THE EFFECT OF DEPTH ON THE USEFUL FIELD OF VIEW

A Dissertation by

John Paul Plummer

Master of Arts, Wichita State University, 2015
Bachelor of Science, Virginia Polytechnic Institute and State University, 2011
Bachelor of Science, Virginia Polytechnic Institute and State University, 2011

Submitted to the Department of Psychology and the faculty of the Graduate School of Wichita State University in partial fulfilment of the requirements for the degree of Doctor of Philosophy

May 2019
THE EFFECT OF DEPTH ON THE USEFUL FIELD OF VIEW

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy, with a major in Psychology.

Rui Ni, Committee Member

Alex Chaparro, Committee Member

Barbara S. Chaparro, Committee Member

Joel M. Suss, Committee Member

Michael J. Jorgensen, Committee Member

Accepted for the Fairmount College of Liberal Arts and Sciences

Andrew Hippisley, Dean

Accepted for the Graduate School

Kerry Wilks, Interim Dean
DEDICATION

In honor of my mother and wife; in memory of my father
With willing hearts and skillful hands, the difficult we do at once; the impossible takes a bit longer.

-Motto of the U.S. Navy Seabees
ACKNOWLEDGMENTS

This dissertation is the culmination of many years of work, and during that time I have been fortunate to have been mentored and advised by some wonderful researchers at Wichita State. First, I would like to thank my advisor, Dr. Rui Ni. Throughout my many years of graduate education, he has given me guidance and encouragement, as well as the perspective of applying basic research to real-world settings. Dr. Barbara Chaparro and Dr. Alex Chaparro have both been sources of support and additional mentoring throughout my graduate career and have greatly influenced my understanding of how to apply human factors methods and principles. Dr. Joel Suss has provided excellent feedback for this project and in general has always been quick and eager to offer assistance. Finally, Dr. Michael Jorgensen has been a wonderful resource in my professional development by offering an engineering perspective to my work. Outside of the committee, I have been influenced by many former human factors faculty, in particular Dr. Joseph Keebler, Dr. Evan Palmer, and Dr. Amanda Smith. The decision to come to WSU has been more fruitful than I could have ever imagined.

My professional development has been more than just from this university, and I would not have gotten here had it not been for the excellent education and advising at Virginia Tech. Specifically, I would like to thanks Dr. Kurt Hoffman for introducing me to psychology, Dr. Scott Geller for showing me a passion for both research and applied psychology, and Dr. Jack Evans for encouraging me to broaden my horizons and take advantage of many professional opportunities. At the Virginia Tech Transportation Institute I received my first experience with real human factors work, and many of the researchers and staff there laid the firm foundation for what would soon become my career: Mr. Jeremy Sudweeks, Dr. Justin Morgan, and Dr. Justin Owens all helped foster my burgeoning interest in the field and it was Dr. Owens who suggested
I apply to Wichita State and introduced me to the concept of the useful field of view. Finally, I have learned an exceptional amount about the field from my internship experience, and I know I would not be the professional I am today without the guidance of my supervisors: Dr. Nichole Morris, Mrs. Katie Mowery, and Mr. John Winters.

The graduate school experience would not have been the same without wonderful and supportive colleagues. First, I would like to thank Bobby Nguyen and Rondell Burge, for acting as big brothers to me throughout their time in school and continuing that role as alumni. My cohort, Dustin Smith, Christina Sui, Duy Nguyen, and Brad Weber, has been a source of support and commiseration since the first day of graduate school. I know all those shared experiences—classes, group projects, conferences, and quals—formed a very unique bond between us all. The group of HF students has in many ways been a family to me and so I would like specifically to thank Mikki Phan, Jen Teves, Sarah Fouquet, Samantha Jansen, Andrew Miranda, David Libby, Jo Jardina, Neethan Siva, Jake Ellis, Tiffany Leverenz, and Anastasia Diamond for the various personal and professional support you all have given me. I would be remiss to not thank my wonderful research assistant, Shivani Nagrecha, for providing great feedback and support through my whole dissertation and Dr. Ryan Ringer for his help and advice throughout the process. I would also like to thank Judy Barnes and Marci Nelson for always being a welcoming presence and help throughout my entire graduate school career.

Finally, I would like to thank my wonderful family. My father, Stan, and mother, Cathy, have always been a source of support for all my educational goals. I’m certain that if it had not been for their example, I would not have had the discipline or motivation that grad school required. My wife, Christina, has been a constant rock of support through the triumphs and trials
that seem to be an inherent part of getting a PhD. I’m not sure that she knew what she was getting into, but I will be eternally grateful that she was an integral part of this journey.

This project was supported in part by a grant from the National Institute of General Medical Science GM113109 of the National Institutes of Health and by the Delano Maggard Jr, Research Grant from Wichita State University.
ABSTRACT

The useful field of view (UFOV) test is a popular measure of visual attention and is often applied to everyday tasks. One aspect of the real world that is not tested by this measure is the influence of depth on the allocation of attention. Previous research has shown that depth does influence some aspects of spatial attention, and capturing that effect in the UFOV test is important, as depth is inherent in real-world environments and tasks. One real-world task to which the UFOV has often been related is driving, and previous research has shown a strong relationship between UFOV performance and crash risk. Understanding the influence of depth may increase the ecological validly of this test, since, in the roadway environment, depth plays a very important role in successful navigation and hazard avoidance.

The goal of this dissertation is to evaluate the influence of depth on the UFOV and apply it to a measure of driving performance. Experiment 1 evaluated the effect of depth on the UFOV through three studies that varied how the depth of peripheral targets was displayed. The first tested the effect when the depth was known before the targets were displayed for divided and selective attention subtasks. The second evaluated this effect when depth was unknown before the targets were displayed for divided and selective attention subtasks. The third evaluated the cost of dividing attention in depth with modifications of the divided attention subtask. The results suggest that depth does have an effect on the allocation of attention in the UFOV test, particularly in the divided attention tasks. Since driving performance is generally the application of the UFOV, Experiment 2 evaluated the predictive power of the 3D UFOV test to a video-based hazard perception test. The results did not show a relationship between the two measures. Practical and theoretical implications of the role of depth on attention and how its application can be tested are discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>5</td>
</tr>
<tr>
<td>2.1 Visual Attention</td>
<td></td>
</tr>
<tr>
<td>2.1.1 The Useful Field of View</td>
<td>7</td>
</tr>
<tr>
<td>2.1.2 Perception in Depth and Stereopsis</td>
<td>11</td>
</tr>
<tr>
<td>2.1.3 Attention in Depth</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Vision and Driving</td>
<td>19</td>
</tr>
<tr>
<td>2.2.1 Problems with Crash Involvement</td>
<td>21</td>
</tr>
<tr>
<td>2.2.2 UFOV and Situation Awareness</td>
<td>23</td>
</tr>
<tr>
<td>2.2.3 Hazard Perception</td>
<td>28</td>
</tr>
<tr>
<td>2.2.4 Using Signal Detection Theory to Measure Hazard Perception</td>
<td>30</td>
</tr>
<tr>
<td>2.3 Present Study and Literature Summary</td>
<td>33</td>
</tr>
<tr>
<td>3. STEREOSCOPIC UFOV STIMULUS DEVELOPMENT</td>
<td>36</td>
</tr>
<tr>
<td>3.1 Presenting Depth through Binocular Disparity</td>
<td>36</td>
</tr>
<tr>
<td>3.1.1 Factors that Affect Stereopsis</td>
<td>37</td>
</tr>
<tr>
<td>3.1.2 Fusion Limits and the Perception of Depth from Disparity</td>
<td>40</td>
</tr>
<tr>
<td>3.1.3 Simulated Depth with Active Shutter Glasses</td>
<td>41</td>
</tr>
<tr>
<td>3.2 UFOV Development and Pilot Testing</td>
<td>41</td>
</tr>
<tr>
<td>3.2.1 First Pilot Test</td>
<td>41</td>
</tr>
<tr>
<td>3.2.2 Retinal Eccentricity as Dependent Measure</td>
<td>46</td>
</tr>
<tr>
<td>3.2.3 Depth Placeholders and Changing the Divided Attention Task</td>
<td>48</td>
</tr>
<tr>
<td>4. EXPERIMENT 1</td>
<td>51</td>
</tr>
<tr>
<td>4.1 Purpose</td>
<td>51</td>
</tr>
<tr>
<td>4.2 Experiment 1A</td>
<td>51</td>
</tr>
<tr>
<td>4.2.1 Method</td>
<td>53</td>
</tr>
<tr>
<td>4.2.2 Results</td>
<td>62</td>
</tr>
<tr>
<td>4.2.3 Discussion</td>
<td>65</td>
</tr>
<tr>
<td>4.3 Experiment 1B</td>
<td>67</td>
</tr>
<tr>
<td>4.3.1 Method</td>
<td>69</td>
</tr>
<tr>
<td>4.3.2 Results</td>
<td>72</td>
</tr>
<tr>
<td>4.3.3 Discussion</td>
<td>74</td>
</tr>
<tr>
<td>4.4 Experiment 1C</td>
<td>76</td>
</tr>
</tbody>
</table>
### TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1 Method</td>
<td>77</td>
</tr>
<tr>
<td>4.4.2 Results</td>
<td>82</td>
</tr>
<tr>
<td>4.4.3 Discussion</td>
<td>84</td>
</tr>
<tr>
<td>4.5 General Discussion</td>
<td>85</td>
</tr>
<tr>
<td>5. EXPERIMENT 2</td>
<td>89</td>
</tr>
<tr>
<td>5.1 Purpose and Hypotheses</td>
<td>89</td>
</tr>
<tr>
<td>5.2 Method</td>
<td>92</td>
</tr>
<tr>
<td>5.2.1 Participants</td>
<td>92</td>
</tr>
<tr>
<td>5.2.2 Design</td>
<td>93</td>
</tr>
<tr>
<td>5.2.3 Apparatus and Stimuli</td>
<td>93</td>
</tr>
<tr>
<td>5.2.4. Procedure</td>
<td>95</td>
</tr>
<tr>
<td>5.3 Results</td>
<td>96</td>
</tr>
<tr>
<td>5.3.1 Useful Field of View</td>
<td>96</td>
</tr>
<tr>
<td>5.3.2 Hazard Perception Scores</td>
<td>97</td>
</tr>
<tr>
<td>5.3.3 Driving Experience Predicting Hazard Perception</td>
<td>99</td>
</tr>
<tr>
<td>5.3.4 UFOV Predicting Hazard Perception</td>
<td>101</td>
</tr>
<tr>
<td>5.4 Discussion</td>
<td>101</td>
</tr>
<tr>
<td>5.4.1 Differences in City and Highway HP Scores</td>
<td>101</td>
</tr>
<tr>
<td>5.4.2 Driving Experience Predicted City Driving Response Lag</td>
<td>103</td>
</tr>
<tr>
<td>5.4.3 3D UFOV and HP Scores</td>
<td>104</td>
</tr>
<tr>
<td>5.4.4 Future Directions</td>
<td>108</td>
</tr>
<tr>
<td>6. CONCLUSIONS</td>
<td>112</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>114</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>129</td>
</tr>
<tr>
<td>A. Stereoscopic Discomfort Survey</td>
<td>130</td>
</tr>
<tr>
<td>B. Post-Trial Hazard Ratings</td>
<td>132</td>
</tr>
<tr>
<td>C. HP Post-Session Questionnaire</td>
<td>133</td>
</tr>
<tr>
<td>D. Fuzzy Signal Detection Calculations</td>
<td>135</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. An illustration of the detrimental effects from the reduction in UFOV size, where panel A represents no restrictions and panel D represents an extreme reduction (Ball &amp; Owsley, 1993).</td>
<td>.......................................................... 10</td>
</tr>
<tr>
<td>2. Major source of depth information, adapted from Sekuler and Blake (1994).</td>
<td>11</td>
</tr>
<tr>
<td>3. A top-down view of binocular vision from Ogle (1952a). The area of single vision is where the objects are fused, and occur around the horopter.</td>
<td>.......................................................... 13</td>
</tr>
<tr>
<td>4. Techniques for image separation in stereoscopic displays (Banks et al., 2016).</td>
<td>15</td>
</tr>
<tr>
<td>5. The stimulus from Andersen et al. (2011). Participants were required to follow the lead vehicle and respond to changes in the low arrays above the road. The depth of the arrays had an effect in addition to the visual angle from the lead vehicle.</td>
<td>.......................................................... 18</td>
</tr>
<tr>
<td>6. Screenshot from real-world UFOV simulators; Gaspar et al. (2016) is on the right side Rogé et al. (2004) is on the left. The visual angle of peripheral targets from the lead vehicle was assessed, but not an effect of the depth of the targets.</td>
<td>.......................................................... 19</td>
</tr>
<tr>
<td>7. The information-processing model of SA (Endsley, 1995c).</td>
<td>.......................................................... 26</td>
</tr>
<tr>
<td>8. The sequence of the UFOV test.</td>
<td>.......................................................... 43</td>
</tr>
<tr>
<td>9. Stimulus display screen for divided (left) and selective (right) attention subtests.</td>
<td>.......................................................... 44</td>
</tr>
<tr>
<td>10. The depth planes and background used in a fixation screen. In this example, the depth is in the “same” plane. For ease in viewing, subsequent diagrams will omit the background grid.</td>
<td>45</td>
</tr>
<tr>
<td>11. The fixation screen showing the array of annuli around the possible peripheral target locations.</td>
<td>.......................................................... 48</td>
</tr>
<tr>
<td>12. Example of the divided attention task where a peripheral target is rotated 90° from the fixation screen (left) to the stimulus screen (right).</td>
<td>.......................................................... 50</td>
</tr>
<tr>
<td>13. The different depths used in the experiment, relative to the central target. The two possible peripheral targets (5° and 15°) are displayed either closer to the observer (A), same as the central target (B), and farther from the observer (C).</td>
<td>.......................................................... 55</td>
</tr>
<tr>
<td>14. The Gabor patches used for peripheral stimuli. The target Gabor patch was always horizontal and the distractor Gabor patches were always vertical.</td>
<td>.......................................................... 56</td>
</tr>
</tbody>
</table>
LIST OF FIGURES (continued)

Figure  
Page

15. Trial sequence for the divided attention test. The central target is “E” and the peripheral target is “1” (top left). .................................................................58

16. Trial sequence for the selective attention test. The central target is “E” and the peripheral target is “1” (top left). .................................................................59

17. The depth presented in both the fixation and the stimulus display. The top of the figure is the divided attention subtask and the bottom is the selective attention subtask. The peripheral target appears in the top left in each example. .........................................................60

18. The main effect of retinal eccentricity on display duration in the divided attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error. .................................................................63

