical and virtual exists based on Travel Time (time it takes the participant to locate shelf of target product) Searching time (time it takes participant to locate product and make purchase decision) and within the virtual store, Error (if the participant gets lost or runs in to a shelf, wall or product). Times were calculated by re-watching each participant’s video and marking intervals based on the 5 shopping list items, followed by taking an average of the total values for each for the physical and virtual store. The physical CUshop travel time break down showed that 58% of the time participants are searching on the shelf for the product while the remaining 42% is spent walking around the store. (Figure 42) In contrast, the virtual store participants spent 52% of the time searching, 44% travelling around the store and 4% error. (Figure 43) Results indicate that participants spend most of their time searching the shelves, followed by travelling around the store in both environments tested. This shows that although the stimuli may
look and be perceived differently between the physical and virtual, people still shop in a similar fashion.

**Figure 42**: Distribution of travel time and searching time taken from total shopping times in the physical CUshop

**Figure 43**: Distribution of travel time and searching time taken from total shopping times in the Virtual CUshop
**Shopping Data**

Participants were asked to choose five different products (BBQ sauce (target product), Sharp Cheddar Cheese, Hot Cocoa Mix, Travel-sized conditioner and Shortbread cookies) during their shopping experience in both studies. Purchase decisions of barbecue sauces were recorded and analyzed to better understand if purchase decision effects where they look at on the shelf between environments.

**Overall Brands Shopping Data Comparison**

*Figure 44: Comparison of purchase decision amongst all brands in the physical and virtual CUshop*
Purchase Decision Discussion

Purchase decision values indicated above in Figure 44 and 45 show a consistency among the barbecue sauces chosen between studies. The top five barbecue sauce purchase decisions illustrated in Figure 45 show that all but BBQ Sauce 4 were chosen in the same order between physical and virtual environments. Results indicate that differences in the environment didn’t seem to affect purchase decision behavior. The significant difference shown for BBQ Sauce 4 barbecue sauce in figure 45 could have been attributed to the limitations in replicating the aesthetics of the label. Certain aspects like metallic reflections and decorations on the bottle were more difficult to accurately reproduce in the virtual store and thus could have affected BBQ Sauce 4’s presence on the shelf. (Figure 46) This also caused the barbecue sauces to render slightly darker than in the physical store due to the lack of metallic reflection.
Post-experiment it was determined post-study that a larger participant pool would have been necessary in order to adequately gather enough information to find a significant difference, if any at all. A power study indicated around 445 total participants would have provided enough data to make stronger comparisons between the two studies.

Data collected from the two experiments support the decision to fail to reject the null hypothesis described in the methodology of this paper. Adjustments were made throughout the progress of this study due to limitations and recommendations from the committee. Due to limitations and recommendations from my committee certain changes were made throughout the study such as: replacing the 3D navigation device with a mouse and keyboard due to coding issues, using a 47” screen TV instead of a 3D TV, errors in fixation data and incentives to encourage an older demographic to participate in the virtual study. Given that this was the best-case scenario possible for current technology available, it is clear that further consideration should be taken in to hardware and overall setup for measuring consumer purchase habits in virtual environments.
Future studies should hope to provide a better, more immersive apparatus in which to conduct Virtual based experiments to better determine correlations between physical and virtual store purchasing habits.

**Conclusion**

The goal of this thesis was to compare attributes collected from eye tracking data such as TTFF, TFD, Purchase Decision and Shopping Times to determine any similarities between Physical CUshop and Virtual CUshop. It also tested how consumers make purchase decisions in physical and virtual contexts in regards to buying barbecue sauce. Eye tracking data in combination with travel time, presence scores and purchase decision were used to draw these conclusions and further decide how well this methodology of virtual eye tracking performed.

An ideal virtual store would score similarly on eye tracking metrics, purchase decision and shopping times when compared to the physical store. Presence scores showed very little deviation between the two studies. Comparing the Las Vegas study to the virtual CUshop showed increases in presence scores with the addition of self guided navigation. This indicates that the addition of navigation in the virtual store can provide increased subjective levels of presence.