19. The main effect of depth on display duration in the divided attention task. Performance was worse only when peripheral targets were displayed behind the central target. Error bars represent standard error. .................................................................63

20. The main effect of retinal eccentricity on display duration in the selective attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error. .................................................................64

21. The main effect of depth on display duration in the selective attention task. Performance was worse only when peripheral targets were displayed behind the central target. Error bars represent standard error. .................................................................65

22. The depth presented in both the fixation and the stimulus display. The top is the divided attention subtask and the bottom is the selective attention subtask. .........................................................71

23. The main effect of retinal eccentricity on display duration in the divided attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error. .................................................................72

24. The main effect of depth on display duration in the divided attention task. Performance was worse when peripheral targets were displayed at a different depth than the central target. Error bars represent standard error. .................................................................73

25. The main effect of retinal eccentricity on display duration in the selective attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error. .................................................................74
### LIST OF FIGURES (continued)

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>26. The fixation (left) and stimulus (right) display for the peripheral-focused attention subtest. The central target is “E” and the peripheral target appears in the top left.</td>
<td>79</td>
</tr>
<tr>
<td>27. The fixation (left) and stimulus (right) display for the divided attention subtest. The central target is “E” and the peripheral target is top left.</td>
<td>80</td>
</tr>
<tr>
<td>28. The fixation (left) and stimulus (right) display for the divided attention subtest. The target appears in the top left.</td>
<td>81</td>
</tr>
<tr>
<td>29. The effect of depth on divided attention. Near and far conditions had significantly higher thresholds than the same depth condition. Error bars represent standard error.</td>
<td>82</td>
</tr>
<tr>
<td>30. The effect of depth on peripheral focused attention. The threshold for the near depth condition was significantly higher than the same depth condition, and the far depth condition was significantly higher than both the near and the same condition. Error bars represent standard error.</td>
<td>83</td>
</tr>
<tr>
<td>32. The effect of depth on the cost of dividing attention. Near and far conditions had significantly higher thresholds than the same depth condition. Error bars represent standard error.</td>
<td>84</td>
</tr>
<tr>
<td>33. Screenshot of a city environment from the HP test. The top shows a highway environment and the bottom shows a city environment.</td>
<td>94</td>
</tr>
<tr>
<td>34. The effect of depth on the UFOV.</td>
<td>97</td>
</tr>
<tr>
<td>35. Effect of environment on response lag. Error bars represent standard error.</td>
<td>99</td>
</tr>
<tr>
<td>37. Relationship between driving experience and response buffer. The dotted line represents the regression line.</td>
<td>100</td>
</tr>
<tr>
<td>38. Comparing the 3D UFOV (at the same depth plane) with available normative data. Error bars represent standard deviation (Bennett et al., 2009; Edwards et al., 2006).</td>
<td>105</td>
</tr>
<tr>
<td>39. Distribution of responses for the experts and participants.</td>
<td>108</td>
</tr>
</tbody>
</table>
# LIST OF TABLE

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Significant (p &lt; .05) Pearson Correlations (r) between Divided and Selective Attention Subtests of the UFOV and Driving Measures on the DriveSim (adapted from Chaparro et al., 1999).</td>
<td>27</td>
</tr>
<tr>
<td>2. An Illustration of the Four Possible Outcomes of a Traditional SDT Task</td>
<td>31</td>
</tr>
<tr>
<td>3. Possible Outcomes from Traditional (top) and Fuzzy (bottom) SDT (adapted from Parasuraman et al., 2000).</td>
<td>32</td>
</tr>
<tr>
<td>4. Factors Affecting Stereoacuity (adapted from Boff &amp; Lincoln, 1988)</td>
<td>38</td>
</tr>
<tr>
<td>5. Summary of Hypotheses for Experiment 1A.</td>
<td>53</td>
</tr>
<tr>
<td>6. Summary of Hypotheses for Experiment 1B.</td>
<td>69</td>
</tr>
<tr>
<td>7. Summary of Hypotheses for Experiment 1C.</td>
<td>77</td>
</tr>
<tr>
<td>8. Summary of Hypotheses for Experiment 2.</td>
<td>92</td>
</tr>
<tr>
<td>9. Pearson Correlation Coefficients between Driving Experience and FSDT Measures</td>
<td>101</td>
</tr>
<tr>
<td>10. Pearson Correlation Coefficient and Significance values for the UFOV Thresholds and HP scores</td>
<td>101</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS AND NOMENCLATURE

° Degrees subtended of a visual angle
2D Two Dimensions
3D Three Dimensional
arcmin Minutes of an arc subtended of a visual angle
arcsec Seconds of an arc subtended of a visual angle
cd/m² Candelas per square meter
cm Centimeter
DBQ Driving Behavior Questionnaire
FSDT Fuzzy Signal Detection Theory
HP Hazard Perception
HRAoD Hierarchical Task Analysis of Driving
Hz Hertz
M Mean
m Meters
ms Milliseconds
OpenGl Open Graphics Library
QT-HPT Queensland Transport Hazard Perception Test
ROI Region of Interest
SA Situation Awareness
SAGAT Situation Awareness Global Assessment Technique
SD Standard deviation
SDT Signal Detection Theory
sec Second
UFOV Useful Field of View
CHAPTER 1

INTRODUCTION

The scientific field of psychology, in its many forms, has a rich history demonstrating the interplay between basic research and applied research. For instance, Helton, Kemp, and Hoffman (2015) provide details where perception research has been directly applied throughout history: military science of the ancient Greeks and Romans, echolocation in bats, color vision and forgery, astronomy, and artillery in the mid-nineteenth century. This intersection of basic and applied research can be found at the definition of applied cognitive psychology: the science that uses theories and methods from experimental psychology in an attempt to gain scientific understand of cognitive phenomena that occur in the real world and solve practical problems (Hoffman & Deffenbacher, 1992).

One concept in the real world that has provided an outlet for research is that of safety, which is the subject and goal of many areas of sub-fields of psychology (Geller, 2001), including human factors psychology. One particular safety-critical task is driving, and the application of psychology to driving has yielded an important increase in basic knowledge as well as used psychological knowledge to design better transportation systems.

Driving is a common everyday task. In 2009 there were more than 210 million licensed drivers in the United States who traveled 3 trillion miles and made 392 billion person-trips during that year (Federal Highway Administration, 2011). Even though driving is commonplace, it is incredibly complex; it is estimated that, for a regular drive, there are over 1600 bottom-level tasks involved (Walker, Stanton, & Salmon, 2015). These tasks involve decisions that must be made at different levels (Michon, 1985): strategic (e.g., route planning), tactical (e.g., hazard avoidance), and control (e.g., reaction to a wind gust).
The complexity of driving leaves much room for the possibility of errors, and these errors can lead to crashes. It is estimated that, in the United States, the total economic cost of yearly crashes is over $240 billion (Blincoe, Miller, Zaloshnja, & Lawrence, 2015). In addition to the economic impact, there is the loss of human life associated with many crashes. In 2016 alone, 37,461 people were killed in the United States in motor vehicle crashes (National Highway Transportation Safety Administration, 2018), and traffic fatalities are one of the leading causes of death (National Center for Injury Prevention and Control, 2011). Since the earliest fatal crashes over a century ago (Fallon & O’Neill, 2005; Porter, 1998; Roberts, 1998), transportation research has developed as one means of reducing these tragedies in a variety of ways (Hagenzieker, Commandeur, & Bijleveld, 2014). This has provided a rich area for human factors research to further understand the psychological components of driving, which has influenced the design of roadway systems, vehicles, and public policy (Green, 2012; Lee, 2008).

One of the psychological aspects of driving is visual perception and cognition (Sivak, 1996). Vision plays an important role in each of the three levels of driving tasks. For example, one has to recognize landmarks for strategic, perceive traffic environment for tactical, and detect slight variations in heading for the control level (Ward, 2000). Specific visual functions, such as acuity, contrast sensitivity, color perception, and motion detection have all been shown to play some role in the ability to drive safely (Shinar, 2017). One area that has received much attention in the driving and vision literature is the useful field of view (Wood & Owsley, 2014).

Generally defined as the area from which information can be extracted in a single glance (Ball & Owsley, 1993), the useful field of view (UFOV) has been measured a number of different ways by computerized tests (Edwards et al., 2005). Unlike many measures of visual function, this test was specifically developed to understand visual cognition as it operates in
everyday life (Ball & Owsley, 1993). One of these everyday tasks that the UFOV was designed to be applied to is driving. Results of the test have shown an ability to predict crashes, with those who do poorly on the test more likely to be involved in crashes in real life (as measured by self-report and crash records) or in simulators (Clay et al., 2005).

Although this is compelling evidence for the importance of the UFOV, there are several areas that can add to the understanding of it. First is furthering the validity of the UFOV test by incorporating more realistic aspects that are relevant to spatial attention. Since the goal of the UFOV was to create a measure that specifically was testing everyday visual functions, the ecological validity is very important. The second area is to further understand the link between UFOV as a cognitive visual function and crash involvement. Crash involvement is a poor variable to use as a reliable and valid measure of driving performance (Horswill & McKenna, 2004; Shinar, 2017). Because of this, it is important to understand why the UFOV is related to crashes and what driving tasks are specifically related to the functions tested in the UFOV.

This dissertation addresses these two issues. First, a test was designed to evaluate the effect of depth on the UFOV. Depth is inherent in everyday life, and there is evidence that depth plays a role in how attention is allocated through space (Atchley, 2005; Howard & Rogers, 2012). If the UFOV is measuring an important aspect of spatial attention for everyday tasks such as driving, the influence of depth should be considered. The development of this test has important implications for how the UFOV is understood, tested, and applied to everyday life. Second, an exploratory study was completed assessing the predictive quality of the 3D UFOV on hazard perception (HP). HP, a specific type of situation awareness, has been shown to be related to crash risk and involves many attentional aspects of driving (Horswill & McKenna, 2004). The second experiment assesses the relationship between the UFOV and a video-based HP. In
addition to specifically relating this test to the UFOV, a further goal was to evaluate its use as an appropriate real-world task of attentional abilities, to lay the ground for future research.

Chapter 2 provides a literature background for the UFOV, attention in depth, and hazard perception. Chapter 3 provides background on the creation of a stereoscopic 3D UFOV test. This includes reviewing the relevant literature on depth and stereopsis as well as on the UFOV and the major pilot studies that lead to the final version. Chapter 4 details experiment 1, which assess the effect of depth on the UFOV in three different ways. Chapter 5 details Experiment 2, which is the exploratory study relating the 3D UFOV to a HP test. Finally, Chapter 6 summarizes the overall findings and implications of the dissertation.
CHAPTER 2
LITERATURE REVIEW

2.1 Visual Attention

Attention can generally be thought of as any one of a wide range of mechanisms that exists to help us to deal with our inability to process all the information presented to our perceptual system (Wolfe, 2014). One important reason is that we simply lack the ability to process all of this information, and so our finite amount of cognitive resources can be allocated to attend to specific pieces of information and filter other information out (Kahneman, 1973; Wickens, 2002). Studies of attention generally fall under one of two broad categories: divided or focused (or selective) attention. In general terms, divided attention refers to the limit of performance of combined tasks and selective attention refers to the ability to ignore irrelevant stimuli (Kahneman & Treisman, 1984). Attention can be applied to all sensory modalities and various cognitive processes.

Attention has been investigated since the infancy of experimental psychology (e.g., James, 1880), with much of that research focusing on visual attention, and research in that field increasing dramatically in the last three decades (Carrasco, 2011). Visual attention can be defined as the process (or set of processes) that allows an observer to recruit resources for processing selected aspects of the retinal images more fully than other areas (Palmer, 1999). Visual attention is often conceptualized as being limited spatially; we can process things within one area of our visual field but not in another area. Various metaphors have been employed by researchers to aid in our understanding of this concept; two prominent metaphors are important for this dissertation and elaborated upon in this review: the spotlight and the zoom lens.
The spotlight metaphor has origins early in the history of attention research. Pierce (2013) reports the earliest use of this metaphor was by Hunter (1919), who described attention as the “degree of light upon an object as a man casts a spotlight here and there” (pp. 113-114). One of the most cited definitions supporting this metaphor for attention was by Posner, Snyder, and Davidson (1980), describing a spotlight that “enhances the efficiency of detection of events within its beam” (p. 175). The spotlight metaphor is meant to imply three properties: attention enhances processing in a spatially localized way, the “spotlight” only processes information in one region of the visual scene at a time, and the spotlight has a spatial limit. A limitation of the spotlight metaphor is that there is an implied fixed size of the spotlight, which does not allow for changes in the size (Duncan, 1984; Treisman & Schmidt, 1982). This limitation has led to the development of the zoom lens metaphor, which proposes that the size of the lens is variable, depending on the demands or requirements of the tasks (Posner et al., 1980). These metaphors liken attention to a kind of beam that can be moved across the visual field to extract information, and the size of that beam may change depending on the various factors.

Both of these metaphors presuppose that attention can be moved through space in order to complete a task. A paradigm to study visual attention in space is visual search, which has been used extensively in both cognitive and human factors psychology. As a general definition, visual search refers to an effort to detect or locate a target item whose presence or position within the search field is not known a priori (Wickens & McCarley, 2008). Important within the context of search is the difference between parallel and serial search. A visual search will begin with a global assessment of the entire field, and then attention shifts to inspect various regions of interest (ROIs) in order to find the target. A serial search involves items being processed one after the other, whereas in a parallel search all items are processed simultaneously. Wolfe and
Horowitz (2004) cataloged over a dozen attributes that have been found to guide parallel search, including color and motion. When parallel searches occur, it does not matter how many distractor items there are (i.e., the set size is irrelevant), as the features tend to “pop out”. In serial search, however, the set size does matter, as the observer must generally inspect the object or area one at a time (Wolfe et al., 2012).

In most cases, unless a target can be identified in the parallel, global stage, some amount of serial processing will occur. The goal of visual search, then, is to have the visual lobe (i.e., the size of attention in which an observer can conduct parallel search) cover the target (Chan & Courtney, 1996). This lobe has various names in the literature, such as the functional field of view (Crundall, Underwood, & Chapman, 1999; Sanders, 1970), perceptual span (Pringle, Irwin, Kramer, & Atchley, 2001), or most commonly, the useful field of view (UFOV). The UFOV was defined and measured by Karlene Ball and her colleagues (Ball, Beard, Roenker, Miller, & Griggs, 1988; Ball & Owsley, 1993), and describes how much information can be searched in parallel in one glance (i.e., without any eye movements).

2.1.1 The Useful Field of View

The UFOV was initially developed as a way to understand the visual impairments that were reported by older adults but were not captured on standard assessments of visual function used at the time (Ball & Owsley, 1993). In particular, Ball and Owsley (1993) wanted to develop a link between visual function and everyday tasks, which happen under various illuminations and contrast levels, in complex and dynamic scenes, and include uncertain targets. Since its development, the UFOV has been found to be related to a number of other cognitive functions, such as attention, executive functioning, general cognition, memory, spatial ability,
visual closure, contrast sensitivity, visual processing speed, and visual acuity, suggesting that it represents a range of important cognitive abilities (Woutersen et al., 2017).

A test, initially developed by Sekuler and Ball (1986), has been used to specifically measure the UFOV, and subsequent versions have been commercialized. The commercial version of this test evaluates central (processing speed), divided, and selective attention. The central task requires participants to identify which of two similar silhouettes of a car or truck was presented in the center of the screen. The divided attention task supplemented the central task with a peripheral stimulus target stimulus (i.e., the silhouette of a car). Participants are required to report both which target was presented in the central field and the visual meridian upon which the peripheral target was presented. The selective attention task differed from the divided attention task only in that the peripheral target was embedded among distractors consisting of inverted triangles. Performance on these tests is defined as the minimum display duration necessary for the targets to be reliably processed. Various adaptations have been made to the commercialized test (Edwards et al., 2005) and other versions have been created (e.g., Crundall et al., 1999; Gaspar et al., 2016; Richards, Bennett, & Sekuler, 2006; Rogé, Pébayle, El Hannachi, & Muzet, 2003), but generally they test for either the spatial extent of the processing at a set display duration or for the speed of processing across set spatial locations (Wood & Owsley, 2014).

The size of the UFOV can be affected by a number of variables, including both environmental and personal factors. In general, research has suggested that the size of the visual lobe is dependent on the overall processing demands of the scene. Additionally, much research has shown that there are significant age effects for the size of the UFOV, where older adults generally have a smaller area from which they can extract information compared to younger
adults (Ball et al., 1988; Owsley, Ball, Sloane, Roenker, & Bruni, 1991; Sekuler, Bennett, & Mamelak, 2000). Because of the importance of being able to attend to task-relevant information in the environment, and the changes based on these various factors, the UFOV has been the subject of much research and applications to the real world.

The UFOV was designed to understand vision as it applied to everyday activity and the application that has received the largest amount of research interest is driving. As outlined by Ball and Owsley (1993), they decided to use driving as the model task for older adults, because it is the most common complex activity completed by older adults and proposed the UFOV test as a measure of the “window of attention”. Figure 1 illustrates the importance of the UFOV to driving at the conceptual level (Ball & Owsley, 1993). With this figure, the authors note in addition to seeing less of an area of the visual field in one glance (i.e., in parallel search), those who have a reduced UFOV size would require more time to scan the same amount of space. In a dynamic situation such a driving, the reduced ability to process important information quickly can be unsafe and is, therefore, an important area of study.
Figure 1: An illustration of the detrimental effects from the reduction in UFOV size, where panel A represents no restrictions and panel D represents an extreme reduction (Ball & Owsley, 1993).

Though the UFOV attempts to measure visual processing in the real world, there is one naturalistic component for which it does not account: depth. The “size” of the UFOV is always measured two-dimensional visual angles on a computer screen; the real world is not laid out that way. Looking at the conceptual diagram of the UFOV in Figure 1, the spatial representation does not just include the two-dimensional, \(x\)- and \(y\)-axis, but also the depth (\(z\)-axis) of the objects and the depth of these objects may be very important to the driver in order to assess the likelihood of a crash. If the UFOV is to truly test visual cognition for everyday tasks, the influence of the depth of objects in the scene should be addressed, as depth is an important
component of our scene perception in everyday life and seems to have an effect on the allocation of attention.

### 2.1.2 Perception in Depth and Stereopsis

The world around us is full of Euclidian geometry (lines remain parallel as they extend through space, objects maintain the same size and shape as they move around, etc.). When this three-dimensional (3D) world is projected onto a curved, two-dimensional (2D) retina, it becomes non-Euclidian (lines do not remain parallel and objects change size and shape as they move around, etc.). The ability of humans to interact with the environment requires an understanding of the structure and layout of the space (i.e., depth). In order to do this, the visual system relies on a number of different heuristics in an attempt to recreate the 3D environment from the image that is projected onto the retina (Wolfe et al., 2012).

![Diagram of depth information]

**Figure 2:** Major source of depth information, adapted from Sekuler and Blake (1994).

A useful way of categorizing these heuristics is by the type of cues that are used (Sekuler & Blake, 1994); Figure 2 provides a summary of the different categories. At the highest level of categorization, there are either oculomotor or visual cues. Oculomotor cues are based on the
degree of convergence of eyes or accommodation of the lens and are derived from the sensation of muscular contraction. Convergence refers to the angle that each eye turns inward, toward each other when a person is fixating on an object. Accommodation refers to a change in the shape of the lens that is based on the distance of fixation. Visual cues are based solely on the image projected onto the retina and can either be monocular or binocular. Monocular cues, which require only one eye, can be static, such as interposition, relative size, and perspective or can be dynamic (i.e., motion parallax). Stereopsis, the binocular depth cue, requires both eyes. These depth cues all work together to create the experience of depth in a scene, and each cue provides different effectiveness at different depths (Cutting & Vishton, 1995).

Of particular importance to this dissertation is the binocular depth cue: stereopsis. In most predatory animals, there is significant overlap in the visual field of both eyes, and that overlap can be as large as 124° in primates (Land & Nilsson, 2012). This overlap causes slightly different images of the same scene to be presented to the retina in each eye. This difference, called binocular disparity, changes in magnitude throughout the visual field in relation to the fixation point, and perception of depth from this difference is called stereopsis. Through stereopsis the observer experiences “at once an obvious and striking sense of . . . depth which is not evident when either target is viewed by one eye alone” (Ogle, 1964, p. 135). There is no disparity at the fixation point, and also in the area around the fixation point called the horopter. Around the horopter is an area, Panum’s fusional area, where disparate images can be fused and perceived as one image; beyond that area two images (diplopia) are perceived (Harwerth & Schor, 2011). Figure 3 shows a top-down view of this area.

Stereopsis can be experienced by an observer in different ways and has been divided into quantitative (or patent) and qualitative stereopsis (Ogle, 1952a, 1952b). Quantitative stereopsis
occurs from smaller disparity (both fused and diplopic) and allows for the perceived depth to vary with disparity. Qualitative stereopsis occurs with diplopia and allows the viewer to simply determine whether one object is behind or in front of the fixation point; Figure 3 shows the relationship between retinal disparities and stereopsis using this categorization. General guidelines for the disparity limits at the fovea for a target of 1 degree are 20 arcmin for fusion, 2 degree for patent stereopsis, and 8 degrees for qualitative stereopsis (Patterson & Martin, 1992), and these limits increase as retinal eccentricity increases (Ogle, 1962).

Figure 3: A top-down view of binocular vision from Ogle (1952a). The area of single vision is where the objects are fused, and occur around the horopter.
Although stereopsis occurs in the world naturally, it can also be artificially induced by presenting separate images to the left and the right eye (Banks, Hoffman, Kim, & Wetzstein, 2016). This was first done by Wheatstone (1838), who used two separate displays to present the different images into the eyes with a mirror haploscope; in more recent times a variation of this method is used in head-mounted displays (Cakmakci & Rolland, 2006). A more common method is an overlapping stereoscopic display, where each eye is presented with different images from the same display through special glasses that modify what is displayed to the observer. This can be done through physical properties of the display (color anaglyph and polarization display) or by synchronizing separate images with active shutter glasses (temporally alternating displays). Figure 4 illustrates these methods of image separation to induce stereopsis. These methods to induce stereopsis have been used in many domains of human factors and performance research and practice (McIntire, Havig, & Geiselman, 2014).
2.1.3 Attention in Depth

Perception in depth is important for viewers to complete many everyday tasks. In addition to simply perceiving depth, it is also important to evaluate the influence of depth on spatial attention—particularly when evaluating an applied construct like the UFOV. There is a small but growing body of research addressing the effect of depth on attention during the past few decades (Atchley, 2005; Howard & Rogers, 2012). The following reviews the major
findings of attention in depth literature related to divided and selective attention across three-dimensional space.

Some of the earliest studies evaluating the effect of depth on attention used modifications of the cueing paradigm, where participants' attention is directed by a cue that is either valid (i.e., cueing to the correct spatial location) or invalid (i.e., cueing to the incorrect spatial location). These tests, traditionally conducted in 2D, measure the ability to shift attention. The first studies used targets’ real depths in a scene (as opposed to artificially induced depth), and found that invalid cues increase reaction time more when the cues indicated a different depth (Downing & Pinker, 1985; Gawryszewski, Riggio, Rizzolatti, & Umiltá, 1987). These results suggested that there was a larger cost in moving attention in the x-, y-, and z-axes than just in the x- and y-axes alone in the scene. To eliminate potential confounds present in using real depths (e.g., accommodation and vergence), Folk and colleagues (Ghirardelli & Folk, 1996; Iavecchia & Folk, 1994) created a stereoscopic three-dimensional cuing paradigm and were not able to replicate the effect of depth. However, Atchley, Kramer, Andersen, and Theeuwes (1997) could replicate the depth-dependent effects when the visual displays included distractors, suggesting that the effect is more pronounced in complex displays. Additionally, Marrara and Moore (2000) found that various aspects of the displays (e.g., display duration and availability of object tokens) can better facilitate the influence of depth on attention when using stereoscopic displays for the cueing paradigm.

In addition to selectively moving attention in depth, there is evidence of the importance of depth of distractors in focused-selective attention. Using a stereoscopic flanker task, Andersen and colleagues (Andersen, 1990; Andersen & Kramer, 1993) varied the depth of the distractors. The studies consistently showed that a distractor’s interference is affected by its
depth, where slower responses were found by compatible distractors at the same depth as the central target than those at different depths. It may be that it is easier not to be influenced by distractors presented at different depths, and therefore, make targets “pop out” from them. Research in the visual search field has suggested that depth is an attribute that affects attention as well. Nakayama and Silverman (1986) found that parallel search can occur when distractors are on a different depth plane presented stereoscopically, meaning that distractors have less of an effect when they are on a different depth plane than the target. Furthermore, it has been suggested that the depth of a target is something that facilitates parallel search (Wolfe & Horowitz, 2004).

Little work has been done on the effect of depth on divided attention, and to the author’s knowledge these studies have been conducted using monocular cues; the results, however, do show support of depth influencing divided attention. Using a driving simulator, Andersen, Ni, Bian, and Kang (2011) required participants to follow a lead vehicle and respond to light changes that varied in depth above the roadway; Figure 5 shows an example display from that study. They found that the response time to identifying the light changes was a function of both its horizontal location and the distance from the driver. Similar results were found using this paradigm in older adults (Pierce & Andersen, 2014). Additionally, this effect was found when peripheral targets displayed for a much shorter period of time (Song, Bennett, Sekuler, & Sun, 2017), which is similar to the peripheral targets used in UFOV tests.
Figure 5: The stimulus from Andersen et al. (2011). Participants were required to follow the lead vehicle and respond to changes in the low arrays above the road. The depth of the arrays had an effect in addition to the visual angle from the lead vehicle.

These study measures attention in a way that is very similar to the UFOV, requiring participants to maintain a certain distance from the lead vehicle (central task) and respond to light changes (peripheral task). Some researchers have made driving-related UFOV (e.g., Gaspar et al., 2016; Rogé et al., 2004); however, they do not incorporate the depth of the target, which may mean they are not using important information in the scene. Figure 6 shows screen captures from those two studies. Because they are realistic scenes, depth is inherent in both versions’ peripheral targets (Gabor patches and vans); that could have effects on how attention is allocated.
Figure 6: Screenshot from real-world UFOV simulators; Gaspar et al. (2016) is on the right side Rogé et al. (2004) is on the left. The visual angle of peripheral targets from the lead vehicle was assessed, but not an effect of the depth of the targets.

To summarize, the research cited above suggests the depth of targets in space does affect the allocation of attention—both divided and selective. If the allocation of attention is affected by depth, it would be important to consider its effect on tests like the UFOV. If there is an effect of depth, incorporating depth in the test would make it more ecologically valid. Additionally, it could have direct relationships to real-world tasks like driving, in which the depth of objects is an important aspect.

2.2 Vision and Driving

Vision overall plays a very large role in driving (Sivak, 1996), and many specific aspects of vision have been related to driving performance and safety (see Shinar, 2017, for a review). One of the earliest studies of vision and driving was done by Gibson and Crooks (1938), who described driving as locomotion by means of a tool that is chiefly guided by vision. Over the past 35 years, the number of articles published on vision and driving has tripled (Owsley, Wood, & McGwin, 2015). Additionally, specific visual aspects such as attention, eye movements, and contrast sensitivity have been investigated in the literature and have been related to driving safety and performance (Owsley & McGwin, 2010). Overall, these studies suggest a firm relationship...
between visual functioning and driving safety. One of the important components of vision that has been related to driving is attention, and researchers have created various models of attention for understanding it specifically in the driving context (e.g., Horrey, Wickens, & Consalus, 2006; Trick, Enns, Mills, & Vavrik, 2004).

From its inception, the UFOV has been linked to driving safety and performance in various ways, and often times these relationships have been found using older populations. For example, Ball, Owsley, Sloane, Roenker, and Bruni (1993) investigated the predictive power of a number of cognitive measures, including the UFOV and compared them to state-reported crashes in the past five years for older adults. The size UFOV by itself predicted crashes with 89% sensitivity (i.e., the likelihood of having a smaller UFOV given crash involvement) and 81% specificity (i.e., the likelihood of having a larger UFOV given no crash involvement). Similar results were found by Goode et al. (1998), who found that UFOV alone predicted state-reported crashes with 86.4% sensitivity and 84.3% specificity. In a meta-analysis, Clay et al. (2005) found that, across eight unique studies that tested older participants for crash risk based on state crash reports, self-reports, simulators, and naturalistic road tests, the UFOV robustly predicted crashes across multiple indices of driving performance.

Despite its predominant use with older adults, the UFOV has value for younger adult populations as well. Creating normative data from children age 5 through 22, Bennett, Gordon, and Dutton (2009) found that the test was sensitive to visual changes throughout childhood and into early adulthood, and Burge et al. (2013) found younger adults can improve on the UFOV with training. McManus, Cox, Vance, and Stavrinos (2015) specifically applied the UFOV to simulated driving in a younger population (mean age 19.7 years). They found that the selective attention subtest significantly predicted collisions, showing that for each 30 ms reduction in
UFOV threshold, there was nearly a 10% increase in crashes. Similar results were found using younger commercial drivers, particularly while they are engaged in secondary tasks (McManus, Heaton, Vance, & Stavrinos, 2016; Wood et al., 2006).