After analyzing eye tracking metrics and travel times through 23 pairwise t-tests (11-TFD, 11-TTFF, 1-Travel time), no significant difference (p>.05) was determined among TTFF, TFD or shopping times. Indicating that there is not enough evidence to
prove that the physical and virtual stores are different in measuring consumer purchase behavior. However, the study did provide some conclusive findings. Based on average total shopping times, it was indicated that shoppers take 14% longer to navigate a virtual shopping environment than a physical store. Additionally, consumers spend a similar ratio of time searching and traveling around in both environments. Finally, although the fidelity of products varied between the physical and virtual store, consumers still chose the same brands in each study. This indicates that with further research, virtual stores can potentially be capable of measuring purchase decision when compared to a physical study of the same nature.

Many limitations existed during the process of this experiment between hardware, data and post processing. The value from this experiment was not in the outcomes of the data but rather in developing the method used to quantify them. This study succeeded in creating a strong methodology that should serve as a model for future studies to compare virtual and physical store consumer shopping behavior under more immersive conditions.

Limitations

This virtual study was conducted as a best-case scenario setup based on the current technology that was available for use at the Sonoco Institute at Clemson University. The participant pool from which we recruited tended to be in the college-age range, causing the demographic from the virtual CUshop study to have a different percentage of age ranges between studies. Efforts were made to recruit an older
demographic from outside of Clemson University but the availability of subjects was limited.

Sample sizes between studies varied in gender and age demographics. This was due to the lack of global diversity available from the participants that were accessible for the experiment. We realize that this makes the two studies very different, thus causing it to be hard to draw stark conclusions about the metrics gathered from the experiment. A consistent sample size between both studies in addition to having a controlled selection of participants would have helped negate this problem.

Hardware limitations existed throughout the experiment. Originally, the study intended to use a 3D TV to add a level of immersion, but certain restrictions led to a 46” LCD high definition TV instead. The navigation device chosen for the experiment was a 3D mouse that would allow the participant to navigate the store through one device. A single device would have been more intuitive and easier to learn, but compatibility issues prevented the mouse from working with Esko Studio Store Visualizer™ resulting in using a keyboard and mouse as a replacement.

Eye tracking metrics between studies varied with how the data was captured for the physical study and the virtual study. Since the eye-tracking glasses could not be used for both studies, the mounted eye tracker (Tobii X2-60) was chosen for the virtual experiment. The Tobii X2-60 tracks at a rate of 60hz compared to the Tobii™ glasses, which record at a rate of 30hz. This means that more data was possibly captured with the Tobii X2-60 and could have had an impact on the fixations recorded. However, it is unclear as to whether this affected the data between the two studies.
Other issues that were recorded in the experiment were problems with people moving away from the mounted eye tracker. Participants would tend to lean in closer to see products on the screen or bend down to write down the product they were purchasing which resulted in errors within their recording. Screen fidelity also became an issue causing users to lean in to the screen because the camera in the environment was at a fixed height. This prevents you from being able to lean forward or down to look at products on the shelves more closely. This factor caused several participants to comment in the post survey that “labels on smaller packages were difficult to read”, “there should be a zoom function” and “I couldn’t crouch down to look at packages more closely” which further reiterates the issues caused by screen resolution and limitations of the program used for the environment.

Furthermore, participants suggested that being able to interact and look at the packages from several angles would have helped them better decide what to purchase. Other participants mentioned that they felt distracted by the experiment room due to not having peripherals in their vision and that it made them feel like it was harder to find products. Additionally, several suggestions were made to include aisle markers for easier location of products. Given the best-case scenario possible from the current technology available for this study, it is clear that further consideration should be taken in to hardware and overall setup for measuring consumer purchase habits in virtual environments.
Recommendations

There are many possibilities for expanding research in eye tracking consumer-shopping habits within virtual environments. Several adjustments could have been made to allow this study to be more immersive and accurate in testing participants shopping habits.