To summarize, research shows strong evidence of a relationship between the UFOV and crash risk, both in older and younger adults. The consistency of the research using different types of driving scenarios suggests a high amount of converging operations (Goodwin & Goodwin, 2013), meaning that there likely is some robust link between the measures of attention and crash risk. However, there has been a lack of research investigating the specific link between UFOV and crash risk by way of driving tasks, particularly as there are some methodological problems with using crash rate as a dependent measure.

2.2.1 Problems with Crash Involvement

Because safety is a concern for driving researchers, crashes are often seen as the “gold standard” of measurement for driving (Groeger, 2011). However, there are some potential weaknesses in using this as an overall metric that represents driving performance. First, it carries with it difficulties in terms of its psychometric reliability. For example, as a measure of test-retest reliability, correlations of crash involvement over successive time periods are very small (Elander, West, & French, 1993), with some researchers finding coefficients as low as .31 comparing two successive one year periods (McKenna, Duncan, & Brown, 1986) and a three year period predicting the following year (French, West, Elander, & Wilding, 1993).

In addition to these psychometric problems, Horswill and McKenna (2004) propose three reasons why crash involvement may be an unreliable measure of a driver’s crash liability. First, crashes are not homogenous (Barmack & Payne, 1961). Any one crash may or may not be the fault of the particular driver and it could be caused by temporary states like drowsiness that is not
necessarily related to the driver’s skill. Second, crashes are caused by interactions of multiple factors, not just one. Third, they are rare events. It is estimated that the average driver has one crash every ten years (Evans, 1991), which means that in any one year, approximately 90% of individuals will have no crash at all.

Additionally, there are simply many biases in how crash involvement is measured. Self-report is a common method, but it has been shown that self-report crashes are underreported (Harano, Peck, & McBride, 1975) or forgotten (Maycock, Lockwood, & Lester, 1991). Though police crash reports are often seen as a more objective alternative, minor crashes are often not reported (Elander et al., 1993). Police reports have been shown not to correlate highly with other object measures of crash involvement, such as video recordings (Chung & Chang, 2015) or crash investigation teams (Shinar, Treat, & McDonald, 1983). Additionally, interfaces of the reporting systems are often designed poorly and, therefore, lack high data quality (Morris, Achtemeier, Ton, Plummer, & Sykes, 2016).

The problems listed above have prompted researchers to relate visual attention to more specific aspects of driving than just crash involvement. For example, Joanne Wood and her colleagues created an overall driving metric, which uses a standardized (z) score of the combined effects of the following driving-related variables: road sign recognition, road hazard recognition and avoidance, gap perception, divided attention task, total driving time, maneuvering, and reversing. They found that, in an older adult population, when the UFOV was combined with motion sensitivity, contrast sensitivity, and dynamic acuity it could explain 50% of the variance in driving scores (Wood, 2002). A similar paradigm was used by Wood, Chaparro, Lacherez, and Hickson (2012), who found that the UFOV was able to predict the standardized driving measures in older adults in the presence of auditory and visual distractors.
The research described above does look at driving as more than just crash involvement; however, it still treats driving performance as one overarching variable of interest, not as sub-sets of driving skills. It does not address what specific aspect of driving the UFOV may be related to, which in turn may lead to a higher rate of crash involvement. Although driving is a very complex task (Walker et al., 2015), there must be specific links between attention that only affect some aspect of driving, whereas other aspects of driving are not related to vision at all. When holistic measures of driving are used to link attention to this applied task, often results are mixed or even counterintuitive. For example, Weaver, Bédard, McAuliffe, and Parkkari (2009) investigated the relationship between a measure of attention (the attention network test) and driving. They used the Manitoba Road Test, which is a demerit-based system around general categories of driving performance, and did not find strong relationships between their attention measure and driving. It is suggested that the reason for the null results is because tests like the Manitoba Road Test do not seem to tap into attention-relevant components of driving when using the overall demerit score (Roca, Crundall, Moreno-Ríos, Castro, & Lupiáñez, 2013). Addressing this link between the UFOV and driving should be an area of importance and a necessary area of research. One such area that may provide a link between attention and driving is that of situation awareness.

2.2.2 UFOV and Situation Awareness

As mentioned above, driving is a very complex task. In the comprehensive Hierarchical Task Analysis of Driving (HRAoD; Walker et al., 2015), it was estimated that over 1,600 bottom-level tasks and 400 plans are conducted while driving a modern car on regular roads to get to a driver’s desired destination. These tasks encompass many aspects of driving. Some examples include:
• Do not allow either hand position to go past twelve o’clock on the steering wheel (while turning)
• Watch for warning signs to slow down (dealing with junctions)
• Act on advice/instruction/rules/guidance provided by the Highway Code (complying with rules)

It seems likely that attention may be related to some of these tasks (e.g., watch for warning signs) and likely not for others (e.g., proper hand position while turning); therefore, only some of the tasks might be related to the UFOV.

Since driving is a complex task, the types of errors made by drivers are many and do not all relate to attention. The driving behavior questionnaire (DBQ; Reason, Manstead, Stradling, Baxter, & Campbell, 1990) addresses 50 errors that drivers could make that varied in type of risk posed to other drivers. Examples of errors that are high risk include:

• Fail to check your mirror before pulling out, changing lanes, turning, etc.
• Lost in thought or distracted, you fail to notice someone waiting at a zebra crossing, or a pelican crossing light that has just turned red.
• Deliberately disregard the speed limits late at night or very early in the morning.
• Get involved in unofficial ‘races’ with other drivers.

As with the HTAoD, the DBQ shows errors that likely are related to attention (e.g., failing to notice a pedestrian in a crosswalk) and some that are likely not (e.g., disregarding speed limits and racing other drivers). In reviewing the literature on human error in driving, Stanton and Salmon (2009) propose five factors that can cause errors in driving: road infrastructure, vehicle, driver, other road users, and environmental conditions. Attention only logically fits within one of those factors. Given the complexity of the driving task and the variety of error types and their
causes, it is important to understand which aspects of driving are related to attention and test the relationship of those driving measures to the UFOV.

A type of skill that seems to be related to many safety-critical tasks is situation awareness (SA). SA is a construct with a wide variety of conceptualizations and definitions (Stanton, Salmon, Walker, Salas, & Hancock, 2017; Vu & Chiappe, 2015). At the high level, it can be thought of as simply knowing what is going on around you or having “the big picture” (Jones, 2015). A popular model for SA was proposed by Mica Endsley (1995c), where SA is divided into three specific sub-stages that act linearly: perception, comprehension, and projection. This model is shown in Figure 7. The first step is to perceive the necessary information for the task at hand (level 1). After information is perceived, it must be comprehended (level 2). Comprehension involves integrating information from multiple sources and interpreting, storing, and retaining that information. Finally, projection (level 3) is the ability to forecast future events. SA then leads to a decision made by the operator and then a performance of an action. All of these stages may be affected by personal or environmental factors, such as interface design, workload, and experience. The performance of an action leads to more information in the environment and feeds back to level 1 of SA to begin the process over again. SA, the decision, and the performance are affected by a number of outside factors from the environment (e.g., system capability, stress, and workload) and individual factors (e.g., goals, automaticity, and experience). Figure 7 shows a conceptualization of the model.
SA has been shown to be an important construct to understand safety-critical tasks. For instance, Wickens and Hollands (2000) argued that SA has received considerable attention because of its relevance for designing displays and for understanding the cause of disasters where situation awareness has been lost. Endsley (1995b) attempts to address the importance of the different levels of this by examining the level of SA that is responsible for the crashes in National Transportation Safety Board reports from 1989-1992. Of the 15 cases, 32 SA errors were found: 72% were found in level 1, 22% in level 2, and 6% in level 3. Though the sample size is small, these findings suggest the importance of level 1 SA, which includes attention and

**Figure 7: The information-processing model of SA (Endsley, 1995c).**
other lower-level cognitive processes in completing tasks successfully and safely. One likely component in level 1 SA is the UFOV.

Theoretical evidence that relates the UFOV and SA has been suggested by Gugerty (2011), who said that the UFOV is the ability to detect salient peripheral events, and then allocate attention to that area. Chaparro, Groff, Tabor, Sifrit, and Gugerty (1999) were the first to investigate a possible relationship of the UFOV to a measure of SA. Using the DriveSim paradigm (Gugerty, 1997), they measured driving performance by crash rate and the recall of both the location of cars and hazards after the simulation is blanked. The significant correlations can be found in Table 1. Overall the authors concluded that those who do poorly on subtests 2 (divided attention) and 3 (selective attention) are less aware of cars and less effective at responding to hazards. These strong effects should demonstrate the potential importance of the UFOV to SA in the driving task and as a possible link between UFOV and crash risk.

<table>
<thead>
<tr>
<th></th>
<th>Divided Attention</th>
<th>Selective Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crash Avoidance</td>
<td>.52</td>
<td>.83</td>
</tr>
<tr>
<td>Reacting to Threat Car</td>
<td>.33</td>
<td>.66</td>
</tr>
<tr>
<td>Proportion of Cars</td>
<td>.71</td>
<td>.62</td>
</tr>
<tr>
<td>Recalled</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Taking it a step further, Kaber, Jin, Zahabi, and Pankok (2016) compared UFOV scores to measures of driving SA measured by a version of the Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1995a). The SAGAT uses a freeze-probe method, where the simulation is periodically blanked and a number of questions are asked of each participant to test their current level of SA. A benefit in this study is that these questions can probe at a specific level of SA, not just a global measure of the construct. For example, the question “What was the
name of the last crossing roadway you passed?” measures level 1 SA, “How long has it been since you passed the last intersection?” measures level 2 SA, and “How long will it take to reach the lead car at your current speed?” measures level 3 SA. They found that lower UFOV scores for divided and selective attention were related to higher levels of SA, but only in level 1. The UFOV did not appear to specifically relate to variance in levels 2 or 3.

The research above suggests that there is indeed a relationship between UFOV and SA and that it may focus the perception of objects in the scene (level 1 SA). Following Endsley’s (1995c) model, having an accurate perception of the situation is necessary for all subsequent steps. This would include other aspects of SA (having higher-level knowledge of the situation and predicting the future events) and also influence functions beyond SA (decision making and the behavioral response). Therefore, even though the UFOV was only found to be related to level one, it should not be discounted as having a small effect on the SA of a driver. This is particularly important because, as noted above, problems with level 1 SA have a high contribution to crashes (Endsley, 1995b).

2.2.3 Hazard Perception

One particular aspect of the environment of which every driver must be aware is a hazard. In a review of a half-century of driving safety research, Lee (2008) concludes that many crashes are caused by the failure to look at the right thing at the right time, and Rumar (1990) terms late detection of hazards as the basic driver error. Large-scale studies of driving find that failure to respond to hazards is a leading cause of crashes (Treat et al., 1979). Therefore, the ability to perceive hazards and react appropriately is an important skill for driving and may, in fact, be the most important aspect of SA for a driver. Therefore hazard perception is an
important component of understand driving performance and safety. This component of SA has a large body of research associated with it in the last 30 years.

One of the reasons for the popularity of hazard perception (HP) is that it is one of the only measures of driving performance that has consistently been found to be related to crash risk (Horswill, 2016a; Horswill & McKenna, 2004). For example, Horswill, Hill, and Wetton (2015) found that those who failed a video-based HP test as part of the licensing requirement in Queensland, Australia, were 25% more likely to be involved in a crash during the year following the test. Horswill, Anstey, Hatherly, and Wood (2010) found that in older adults, those who had a mean response latency of 6.68 seconds or slower were 2.32 times more likely to have been in a self-reported crash in the previous five years. Though, as mentioned above, crash involvement is often seen as unreliable as a measure of driving performance, HP has also been found to be related to other crash-related measures, such as distraction (Borowsky et al., 2015; Borowsky et al., 2014), drowsiness (Smith, Horswill, Chambers, & Wetton, 2009), and alcohol consumption (Deery & Love, 1996). Taken as a whole, this suggests that HP likely has a casual implication in many crashes (Horswill, 2016b).

A majority of HP tests generally involve drivers watching images or films of traffic situations and having them indicate their awareness of potentially hazardous events (Horswill, 2016b). In these scenes, drivers will have to identify hazards (through response latency, response sensitivity, etc.) or rating the hazard, by categorizing various situations (Oron-Gilad & Borowsky, 2015). A common way to display the traffic situation is through films captured from a driver’s perspective, and are typical of HP tests that are part of licensing requirements (Horswill, 2016b). For example, the Queensland Transport Hazard Perception Test (QT-HPT) has videos of genuine traffic situations, and drivers use their mouse to click on a road user whom
they anticipate to be in a traffic conflict as quickly as they can. A participant’s HP score is the mean response time to these conflicts or hazards (Wetton, Hill, & Horswill, 2011).

In these video-based clips, the variable of interest is generally latency, focusing not if a participant can identify a hazard, but when he or she can. This allows researchers to assess hazard perception on a continuous variable that is likely sensitive to individual differences. In addition to the response latency, HP tests often include a measure of correct identification of hazard. The task then falls into an event detection task and is often analyzed as the number of correctly identified hazards. However, using a method called signal detection theory, HP detection tests can separate the two aspects of the response decision: sensitivity and response criteria.

### 2.2.4 Using Signal Detection Theory to Measure Hazard Perception

Over the past 60 years, signal detection theory (SDT; Green & Swets, 1966) has become one of the most prominent paradigms for study detection tasks (Bohil, Szalma, & Hancock, 2015). In this paradigm, it is assumed that two states of the world exist, “noise” and “noise + signal”. The decision to identify signal from noise is based on two factors: sensitivity and response criterion. Sensitively refers to the capacity of the observer to discriminate among categories, and response criteria to the internal bias for selecting one category over the other (Macmillan, 2002). SDT, therefore, attempts to separate the bottom-up processes of perception with more top-down, higher level cognitive processes (Wickens, Lee, Liu, & Becker, 2004).

The simplest SDT experiment is the “yes-no” experiment, where the observer must just respond whether a target was present or not. This means that there are two types of signals (noise or noise + signal) and two types of responses (yes or no), and therefore four possible outcomes: correct response (hit and correct rejections) or incorrect responses (false alarm or miss). Table 2
illustrates these four possible outcomes. Using the hit and false alarm rates, various measures of sensitivity and response criterion can be calculated (See, Warm, Dember, & Howe, 1997).

Outside the laboratory, SDT has been applied to many domains as diverse as air-to-air combat training (Eubanks & Killeen, 1983), medical diagnoses by radiologists (Swets et al., 1979) and psychiatrists (Wyshak, Barsky, & Klerman, 1991), and forensic science and training (Phillips, Saks, & Peterson, 2001).