Using more immersive hardware is the first step to getting more finite results when testing in virtual environments. Within the past two years the virtual reality (VR) world has vastly expanded with the introduction of hardware like the Oculus Rift headset. This stereoscopic VR headset completely encompasses a users vision and provides a range of peripheral vision allowing them to able to look around an environment using only their head. Many other devices are being designed to work with the Oculus such as Leap Motion and the Omni treadmill, two devices, one that allows you to use your legs to walk and the other to use your hands to pick up objects. Adding in small abilities in to a study would not only cause someone to feel immersed when shopping but also make them have more natural responses to stimuli.

When choosing product to use as target items it is important to distribute the planogram effectively to ensure that certain products don’t score higher due to more shelf presence. Additionally, spending more time in refining small details such as lighting, reflection and resolution of the imagery placed on products could help participants shop easier when moving throughout virtual environments. Lastly, it would be beneficial to include an incentive for encouraging a varied demographic to participate and to incorporate more participants in the study.
APPENDICES
Appendix A

Survey Questions

Pre-survey

*1. What is your participant number?

*2. What is your biological sex?
   - Female
   - Male

*3. How old are you?
   - 18-20
   - 21-29
   - 30-39
   - 40-49
   - 50-59
   - 60-64
   - 65 or older

*4. Do you have any children?
   - Yes
   - No

*5. What is the highest level of education you have completed?
   - Less than high school degree
   - High school degree or equivalent (GED)
   - Some college but no degree
   - Associate degree
   - Bachelor degree
   - Graduate degree
   - Higher than graduate degree
Figure A-1: Questions 1-5 of Pre-survey.

**6. Which of the following best describes your employment status? (Check all that apply)**
- [ ] Employed, working 1-39 hours per week
- [ ] Employed, working more than 40 hours per week
- [ ] Not employed, looking for work
- [ ] Not employed, not looking for work
- [ ] Retired
- [ ] Disabled, not able to work

**7. What is your annual household income?**
- [ ] Less than $20,000
- [ ] $20,000 to $34,999
- [ ] $35,000 to $49,999
- [ ] $50,000 to $74,999
- [ ] $75,000 to $99,999
- [ ] $100,000 to $149,999
- [ ] $150,000 to $199,999
- [ ] $200,000 or more

**8. Are you the primary shopper for your household?**
- [ ] Yes
- [ ] No
- [ ] Sometimes

**9. How many people are in your household?**
- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5 or greater
Figure A-2: Questions 6-9 of Pre-survey.

**10. How often do you grill during the grilling season?**
- Multiple times a week
- Once a week
- Once every two weeks
- Once every month
- Very rarely
- Never

**11. How often do you shop for barbecue sauce?**
- Once a week or more
- Once every 2-3 weeks
- Once a month
- Once every 2-3 months
- Less than every 3 months

Figure A-3: Questions 10-11 of Pre-survey.
Post Survey

1. What is your participant number?

2. Which brands of barbecue sauce do you commonly use?
   - Sweeet Baby Rays
   - Weber
   - Kraft
   - KC Masterpiece

3. Rate on a scale of 1-5, how easy it was to find the Weber Barbecue packages on the shelf with 1 being very hard to find and 5 being very easy to find
   - 1
   - 2
   - 3
   - 4
   - 5

4. How many hours do you spend using 3D programs and/or computer games during a month?
   - 0-4
   - 5-9
   - 10-14
   - 15-20
   - 20+

5. Do you have any experience with video games?
   - Yes
   - No

6. If so, please check off the types of video games you have played?
   - First Person Shooters
   - Puzzles
   - Role Playing Games
   - Other (please specify)

Figure A-4: Questions 1-6 of Post-survey.
7. What video game platforms have you had experience with?
   - Wii
   - Xbox
   - Playstation
   - PC
   - Mobile Games

**8. My interactions with the shelving environment seemed natural. (1 indicating lowest level of agreement and 7 indicating highest level of agreement)**

**9. All my senses were completely engaged. (1 indicating lowest level of agreement and 7 indicating highest level of agreement)**

**10. My experience with the shelving system seemed consistent with my real-world experience. (1 indicating lowest level of agreement and 7 indicating highest level of agreement)**