Table 2: An Illustration of the Four Possible Outcomes of a Traditional SDT Task

<table>
<thead>
<tr>
<th>Observer Response</th>
<th>State of the World</th>
<th>Signal Present (Signal + Noise)</th>
<th>Signal Absent (Noise Only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Yes” (Signal Seen)</td>
<td>Hit</td>
<td>False Alarm</td>
<td></td>
</tr>
<tr>
<td>“No” (No Signal Seen)</td>
<td>Miss</td>
<td>Correct Rejection</td>
<td></td>
</tr>
</tbody>
</table>

Of importance to the current study, SDT has been applied to HP (e.g., Burge & Chaparro, 2012). In these studies, the hazard was either present or not present in each trial (state of the world), and the participants had to respond if they saw or did not see the hazard (observer response) under differing conditions. A difficulty in this approach is that in the real world, hazards are not binary classifications but instead lie on a continuum, with each traffic situation containing a different level of potentiality into developing into a crash. Therefore, using traditional SDT may be inappropriate for HP, as there is no objective way to measure the true state of the world (Wallis & Horswill, 2007).

A possible solution is to use fuzzy signal detection theory (FSDT; Parasuraman, Masalonis, & Hancock, 2000), which builds upon traditional SDT by including fuzzy logic (Zadeh, 1965). In this paradigm, each trial is not limited to belonging solely to one of those outcomes and instead each trial can have partial membership for both the state of the world (0 =
“entirely non-signal-like” to 1 = “entirely signal-like”) and the response (0 = “entirely no-like” to 1 “entirely yes-like”). The four possible response of traditional SDT are used (hit, false alarm, miss, correct rejection), but each response can claim partial membership in those categories. For example (from Wallis & Horswill, 2007), the state of the world for a trial may be 60% signal-like (and therefore 40% non-signal-like), and a response could be 80% “yes”. The individual is mostly correct (hit = 60%), but over responds to it, so a proportion of the response (20%) is false alarm. There is also membership that belongs as a correct rejection (20%), but no membership (0%) in the miss category. Each trial’s total equals 100% with membership potentially spread across all four categories. Using these values, researchers can calculate sensitivity and response criterion in a similar way as is done in traditional SDT (see Bohil et al., 2015 for a summary of FSDT validation research). Table 3 provides an example comparing responses from traditional SDT and FSDT.

Table 3: Possible Outcomes from Traditional (top) and Fuzzy (bottom) SDT (adapted from Parasuraman et al., 2000).

<table>
<thead>
<tr>
<th>State of World</th>
<th>Observer Response</th>
<th>Hit</th>
<th>False Alarm</th>
<th>Miss</th>
<th>Correct Rejection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional SDT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fuzzy SDT</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.9</td>
<td>0.1</td>
<td>0.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The application of FSDT to HP video clip tests has been done both to evaluate expert vs. novice differences (Wallis & Horswill, 2007) and to assess the impact of secondary tasks on both sensitivity and response criterion (Burge & Chaparro, 2018). In each of these studies, it was important to separate the response into sensitivity and response criterion. In each of these
studies, the hazard films were rated by experts for their hazard potentiality, which is the likelihood that the scenes would lead to a situation in which a collision or near collision with another road user would occur unless some evasive action is taken by the driver. Participants responded as soon as they perceived a hazard, and then rated the likelihood of the hazard causing a collision.

2.3 Present Study and Literature Summary

The current study is divided into two experiments. The first addresses a more basic aspect of visual attention by evaluating the effect of depth on the UFOV. The second attempts to apply the UFOV to hazard perception as a way of testing the effects on a real-world task.

The first study evaluates the effect of depth on a UFOV test. The UFOV was developed to be an applied test, particularly to driving (Ball & Owsley, 1993), and therefore there should be significant concern for how well it is assessing attention as it happens in the real world. The traditional test only measures special attention (x and y) and therefore treats attention as only consisting of two dimensions. The world consists of three dimensions, however, and perceiving is an important aspect of an individual successfully interacting with the objects in the world. There is evidence that depth has an effect on the allocation of attention (Atchley, 2005; Howard & Rogers, 2012), and, therefore, the UFOV may be impacted by the depth of the targets in the scene.

Chapter 3 describes the development of a new 3D UFOV test using stereoscopic displays. The goal was to create a stimulus that could measure the UFOV (in the divided and selective attention tasks) and manipulate the depth of the peripheral targets in relation to the central target. Stereopsis was included as the only depth cue because in using that, other confounding variables associated with depth (e.g., size and brightness) can be controlled. Chapter 4 describes
Experiment 1 and consists of two parts: Experiment 1A address allocating attention across known depths and Experiment 1B addresses allocating attention across depths that are unknown. The former is relevant to allocating attention across a scene (e.g., a roadway environment with important objects at various depths) and the latter to reacting to stimuli that occur at an unexpected depth (e.g., a hazard that occurs much closer than the focus of attention).

The second experiment, described in Chapter 5, applies the results of the 3D UFOV to driving performance. UFOV has traditionally been applied to driving, and strong correlations have been found before test performance and crash involvement (Ball et al., 1993; Goode et al., 1998). Although this is promising, there is a significant concern when using crash as a variable (e.g., Horswill & McKenna, 2004), and many causes of crashes are not related to attention involvement (Reason et al., 1990; Stanton & Salmon, 2009). Therefore, it is proposed that the UFOV relates to SA and HP, which in turn relates to crash involvements.

HP, as an extension of SA, is strongly related to visual attention. Indeed, the quick identification and localization of the objects in a wide visual field (i.e., the UFOV) are very similar to the real world task of identifying and localizing hazards that may occur in the periphery. Since the UFOV has been shown to predict SA in driving tasks (Chaparro et al., 1999; Kaber et al., 2016), it is important to assess if it would predict HP, since that construct has been shown to be a reliable predictor of crash involvement (Horswill, 2016a; Horswill & McKenna, 2004).

Since HP is a detection task that requires a decision, it is measured using a modification of the video-based HP test developed by Burge and Chaparro (2018), which uses FSDT to evaluate both the sensitivity and response criterion to the hazard, comparing these two components of response decisions to individual differences in the UFOV. By doing this, the
predictive quality of a standard, 2D UFOV test was made as well, as assessing the extra predictive power of the 3D UFOV.

Overall, this study attempts to create a more holistic understanding of the UFOV. For the basic research component, it assesses the effect of depth on the UFOV. From the applied component, it assess if the UFOV predicts HP performance and if including depth into the UFOV improves the predictive ability. The results of both experiments 1 and 2 will increase knowledge and understanding about the UFOV and how it can be applied to driving safety.
CHAPTER 3

STEREOSCOPIC UFOV STIMULUS DEVELOPMENT

To the author’s knowledge, this study is the first attempt to create a stereoscopic stimulus to evaluate the effect of depth of the useful field of view (UFOV). Because of that, extensive literature reviewing and pilot testing was done to ensure that the specific requirements of stereopsis could be met while still measuring the UFOV construct. This chapter explains the background and development of the stimulus that was used throughout the dissertation. The first section details the pertinent aspects of stereopsis as it applies to the UFOV stimulus. The second section details the developed of the UFOV test as it progressed from a 2D version into the 3D test used in the dissertation.

3.1 Presenting Depth through Binocular Disparity

As discussed in Section 2.1.2, the 3D world is presented on our 2D retinae, and our visual system must use a variety of cues to infer depth from that 2D representation. These cues vary in terms of the source of information (ocular-motor vs. visual), depth information (relative vs. quantitative), and include specific cues like motion parallax, occlusion, and stereopsis (Sekuler & Blake, 1994). In natural scenes, these cues work together at different depths to give the rich impression of depth to the viewer (Cutting & Vishton, 1995).

Stereopsis is one of the forms of depth perception and refers to the ability of our visual systems to use binocular disparity (i.e., the different images projected on to our two retinae) to give a vivid impression of depth (Ogle, 1964). In the real world, it occurs naturally, but it can also be mimicked in the laboratory showing separate displays to each eye (Banks et al., 2016). For further information on these displays, see Section 2.1.2.
A benefit of using stereopsis on its own is that it can control for confounds that are inherent with other depth cues (either real or simulated). For example, if monocular cues are used, other aspects of the target will also likely change. As an object moves closer to the observer in the real world, that object will also appear larger. In stereopsis, the only difference is the slight different position of images on the two retinae; all 2D aspects look identical. For this reason, a vast majority of research evaluating attention and depth have used stereopsis as the only method participants have to perceive depth (e.g., Andersen, 1990; Andersen & Kramer, 1993; Atchley et al., 1997). The current version of the 3D UFOV was created with these advantages in mind, and from the onset of development, stereopsis was the only depth cue planned to be used. The remainder of this section summarizes the pertinent issues related to stereopsis that must be addressed in order to incorporate stereoscopic displays into a visual attention test like the UFOV.

3.1.1 Factors that Affect Stereopsis

Stereopsis generally occurs because of binocular disparity, however there are conditions that increase the precision that depth will be experienced by the observer (or if depth will be experienced at all). These are particularly important when artificially displaying stereopsis in isolation (i.e., without any other depth cues) because the only depth information comes from the binocular disparity. If some attribute of the display does not facilitate stereopsis and the viewer does not perceive depth, it is as though the manipulation did not occur. Boff and Lincoln (1988) catalog various factors that affect stereoacuity; this can be seen in Table 4. Though all are important to consider when constructing stereoscopic displays, two of them, the presence of depth references and exposure duration, are particularly relevant to the design of a 3D UFOV display and are discussed in more detail below.
Table 4: Factors Affecting Stereoacuity (adapted from Boff & Lincoln, 1988)

<table>
<thead>
<tr>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Illumination level</td>
</tr>
<tr>
<td>- Retinal location</td>
</tr>
<tr>
<td>- Relative disparity</td>
</tr>
<tr>
<td>- Target/background contrast</td>
</tr>
<tr>
<td>- Presence of depth references</td>
</tr>
<tr>
<td>- Configuration of reference contours</td>
</tr>
<tr>
<td>- Lateral separation of adjacent contours</td>
</tr>
<tr>
<td>- Viewing distance</td>
</tr>
<tr>
<td>- Field of view</td>
</tr>
<tr>
<td>- Fixation conditions</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>- Length of target</td>
</tr>
<tr>
<td>- Width of target</td>
</tr>
<tr>
<td>- Orientation of frontal plane</td>
</tr>
<tr>
<td>- Lateral motion</td>
</tr>
<tr>
<td>- Motion in depth</td>
</tr>
<tr>
<td>- Spatial frequency</td>
</tr>
<tr>
<td>- Exposure duration</td>
</tr>
<tr>
<td>- Target/comparison onset asynchrony</td>
</tr>
<tr>
<td>- Right/left retinal illuminances</td>
</tr>
</tbody>
</table>

Though some amount of stereopsis can be obtained with extremely short displays (e.g., Langlands, 1926), the extremely short duration does not give optimal stereoacuity to the observer. Ogle (1964) reports that optimal stereoacuity is found when targets are displayed for 3 sec. Targets displayed for a shorter period of time have a four- to five-fold decrease in acuity down to 0.2 sec, after which the thresholds remain relatively constant. Marrara and Moore (2000) specifically looked at stimulus duration in a 3D cueing paradigm study while investigating display components that played a role in how attention was allocated in 3D. They found that when a depth placeholder was displayed before the stimulus for a longer period of time (1600 ms) there was an effect of depth and when it was displayed for a shorter period of time (50 ms) there was not.

Display duration is therefore an important constraint for the UFOV, particularly as the stimulus is displayed very briefly to prevent eye movements; the commercial version is generally limited to 500 ms (Edwards et al., 2005). Therefore, it may be advantageous if some aspect of depth is experienced before the onset of the stimulus. Creating an environment where stereopsis is experienced before the target is displayed is common in the literature. For example, Andersen (1990) had a “fusion” screen for the 3D flanker task, where participants responded when fusion
occurred at the beginning of each trial and then the stimulus display appeared. In this paradigm, the actual stimulus display was very brief and the author ensured that participants experienced the depth location of the target by making sure that fusion occurred before the stimulus was displayed.

The presence of depth references (i.e., other objects or surfaces in the scene), is the second factor that may be important to consider when creating a 3D UFOV. If the 2D UFOV were replicated in depth, the central and peripheral targets would essentially float in space against a very dark background. The only references to depth in that display are the central target and the peripheral target. Spread across the visual field, as they are in the UFOV, they may not give strong enough references to really experience the depth difference because, in a sparse scene, the opportunities for references is inherently sparse as well.

Compared to the literature on attention depth, much of the work has been done using a random dot stereograph (e.g., Andersen, 1990; Andersen & Kramer, 1993). These displays create objects that are perceived to “float out” from a background of white noise (Fricke & Siderov, 1997). Therefore, a background reference is always available to the observer across the entire visual field. Studies that do not use random dot stereograms generally use more complex scenes (e.g., Atchley et al., 1997; Marrara & Moore, 2000; Theeuwes, Atchley, & Kramer, 1998). Complexity (perceptual load) has specifically suggested as something that facilities the influence of depth on attention (Atchley, 2005; Atchley et al., 1997). The UFOV scene, subtended roughly 40° around the central target has comparatively little visual information and therefore is not much in terms of reference or complexity.
Both the temporal aspect of stereoacuity and the benefit of relative depth cues are very important for the development of the 3D UFOV program. These factors must be incorporated because, in order for the depth manipulation to cause any effect, it must be experienced by the observer. If participants are simply able to correctly localize and identify targets without perceiving the depth, the manipulation will in effect not have occurred.

3.1.2 Fusion Limits and the Perception of Depth from Disparity

In addition to facilitating stereopsis, it is also important to determine what depths should be displayed. Stereopsis allows for the perception of depth simply from binocular disparity where the larger the difference in the two retinal images, the larger the perceived depth will be. However, there are limits to how much disparity there can be and, therefore, there are limits to the depth that can be experienced. As reviewed in Section 2.1.2, these limits involved two aspects: stereopsis and fusion. Fusion refers to the ability to combine the two separate images into one and stereopsis is the ability to resolve depth from that fusion. Quantitative stereopsis occurs from smaller disparity (both fused and diplopic) and allows for the perceived depth to vary with disparity. Qualitative stereopsis occurs with diplopia and allows the viewer to simply determine whether one object is behind or in front of the fixation point. This is illustrated in Figure 3 in the Literature Review.

General guidelines for these limits were given by Patterson and Martin (1992), where for large targets (1.0°-6.6°), the limits in the fovea were fusion, patent, and qualitative stereopsis are 20 arcmin, 2°, and 8°, respectively. Using these guidelines and through pilot studies, it was determined that 21 arcmin was an appropriate disparity limit for fusion in the UFOV display. These limits are in line with what has been used by other researchers to study attention in depth.
(e.g., Andersen, 1990; Atchley & Kramer, 2000; He & Nakayama, 1995), which suggests an importance of fusion, not just stereopsis, in the displays.

3.1.3 Simulated Depth with Active Shutter Glasses

As discussed in Section 2.1.2, disparity can be manipulated with active temporally alternating displays, or shutter glasses. With this equipment, slightly different views of the same scene are displayed to each eye by manipulating when each eye is exposed to the screen. This allows for considerable freedom in what is displayed in depth. This is much more freedom than in random dot displays, as depth can only be displayed for objects, not for textures. This can then be relatively easy to implement in 2D programs of the UFOV that already exist.

One component necessary for displaying stereopsis with active shutter glasses is the interpupillary distance, which is the distance between the two pupils (Gordon et al., 2014). The standard distance used in the literature is 6.5 cm (Ogle, 1964; Patterson & Martin, 1992). As with all anthropometric information, there is some variation in the population (Dodgson, 2004); however, any slight differences do not seem to cause large negative effects in performance or create visual fatigue when viewing stereoscopic displays with the standard measure (Rosenberg, 1993).

3.2 UFOV Development and Pilot Testing

3.2.1 First Pilot Test

There are many variations of the UFOV that appear in the literature. All of them generally include detection, localization, or identification of targets among different backgrounds (Ball & Rebok, 1994). The UFOV stimulus that was developed for this dissertation is primarily based on the versions of the divided attention subtest used by Richards et al. (2006) and Plummer and Ni (2014, 2015).
The general UFOV stimulus display consists of two targets that are displayed simultaneously: central and peripheral. As in the version by Richards et al. (2006), the central targets were letters, either “E”, “F”, “H”, or “L”. The peripheral target was a Gaussian windowed sinusoidal grating (i.e., Gabor patch). This is different from the white disk used as the peripheral targets in previous research (Plummer & Ni, 2014; Richards et al., 2006), but the Gabor patch offers flexibility in how the peripheral targets can be manipulated. For instance, orientation is a feature that guides attention (Wolfe & Horowitz, 2004) and therefore the Gabor patch could be easily changed to create targets and distractors.

Based on this, a 2D UFOV test was created as a first step towards a 3D version. In this paradigm, there are five display screens in the UFOV stimulus, and examples can be seen in Figure 8. In the first screen, the text “Ready?” appears until the trial is initiated by the participant via a button press. The second screen is a fixation, where the participant focuses on the center of the screen (where crosshairs are displayed) and annuli act as placeholders of where the peripheral targets may appear. The third screen is the stimulus, showing the central and peripheral stimuli simultaneously. Fourth is the mask, which is followed by two response screens, first for the central target identification and then for the peripheral target localization.
The divided attention task has only a peripheral target. In the present version, the peripheral target is a Gabor patch. Many UFOV tests, including the commercial versions (Edwards et al., 2005), includes an additional subtest of selective attention. In these tests, observers must identify the peripheral target while filtrating out peripheral distractors. The commercial version has a silhouette of a car (the target) and a silhouette of a truck (distractor). For this version, the divided attention tasks were adapted to include distractors; the targets as Gabor patches oriented horizontally and the distractors oriented vertically. Figure 9 shows examples of the two stimulus displays.
Incorporating information found in the literature, two features were added from the base 2D version for the first attempt to create the 3D version. First, to create a reference surface a slight grating was added at the far depth plane (−21 arcmin). Each line was less than 0.05° thick; the lines occurred every 2° from the center both horizontally and vertically. This grating had the effect of there being a “wall” behind the display throughout the entire test and made the central targets appear to pop out. This is similar to the effect that a random-dot stereogram would have, where the shapes that appear in stereoscopic relief in relation to a background (Fricke & Siderov, 1997). Based on informal pilot testing, this created a much stronger feeling of depth than without the background by giving a reference surface to all the UFOV displays. Figure 10 illustrates the three depth planes used in the study with the background as seen during a fixation screen.
Second, the duration of the fixation screen was set at 3 sec. This allowed for a sufficient period of time for optimal stereoacuity, and therefore greatly increased the likelihood that depth is experienced by the observer before the stimulus is displayed. The depth was manipulated only in the peripheral targets and was changed relative to the central target, either on a near depth plane, the same depth plane, or a far depth plane (21, 0, and −21 arcmin, respectively). The change in depth (near, same, far) occurred only on the stimulus display screen.

The original versions of the UFOV test used a dependent measure of the percentage of reduction of some maximum radius of a visual field at a set display duration (e.g., Ball et al.,
Using this dependent variable, the higher performance is indicated by a larger percentage threshold of the total possible area. This first version of the 3D UFOV replicated those studies’ dependent measure by using the distance from the central target, from 4° to 20°, as the dependent measures. After running pilot participants, there was evidence of both floor and ceiling effects. Floor effect and ceiling effects are types of range restriction, which is a threat to statistical conclusion validity (Shadish, Cook, & Campbell, 2001). A ceiling effect occurs when participants cluster at the highest possible score; a floor effect occurs when they cluster at the lowest possible score. These effects mean that the measurement used is not sensitive to validly differential among the participants on the construct being measured (Pedhazur & Schmelkin, 1991).

The ceiling effects were found in the divided attention task, where participants regularly responded correctly on each trial and received the maximum possible threshold. A floor effect was found in the selective attention task, where participants struggled to complete the task and thresholds were regularly at the lowest possible value. These effects remained despite various iterations of piloting with different display durations (ranging from 50 to 150 ms for divided attention and 150 to 300 ms for selective attention subtests). In light of these findings, a change in the approach of how to measure the UFOV was implemented.

3.2.2 Retinal Eccentricity as Dependent Measure

Though the original concept of the “functional” or “useful” field of view in terms of size, recent versions of the test have often tested speed of processing at specific retinal locations (Wood & Owsley, 2014). This can be done where peripheral targets are treated as one value (Richards et al., 2006; Sekuler et al., 2000) or with different retinal eccentricities independently evaluated for thresholds (Plummer & Ni, 2014, 2015). In the latter experiment, Plummer and
colleagues tested separate retinal locations and used the display duration as the threshold. Therefore, a speed-of-processing score at each eccentricity was evaluated. This approach was replicated for the current study, and two retinal eccentricities, 5° and 15° were used for the peripheral targets.

Using the display duration threshold eliminates the problem of finding the correct display duration because the duration adapts to each observer’s processing speed at the given location. The only possible range restriction could be due to the display duration limits of the program. On the one extreme, the range is restricted based on the refresh rate of the screen. The display is set to 120 Hz while doing the temporally alternating display to achieve stereopsis with shutter glasses. This means that for each eye, the refresh rate is 60 Hz, with one frame lasting 16.66 ms. Therefore it is a constraint of the technology that the lowest measurable threshold is 16.66 ms. The commercial UFOV test has an upper limit of 500 ms (Edwards et al., 2005), which gives a wide range of possible threshold values.

In the previous version, the location of the four peripheral target locations moved depending on each participant’s threshold. Because they moved incrementally, only the four annuli could be used. In the current version, the peripheral targets stay in the same location throughout the entire test. Instead of just putting annuli at the two possible peripheral target locations, six additional annuli were added to each quadrant to give the appearance of an “arm” extending from the central target. These annuli were always displayed at the central depth location and provided more depth references. Figure 11 shows an example fixation screen.
Figure 11: The fixation screen showing the array of annuli around the possible peripheral target locations.

A pilot study with this design showed there was no longer a ceiling and floor effect, meaning that this was a better dependent measure to use going forward. Additionally, the participants reported a stronger experience of depth with the extra annuli than they had in the previous version. However, no difference was found across the depth planes and the selective attention task was very difficult for participants. Because of this, it was necessary to re-evaluate the divided attention task.

3.2.3 Depth Placeholders and Changing the Divided Attention Task

Informal observations of the divided attention task suggested that the short duration of the peripheral targets made it difficult to actually resolve the depth of the target. If that were the case and depth was not resolved, it would be as though the depth manipulation did not occur at
all. Since the peripheral target in the divided attention task has no distractors, the participants could essentially be responding to a flash of light in the periphery, and that is enough for them to localize it even though they had not necessarily resolved its depth.

The general difference between a test of divided and selective attention is that divided attention measures the ability to simultaneously complete two tasks, whereas selective attention tasks measure the ability to filter out irrelevant stimuli (Kahneman & Treisman, 1984). In the commercial versions of the UFOV, the selective attention builds upon the divided attention task by adding distractors to the peripheral target. However, any peripheral change that captures attention and affords parallel search should be able to be used as a way to measure divided attention.

Studies that investigate peripheral attention have often used feature changes in objects that are already in the scene instead of the onset of whole new objects: Rogé et al. (2004) used the onset of a light and Andersen et al. (2011) used the change in light color in a light array (screenshots of these displays can be seen in Figure 5 and Figure 6). An important feature of the Gabor patch that could be changed to guide attention is orientation of the gratings (Wolfe & Horowitz, 2004). As a way of making the divided attention tasks more complex, changing the orientation of a Gabor patch could be used as the peripheral target.

To incorporate this idea in the 3D UFOV test, eight vertical Gabor patches were put in the possible peripheral target locations during the fixation screen. These served as the possible target locations. During the stimulus display, one of those eight patches becomes horizontal. Figure 12 shows an example display screen.
Figure 12: Example of the divided attention task where a peripheral target is rotated 90° from the fixation screen (left) to the stimulus screen (right).

The result of the pilot studies and literature review led to the creation of this final version of the UFOV with divided and selective attention subtests. The results of the final pilot study show promise in the ability of this UFOV design to evaluate an effect of depth without significant confounds.
4.1 Purpose

The results of the literature review (see Chapter 2) suggest that depth has an effect on the way that attention is allocated across a scene. This effect should also be present in the useful field of view (UFOV) test. Therefore, the overall purpose of Experiment 1 is to evaluate the effect of depth on the UFOV using the stimulus described in Chapter 3.

Depth can be present in a scene in many different ways and the current UFOV test manipulates the depths of the peripheral targets relative to the central target in three different ways. Experiment 1A evaluated the effect of depth when depth of the peripheral targets and distractors are known before they are displayed for the divided and selective attention subtask. Experiment 1B evaluated the effect of depth when the depth of the peripheral targets is not known beforehand in divided and selective attention subtests. Finally, Experiment 1C evaluated the cost of dividing attention between the central and peripheral targets when the peripheral targets vary in depth.

4.2 Experiment 1A

Experiment 1A evaluated the effect of depth on the UFOV when the depths of the possible peripheral target locations were known before the onset of the target. This is done by displaying the depths of potential target locations during the fixation screen so that participants must divide attention between locations in all three dimensions. This depth manipulation should increase the ecological validity of this test by incorporating how depth is often present in real-world scenes. Oftentimes, real-world scenes involve objects that are in depth, and the critical
cues are some feature change of the objects, such as a traffic light changing color. In this experiment, attention must be allocated through all three dimensions.

Previous research by Andersen and colleagues (Andersen et al., 2011; Pierce & Andersen, 2014) found an effect of depth on attention when the peripheral target was a feature change in an array of lights that appeared at different simulated depths from the observer. Like the depth paradigm used in Experiment 1A, the depths of all of the objects that could become peripheral targets were known before the stimulus actually appeared. In addition, having some sort of object token seems to facilitate and effect the influence of depth in attention (Marrara & Moore, 2000), and therefore makes this paradigm appropriate for the first evaluation on the effect of depth on the UFOV.

The first hypothesis of this experiment (H1) is that there will be an effect of retinal eccentricity, where participants will perform better the closer the peripheral target is to the central target. In general, the UFOV test is more difficult when peripheral targets are displayed farther away from the central target (e.g., Ball et al., 1988; Plummer & Ni, 2014, 2015). If this hypothesis is supported, it would give confidence that this test is actually measuring the UFOV construct. This is important, since the stimulus is different than many traditional tests; see Chapter 3 for more information about the stimulus development.

The second hypothesis (H2) predicts that there will be an effect of depth of the peripheral targets on the display duration threshold in that performance will be worse when the targets are displayed on different depth planes. It was unknown if the two different depth planes (near or far) will have different thresholds or of the depth difference is asymmetrical (e.g., difference when target is behind but not in front of central target). Table 5 summarizes the hypotheses.
Table 5: Summary of Hypotheses for Experiment 1A.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>There will be an effect of retinal eccentricity on the thresholds, with participants exhibiting a higher display duration threshold when peripheral targets are displayed farther from the central target.</td>
</tr>
<tr>
<td>H2</td>
<td>There will be an effect of depth; specifically, with participants exhibiting a higher display duration threshold when peripheral targets are displayed at a different depth plane than the central target (either near or far) then when peripheral and central targets are on the same depth plane.</td>
</tr>
</tbody>
</table>

4.2.1 Method

4.2.1.1 Participants

To determine the number of participants necessary, an *a priori* power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) was used to conduct. Piloting suggested the effect size would be small, and therefore a small eta square was chosen, showing at least 27 participants. For Experiment 1A, 30 participants (21 females, 9 males) aged 18 to 32 years (*M* = 20.2, *SD* = 3.4) were recruited to participant in the study. All the participants were recruited from the Wichita State University SONA Experiment Management System, where they were granted psychology course credit for their time. Before volunteering, they were informed of the eligibility requirements and were tested for visual function prior to the experiment. All participants had normal or corrected-to-normal visual acuity, with all acuity lower than 25/30 as measured by a Snellen acuity chart (Lighthouse Low Vision Product, Long Island City, NY). All participants also showed stereoacuity of at least 50 seconds of an arc, as measured by the Randot Stereo Test (Stereo Optical Company, Inc., Chicago, IL). Additionally, no participant experienced severe stereoscopic discomfort during the session, as measured by symptom questionnaire (Hoffman, Girshick, Akeley, & Banks, 2008). All participants completed the entire study.
4.2.1.2 Design

The first independent variable was the retinal eccentricity of the peripheral target relative to the central target. Two eccentricities were chosen, 5° and 15° from the central target. Only two eccentricities were used to ensure that the experiment could be completed in a reasonable amount of time (i.e., no more than one hour). The specific locations were within standard UFOV areas and did not show floor or ceiling effects while pilot testing.