**11. Visual display quality interfered or distracted me from completing my task. (1 indicating lowest level of agreement and 7 indicating highest level of agreement)**

**12. I was able to concentrate on the search task and not on devices use to perform the task. (1 indicating lowest level of agreement and 7 indicating highest level of agreement)**

13. Did any of the navigation devices (keyboard/mouse) make moving and looking around the store difficult?

14. Were there any visual aspects of the store that prevented you from shopping naturally?

15. Do you have any additional comments or feedback on your experience?

Figure A-5: Questions 7-15 of Post-survey.
Appendix B

Demographic and Survey Results: Physical CUshop

Do you have any children?

Figure B-1: Q4- Do you have any children?

Education

- Bachelor degree 32%
- Graduate degree 35%
- Some college but no degree 10%
- Associate degree 8%
- High school degree or equivalent (GED) 6%
- Higher than graduate degree 9%

Figure B-2: Q5- What is the highest level of education you have completed?
Figure B-3: Q6: What best describes your current employment status?

Figure B-4: Q7: What is your annual household income?
Figure B-5: Q8: Are you the primary shopper for your household?

**Primary Shopper**

- *Yes*: 64%
- *Sometimes*: 27%
- *No*: 7%
- *Did Not Answer*: 2%

Figure B-6: Q9: How many people live in your household?

**People in Household**

- *1*: 15%
- *2*: 40%
- *3*: 24%
- *4*: 18%
- *5 or greater*: 3%
- *Did Not Answer*: 2%
Figure B-7: Q10- How often do you grill during the grilling season?

**Grilling Frequency**

- Very rarely: 2%
- Never: 1%
- Once every month: 9%
- Once every two weeks: 27%
- Multiple times a week: 34%
- Once a week: 27%
- Once a week or more: 3%

Figure B-8: Q11: How often do you shop for barbecue sauce?

**BBQ Sauce Shopping Frequency**

- Once a month: 41%
- Once every 2-3 months: 20%
- Once every 2-3 weeks: 28%
- Did Not Answer: 4%
- Less than every 3 months: 4%
- Did Not Answer: 4%
Appendix C

Demographic and Survey Results: Virtual CUshop

Do you have any children?

Figure C-1: Q4- Do you have any children?

Education

- Bachelor degree: 38%
- Associate degree: 2%
- Some college but no degree: 28%
- Graduate degree: 14%
- Higher than graduate degree or equivalent (GED): 10%
- High school degree or equivalent (GED): 10%
- Higher than graduate degree: 8%
Figure C-2: Q5- What is the highest level of education you have completed?

![Pie chart showing employment status]

- Employed, working 1-39 hours per week: 48%
- Employed, working more than 40 hours per week: 23%
- Not employed, looking for work: 8%
- Not employed, not looking for work: 19%
- Retired: 2%

Figure C-3: Q6: What best describes your current employment status?

![Pie chart showing participant annual income]

- Less than $20,000: 48%
- $20,000 to $34,999: 12%
- $35,000 to $49,999: 10%
- $50,000 to $74,999: 8%
- $75,000 to $99,999: 4%
- $100,000 to $149,999: 10%
- $150,000 to $199,999: 4%
- $200,000 or more: 4%

Figure C-4: Q7: What is your annual household income?
Figure C-5: Q8: Are you the primary shopper for your household?

Primary Shopper

Yes 62%
Sometimes 16%
No 22%

Figure C-6: Q9: How many people live in your household?

People in Household

1 38%
2 12%
3 24%
4 16%
5 or greater 10%
Figure C-7: Q10- How often do you grill during the grilling season?

Grilling Frequency

Figure C-8: Q11: How often do you shop for barbecue sauce?

BBQ Sauce Shopping Frequency
Appendix D

Product Shelf Images

Figure D-1: Barbecue products from experiment

Figure D-2: Shortbread cookie products from experiment
Figure D-3: Hot Cocoa products from experiment

Figure D-4: Shampoo products from experiment
REFERENCES