The second independent variable was the depth of the peripheral targets relative to the central target. The central target was always presented at the same depth (simulated 270 cm), and the peripheral targets could appear at that same depth, or near (simulated 215 cm) or far (simulated 360 cm) depth planes. Distance was simulated through stereopsis, with the disparities of the near and far targets 21 arcmin and −21 arcmin, respectively, and the same depth plane shown at 0 arcmin disparity. Stereopsis was used to reduce the chance of confounds associated with other depth cues (e.g., size, brightness) that can be found when other depth cues are used. In order for there to be adequate reference surfaces, a slight grid was displayed on the far depth plane; refer to Chapter 3 for further information on the UFOV development. Figure 13 shows the depths of the possible peripheral target locations.
Figure 13: The different depths used in the experiment, relative to the central target. The two possible peripheral targets (5° and 15°) are displayed either closer to the observer (A), same as the central target (B), and farther from the observer (C).

4.2.1.3 Apparatus and Stimuli

All displays were generated by customized programs written in C++ and Open Graphics Library (OpenGL) in the Microsoft® Visual Studio® 2008 environment. The stimuli were presented on a Panasonic Real 3D DLP Projector (Model PT-RW430UK) at a resolution of 1024 × 768 and with a refresh rate of 120 Hz. The projector was driven by a NVIDIA Quadro FX1700 graphics card and controlled by a Dell Precision T5400 PC running a Windows XP Professional operating system. NVIDIA 3D Vision (Model P854) active shutter glasses were synchronized with the projector (60 Hz in each eye) to enable the stereoscopic displays. Participants viewed the screen from a distance of 270 cm.

The UFOV was measured using an adaption of the test used by Richards et al. (2006) to test divided and selective attention; the development of this version of the UFOV is fully explained in Chapter 3. The background of the display was dark (.4 cd/m² through glasses) with a slight grid and all objects were displayed in brighter white color (6.9 cd/m² through glasses).

Each trial in each test began with a prompt screen that showed the text “Ready”, and participants were instructed to press the space bar to initiate the trial. After the trial was
initialized, a fixation cross appeared in the center of the screen after the key press for 3000 ms, which is enough time for maximum stereoacuity to occur. The divided attention task consisted of both central and peripheral targets and the selective attention tasks consisted of central and peripheral targets with peripheral distractors. Central targets were one of four randomly selected letters (“E”, “F”, “H”, or “L”) and were displayed in the center of the screen subtending 1.1° of a visual angle. Peripheral stimuli consisted of 36 white circular annuli that each subtended 1.1 of a visual angle, creating an “X” pattern with each arm containing eight annuli in one quadrant surrounding the central target.

Two of these eight annuli on each arm represented possible target locations (5° and 15°). The peripheral targets consisted of a Gaussian windowed sinusoidal grating (i.e., a Gabor patch), oriented horizontally. Distractor Gabor patches were oriented vertically and were displayed at other possible target positions. In addition, a central identification task was presented, simultaneously showing one of four possible letter options (E, F, H, or L) in the center of the screen. The combination of central and peripheral targets was presented in a random order to each participant. The stimulus was displayed for a varying amount of time, depending on the level of threshold of the participants. See Figure 14 for example of the target and distractor Gabor patches.

![Figure 14: The Gabor patches used for peripheral stimuli. The target Gabor patch was always horizontal and the distractor Gabor patches were always vertical.](image)
After the stimulus was presented, a checkerboard mask was displayed for 1000 ms covering all stimuli. For both tests, participants first responded to the central letter task and then the peripheral target location task using a keyboard. Auditory feedback was given to the participant if the response was incorrect.

Divided and selective attention tasks differed by when and how the targets and distractors appeared. These were determined through the extensive literature review and pilot studies described in Chapter 3. For the divided attention task, distractor Gabor patches were presented during the fixation screen. In the stimulus screen, one of the eight distractor patches changed to a target patch. For selective attention, the fixation screen showed only annuli in the possible target locations. In the stimulus screen, the target patch appeared along with three distractor patches at the same retinal eccentricity. Figure 15 shows the sequence for divided attention and Figure 16 shows the sequence for selective attention. In both subtests, the depth of the possible peripheral targets is presented at both the fixation and the stimulus screen in all eight potential target locations (by the Gabor patches in divided attention and annuli in selective attention). The depth of fixation and stimulus screens is shown on Figure 17.
Figure 15: Trial sequence for the divided attention test. The central target is “E” and the peripheral target is “1” (top left).
Figure 16: Trial sequence for the selective attention test. The central target is “E” and the peripheral target is “1” (top left).
Figure 17: The depth presented in both the fixation and the stimulus display. The top of the figure is the divided attention subtask and the bottom is the selective attention subtask. The peripheral target appears in the top left in each example.
Threshold stimulus duration was measured using the best PEST procedure (Lieberman & Pentland, 1982; Pentland, 1980), a maximum likelihood technique that requires fewer observations than standard staircase procedures. Both the divided and selective attention task had six thresholds: two retinal eccentricities (5° and 15°) for three depths (21, 0, and −21 arcmin disparity).

4.2.1.4 Procedure

The experimental session lasted approximately 60 minutes. Participants first completed the consent form and had any questions addressed by the researcher. They were led into the experiment room where they performed both the Snellen acuity test and the Randot stereo test. After administering these screening tests, the lights were turned off so participants could begin dark adaption, during which time they were able to adjust their seat for maximum comfort and were given overall instructions about the session.

The order of tasks (divided or selective attention) was counterbalanced to control for order effects. At the beginning of each task, participants were given a verbal explanation, a demo, and then allowed to practice for a few minutes until they understood the task and demonstrated a least five correct trials in a row. After completing the demo and the practice, they began the experimental task, which lasted approximately 20 minutes. They were instructed to take breaks when needed to avoid visual discomfort or eye strain. After they completed the first task, participants completed the next task in the same way. After completing both UFOV subtasks, participants completed an adaption of the stereopsis symptom questionnaire (Hoffman et al., 2008), which can be found in Appendix A. The entire session lasted approximately 60 minutes in which the researcher remained in the room with the participant the entire time.
4.2.2 Results

Display duration threshold estimates were calculated for each participant under each condition; higher threshold values indicate worse performance. For ease of reading, depth is indicated by the peripheral target’s relationship to the central target (near, same, or far). Two within-subject analysis of variances (ANOVAs) were conducted, one for each of the divided and selective attention tasks. These ANOVAs tested for main effects of retinal eccentricity and depth, and the interaction of retinal eccentricity and depth. For each ANOVA, Mauchly’s test for sphericity was conducted and if the assumption were violated, an appropriate correction was applied (Roberts & Russo, 1999).

4.2.2.1 Divided Attention

A 2 × 3 within-subjects ANOVA was conducted to evaluate the effect of retinal eccentricity and depth on the display duration threshold, which was measured in milliseconds. The main effect of retinal eccentricity was significant, $F(1, 29) = 51.12$, $p < .001$, $\eta^2_p = .64$, where the display duration threshold increased when peripheral targets were displayed farther from the central target; Figure 18 shows the means and standard errors. The main effect of depth was significant, $F(2, 58) = 4.70$, $p = .013$, $\eta^2_p = .14$; Figure 19 shows the means and standard errors. Post hoc analysis showed that the far depth was significantly different from the same depth ($p = .008$), but that neither far and near ($p = .106$) and same and near depths ($p = .141$) were significantly different from each other. The interaction of depth and eccentricity was not significant ($p = .061$).
Figure 18: The main effect of retinal eccentricity on display duration in the divided attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error.

Figure 19: The main effect of depth on display duration in the divided attention task. Performance was worse only when peripheral targets were displayed behind the central target. Error bars represent standard error.
4.2.2.2 Selective Attention

A $2 \times 3$ within-subjects ANOVA was conducted to evaluate the effect of retinal eccentricity and depth on the selective attention task. The main effect of retinal eccentricity was significant, $F(1, 29) = 71.69$, $p < .001$, $\eta_p^2 = .71$, where the display duration threshold increased when peripheral targets were displayed farther from the central target; Figure 20 shows the means and standard errors. The main effect of depth was significant, $F(2, 58) = 4.70$, $p = .013$, $\eta_p^2 = .14$; Figure 21 shows the means and standard errors. Post hoc analysis showed that the far depth was significantly different from the same depth ($p = .006$) and the near depth ($p = .030$); there was no significant difference between same and near depth planes ($p = .518$). The interaction of depth and eccentricity was not significant ($p = .435$).

Figure 20: The main effect of retinal eccentricity on display duration in the selective attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error.
Figure 21: The main effect of depth on display duration in the selective attention task. Performance was worse only when peripheral targets were displayed behind the central target. Error bars represent standard error.

4.2.3 Discussion

Experiment 1A assessed the effect of depth on the UFOV when the depth of the peripheral targets is known before the stimulus display for each trial as well as the effect of retinal eccentricity. First, there was a large effect of retinal eccentricity, which is in line with previous research on the UFOV (Plummer & Ni, 2014, 2015) and supports H1. The replicability of this effect suggests that the new stimulus is measuring the UFOV construct. More importantly, the results suggest that there is an effect of depth, but only when the peripheral target is presented behind than the central target. Therefore, Experiment 1A shows partial support for H2.

This effect of depth is interesting, and that this effect was only evident for the far targets was not expected. Relating to previous work in attention and depth, Andersen et al. (2011) only tested depth that was farther away from the focus of attention (the lead vehicle), but other researchers (e.g., Andersen, 1990; Atchley et al., 1997) found an effect of depth both behind and
in front of the fixation. In the Andersen et al. (2011) paradigm, the depth of the targets was all
known beforehand (as it is in this experiment), so it is unknown if a similar effect would have
been found if the peripheral targets were closer. Additionally, there is some evidence that when
stimuli lie in different depth planes, the ones that are presented closest to the observer is
processed preferentially by the visual system (Lehmkuhle & Fox, 1980), which may explain why
closer stimuli seemed to not be processed slower than the far stimuli. The overall results suggest
that it is more difficult to process targets at depths farther than fixation when these depths of
possible target locations are known before the target is displayed. In other words, it takes more
effort to divide attention across objects located farther than the fixation than if they were at the
same depth as the fixation.

The small effect size of depth ($\eta^2_p = .14$ for both tasks) may be in part because of the
limits of testing depth with stereopsis. Because of the limits on stereoscopic fusion (Ogle,
1952a; Patterson & Martin, 1992), the simulated depth planes for near and far peripheral targets
were relatively small (54 cm and 90 cm, respectively, from the central target). Presenting depths
that were larger would likely have resulted in consistent diplopia and lack of quantitative
stereopsis, and created a significant confound in interpreting the results. The depth of real and
simulated scenes can be much greater when using other depth cues, and the larger the depth is,
the larger the effect may be. For example, the largest monocular depth used by Andersen et al.
(2011) was 60 m, which is many times larger than used in the present experiment. Extending the
depth difference between central and peripheral targets may increase the effect of depth of
known targets. It also may be caused by the scene not being as complex, what has been shown
to be important in depth’s influence on attention (Atchley et al., 1997).
Finally, participants overall did not do well on the selective attention task, with some getting the highest possible threshold (i.e., worst performance) and therefore may have exhibited a ceiling effect. It is possible that this restriction of the range reduced possible differences in performance on the different depth planes. There are two ways of addressing this problem. The first is to make the selective attention task easier (e.g., increasing the difference between target distractor, displaying peripheral targets closer to central target), and therefore lowering the overall threshold. The second way is to increase the highest possible threshold, giving participants a greater chance of perceiving the targets. The current threshold of 415 ms could be extended to 500 ms, which is what the commercial UFOV tests generally use as a limit (Edwards et al., 2005). Using more time than that might facilitate eye movements and therefore not actually measure the UFOV.

Overall, these results confirm that this 3D UFOV test is assessing the underlying UFOV construct because of the effect of retinal eccentricity. Additionally, there is evidence that it is more difficult to allocate attention in depth when peripheral targets are displayed farther away from the focus of attention, even when controlling for size, luminance, and retinal location.

4.3 Experiment 1B

The results of Experiment 1A suggest that when the depth of potential peripheral targets is known before they are displayed, the targets presented farther away from the central depth plane take longer to process. Another aspect of depth that is important to understand is when it is not cued. This is what Experiment 1B attempts to answer.

Whereas Experiment 1A cued participants to the peripheral depth during the fixation period, Experiment 1B positions all peripheral depth cues on the same depth plane as the central
target at that screen, and therefore no depth is cued. The effect of this is the presentation of targets at a completely unknown depth for each trial and may be similar to the onset of sudden objects in a scene, such as hazards occurring in the roadway.

Having objects appear in a display at unknown depth is common in the attention literature. The 3D flanker task used by Andersen (1990) had the distractors appear at different depths. These depths were completely unknown to the participants before each trial, and could appear at one of five different depths relative to the central target. Similarly, the cueing paradigm used by Atchley et al. (1997) specifically cued to incorrect depths before the target was displayed. Because of this, the results of Experiment 1B may be more comparable to other research evaluating attention and depth.

As in Experiment 1A, the first hypothesis (H1) is that there will be an effect of retinal eccentricity on UFOV performance. It is expected that as peripheral targets get farther from the central target, performance will get worse and it will take longer for participants to process them. The second hypothesis (H2) is similar to Experiment 1A as well, in that an effect of depth is expected. Because the depth of the targets is un-cued, the effect in the present study is hypothesized to be greater than in experiment 1A, but still no specific hypothesis is given as to which depths might show a greater effect. Much of the literature suggests that, when there is no depth cued before target presentation, depth in front of and behind a target leads to changes in performance. The different hypotheses are summarized in Table 6.
Table 6: Summary of Hypotheses for Experiment 1B.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>There will be an effect of retinal eccentricity on the thresholds, with participants exhibiting a higher display duration threshold when peripheral targets are displayed farther from the central target.</td>
</tr>
<tr>
<td>H2</td>
<td>There will be an effect of depth with participants exhibiting a higher display duration threshold when peripheral targets are displayed at a different depth plane than the central target (either near or far) than when peripheral and central targets are on the same depth plane.</td>
</tr>
</tbody>
</table>

4.3.1 Method

4.3.1.1 Participants

A power analysis was again conducted to determine the required number of participants. Because depth presented at an un-cued position should have a larger effect than found in Experiment 1A, an effect size of .06 was chosen, showing that at least 18 participants would be required to find an effect. Eighteen participants were recruited to participate, and fifteen completed the study (12 females, 3 males), aged 18 to 32 years ($M = 22.1$, $SD = 5.0$). All the participants were recruited from the Wichita State University SONA Experiment Management System, where they were granted psychology course credit for their time. Before volunteering, all participants were informed of the eligibility requirements and were tested for visual function prior to the experiment. All participants had normal or corrected-to-normal visual acuity, with all acuity lower than 25/30 as measured by a Snellen acuity chart (Lighthouse Low Vision Product, Long Island City, NY). All but three participants also showed stereoacuity at 50 seconds of an arc, as measured by the Randot Stereo Test (Stereo Optical Company, Inc., Chicago, IL); those three participants did not continue the study. Additionally, no participant experienced severe stereoscopic discomfort, as measured by symptom questionnaire (Hoffman et
al., 2008). All participants completed the entire study and were compensated with class credit for courses in the Psychology Department.

4.3.1.2 Design

The research design and variables for Experiment 1B were identical to Experiment 1A.

4.3.1.3 Apparatus and Stimuli

The apparatus and UFOV stimulus for Experiment 1B were identical to Experiment 1A with the following exceptions. In order to test the effect of an unknown target depth, the fixation screen always displayed the peripheral annuli at the same depth as the central target. During the stimulus display, the targets appeared in one of the three relative depths. To show the range of possible display depths, one non-target annuli was always in front of the target and one was always behind the target on each of the four quadrants. Figure 22 shows an example.
**Figure 22:** The depth presented in both the fixation and the stimulus display. The top is the divided attention subtask and the bottom is the selective attention subtask.

### 4.3.1.4 Procedure

The procedure for Experiment 1B was identical to Experiment 1A.
4.3.2 Results

The data for Experiment 1B was analyzed the same way as it was for Experiment 1A.

4.3.2.1 Divided Attention

A 2 × 3 within-subjects ANOVA was conducted to evaluate the effect of retinal eccentricity and depth on the divided attention task. The main effect of retinal eccentricity was significant, \( F(1, 14) = 13.59, p = .002, \eta^2_p = .49 \), where the display duration threshold increased when peripheral targets were displayed farther from the central target; Figure 23 shows the means and standard errors. The main effect of depth was significant, \( F(2, 28) = 13.12, p < .001, \eta^2_p = .48 \); Figure 24 shows the means and standard errors. Post hoc analysis showed significant differences between far and same \( (p = .002) \) and near and same \( (p < .001) \); no differences were found between near and far \( (p = .838) \).

![Figure 23: The main effect of retinal eccentricity on display duration in the divided attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error.](image-url)
Figure 24: The main effect of depth on display duration in the divided attention task. Performance was worse when peripheral targets were displayed at a different depth than the central target. Error bars represent standard error.

4.3.2.2 Selective Attention

A 2 × 3 within-subjects ANOVA was conducted to evaluate the effect of retinal eccentricity and depth on the selective attention task. The main effect of retinal eccentricity was significant, $F(1, 14) = 37.28, p < .001, \eta_p^2 = .73$, where the display duration threshold increased when peripheral targets were displayed farther from the central target; Figure 25 shows the means and standard errors. Neither the main effect of depth ($p = .279$) nor the interaction of eccentricity and depth ($p = .278$) were significant.
Figure 25: The main effect of retinal eccentricity on display duration in the selective attention task. Performance was worse when peripheral targets were displayed farther from the central target. Error bars represent standard error.

4.3.3 Discussion

Experiment 1B showed an effect of retinal eccentricity for both tests, which supports H1. Experiment 1B shows two very different results for the divided and selective attention, therefore giving mixed results for H2. For divided attention, there is a very large effect of depth, where peripheral targets displayed either closer or farther from the central target were more difficult to process. The interaction of depth and eccentricity, although significant, is likely because of a floor effect for the larger eccentricity. For the selective attention task, there did not appear to be any effect of depth.

It was surprising that the selective attention task did not show an effect of depth; however, there are a number of possible explanations. Primarily, these explanations are from the general limitations of using stereopsis to display depth. One important limitation has to do with the temporal aspects of stereopsis. As discussed in Chapter 3, it takes approximately 3 sec to correctly perceive depth from disparity. Shorter times can lead to effective fusion, but the
perception of depth is limited (Boff & Lincoln, 1988; Ogle, 1964). In Experiment 1A, the selective attention test cued for depth by presenting annuli as placeholders for the possible targets. The targets always appeared at the cued depth, and therefore the fixation period ensured disparity was displayed for at least 3 sec and the target appeared at that depth. In Experiment 1B, the depth was not cued. Because of the very short (less than 415 ms) presentation of the peripheral target, they may not have been perceived in depth but still able to perceive the orientation (to discriminate target or distractor). This type of limitation has been found by other researchers who have investigated the conditions necessary for attention to be affected by stereoscopic depth (e.g., Marrara & Moore, 2000).

In the divided attention task, there was a large effect of depth ($\eta_p^2 = .48$). This suggests that when individuals are attending to one depth, it takes more time to react to peripheral targets presented at an unexpected depth. This was the anticipated result and fits well with previous research showing an effect of depth on the allocation of attention. Anecdotally, there was a very vivid perception of depth when all four Gabor patches change depths. This suggests that limitations encountered by selective attention would not be present in divided attention, as the depth was more likely to be experienced in this task.

Overall, Experiment 1B adds to what was found in Experiment 1A in suggesting that there is an effect of depth from the UFOV to some degree. Particularly, when depth can be presented with stereopsis, there seems to be a very large performance decrement when targets appear on a depth plane that is different from the central target. Related to real-world scenarios, this is likely similar to unanticipated objects appearing. An example of this could be a driver attending to a car in front of him or her and an unanticipated hazard appearing in front of the car, and much closer than the lead vehicle.
4.4 Experiment 1C

Taken together, the results of Experiment 1A and 1B suggest that depth has an effect on how attention is allocated under some conditions. The effect of depth in the divided attention subtest for Experiment 1B was very large ($\eta_p^2 = .48$), and invites further study. In that paradigm, attention is divided between different depth planes. Since this is a divided attention task, an important area for further study is the cost of dividing attention in depth.

The cost of dividing attention (i.e., having to complete two tasks at once) has been found in many different tasks and conditions (e.g., Ikeda & Takeuchi, 1975; Leibowitz & Appelle, 1969; Sekuler & Ball, 1986). Despite this, the cost is generally not investigated in studies of the UFOV as they often only measure under divided or selective attention conditions. A UFOV paradigm was developed by Sekuler and colleagues (Richards et al., 2006; Sekuler et al., 2000) to evaluate this effect. In these studies, participants are tested with a subtest of divided attention and also a subtest that tests the focused attention on just the peripheral displays. The cost of dividing attention is determined by subtracting the scores of the peripheral focused attention subtest from the divided attention subtest.

This paradigm is adapted to the divided attention subtest of the 3D UFOV. Two additional changes were made to simplify the test. First, only one eccentricity for peripheral targets was tested ($10^\circ$, which is the in between the two used in Experiments 1A and 1B). The results of both experiments 1A and 1B strongly show this effect, so there is no need to test it again. Additionally, only having one peripheral condition significantly reduces the time necessary to administer the test. The second change is that it is displayed on a monitor instead of a projector. Changes in viewing distance in this way should not have an effect on the perception
of depth (Boff & Lincoln, 1988), so the findings should be replicated. This change also makes it much more versatile in where this test can be administered.

The first hypothesis (H1) is that there will be an effect of depth on the simplified divided attention task. The second hypothesis (H2) is that there will also be an effect of depth for the focused attention task. This would be consistent with previous work looking at focused attention in depth (Andersen, 1990; Andersen & Kramer, 1993). The final hypothesis is that there will be a cost associated with divided attention, and it will be greater when the peripheral targets occur at a different depth than the central target (H3). The different hypotheses are summarized in Table 7.

Table 7: Summary of Hypotheses for Experiment 1C.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>There will be an effect of depth with participants exhibiting a higher display duration threshold when peripheral targets are displayed at a different depth plane than the central target (either near or far) then when peripheral and central targets are on the same depth plane.</td>
</tr>
<tr>
<td>H2</td>
<td>There will be an effect of depth with participants exhibiting a higher display duration threshold when peripheral targets are displayed at a different depth plane than the fixation point (either near or far) then when peripheral target appears at the same depth plane as the fixation point.</td>
</tr>
<tr>
<td>H3</td>
<td>The cost of dividing attention will be greater when the attention is divided in depth.</td>
</tr>
</tbody>
</table>

4.4.1 Method

4.4.1.1 Participants

Because the effect size was larger than anticipated in Experiment 1B a larger eta-squared (.12) was chosen for a power analysis for this experiment; eight participants were necessary to find an effect. Eight participants (4 female, 4 male), aged 23 to 34 years ($M = 28.9, SD = 2.8$) participated in the study. All the participants were recruited from the greater Wichita community.
and were compensated $15 for their time. This participant pool was used because classes were not in session and therefore the student pool typically recruited through SONA was unavailable. All participants had normal or corrected-to-normal visual acuity, with all acuity lower than 25/30 as measured by a Snellen acuity chart (Lighthouse Low Vision Product, Long Island City, NY). All participants also showed stereoacuity at 50 seconds of an arc, as measured by the Randot Stereo Test (Stereo Optical Company, Inc., Chicago, IL).

4.4.1.2 Design

The 3D UFOV test was adapted from the divided attention subtest of Experiment 1B. The only difference in the stimulus display was that instead of having conditions for the retinal eccentricity of the target (at 5° and 15° of a visual angle), the Experiment 1C version had only one retinal eccentricity (at 10° of a visual angle).

4.4.1.3 Apparatus and Stimuli

The UFOV was generated by customized programs written in C++ and Open Graphics Library (OpenGL) in the Microsoft® Visual Studio® 2008 environment. The stimuli were presented on an ASUS VG248QE LCD 3D monitor at a resolution of 1024 × 768 and with a refresh rate of 120 Hz. The projector was driven by a NVIDIA Quadro FX1700 graphics card and controlled by a Dell Precision T5400 PC running a Windows XP Professional operating system. NVIDIA 3D Vision (Model P854) active shutter glasses were synchronized with the projector (60 Hz in each eye) to enable the stereoscopic displays. Participants viewed the screen from a distance of 45 cm. Adjustments to the program were made so that the retinal size of all aspects of the display would be consistent to that of experiments 1A and 1B.

The divided attention subtest was similar to that used in Experiment 1B. The only difference is that peripheral targets were displayed at one retinal eccentricity instead of two.
Like Experiment 1B, the depth of the peripheral targets was unknown before the stimulus display. At the fixation screen, four Gabor patches were presented at the same depth plane as the central target. Figure 15 shows the 2D representation and Figure 27 shows the 3D representations of the display.

Figure 26: The fixation (left) and stimulus (right) display for the peripheral-focused attention subtest. The central target is “E” and the peripheral target appears in the top left.
Figure 27: The fixation (left) and stimulus (right) display for the divided attention subtest. The central target is “E” and the peripheral target is top left.

The focused peripheral attention task was similar to the divided attention task except that there was no central target. Otherwise, the trial sequence was the same. Figure 28 shows the fixation and stimulus display.
Figure 28: The fixation (left) and stimulus (right) display for the divided attention subtest. The target appears in the top left.

4.4.1.4 Procedure

The experimental session lasted approximately 30 minutes. As in the previous experiments, participants first completed the consent form and had any questions addressed by the researcher. They were led into the experiment room where they performed both the Snellen acuity test and the Randot stereo test. After administering these tests, the lights were turned off so participants could begin dark adaption, during which time they were able to adjust their seat for maximum comfort and were given overall instructions about the session.

The order of tasks (divided or peripheral focused attention) was counterbalanced to control for order effects. At the beginning of each task, participants were given a verbal explanation, a demo, and then allowed to practice for a few minutes until they understood the task. After completing the demo and the practice, they began the experimental task, which lasted approximately 10 minutes. They were instructed to take breaks when needed to avoid visual discomfort or eye strain. After they completed the first subtest, participants completed the next
in the same way. Since no participants experienced severe visual discomfort or eye strain in experiments 1A and 1B, the stereopsis symptom questionnaire was not administered. The entire session lasted approximately 30 minutes.

4.4.2 Results

4.4.2.1 Divided Attention

A one-way within-subjects ANOVA was conducted to evaluate the effect of depth on the divided attention subtest. The effect of depth was significant, \( F(2, 14) = 8.6, p = .004, \eta_p^2 = .55 \). Post hoc tests revealed that thresholds for the near \( (p = .032) \) and far \( (p = .008) \) depth conditions were significantly higher than the same depth condition; no significant difference was found between near and far \( (p = .250) \). Figure 29 shows the means and standard errors for this effect.

![Figure 29: The effect of depth on divided attention. Near and far conditions had significantly higher thresholds than the same depth condition. Error bars represent standard error.](image)

4.4.2.2 Peripheral Focused Attention

A two-way within-subjects ANOVA was conducted to evaluate the effect of retinal eccentricity and depth on the divided attention task. The main effect of depth was significant,
$F(1.19, 8.37) = 9.33, p = .003, \eta^2_p = .57$. *Post hoc* tests revealed that thresholds for the near ($p = .018$) and far ($p = .015$) depth conditions were significantly higher than the same depth condition; far and near were also significantly different ($p = .044$). Figure 30 shows the means and standard errors for this effect.

![Figure 30: The effect of depth on peripheral focused attention. The threshold for the near depth condition was significantly higher than the same depth condition, and the far depth condition was significantly higher than both the near and the same condition. Error bars represent standard error.]

### 4.4.2.3 Attention Cost

Attention cost was calculated by subtracting the thresholds of the peripheral-focused attention from those of the divided attention for each condition (Richards et al., 2006). A one-way within-subjects ANOVA was conducted to evaluate the effect of depth on the cost of dividing attention. The effect of depth was significant, $F(2, 14) = 4.7, p = .028, \eta^2_p = .40$. *Post hoc* tests revealed that thresholds for the far and near depth conditions were significantly different ($p = .027$); no significant difference was found between near and far ($p = .681$) or near and same ($p = .062$). Figure 31 shows the means and standard errors for this effect.
Figure 31: The effect of depth on the cost of dividing attention. Near and far conditions had significantly higher thresholds than the same depth condition. Error bars represent standard error.

4.4.3 Discussion

Even though this version of the UFOV was tested on a monitor at much closer viewing distance, an effect of depth was found for both the divided and focused peripheral subtests, supporting H1 and H2. What is most interesting is that there was an effect of depth on the cost of dividing attention, supporting H3.

The results of this experiment support the idea that there is a cost associated with having to divide attention between different depths. This suggests that at least part of the effect of depth is because of dividing attention, not just because peripheral targets appear at an un-cued depth. A component of the depth effect is because the participant must divide attention between the two locations.
4.5 General Discussion

All three experiments found support for the general research question that depth plays a role in how attention is allocated in the useful field of view. Experiment 1A found that when observers are given the depth beforehand, it takes longer to correctly resolve peripheral targets that occur farther than the central target. In Experiment 1B, the divided attention task found large effects of depth (both closer and farther) but the selective attention task did not. Finally, Experiment 1C demonstrated that it was costly to divide attention in depth. Taken together, it suggests there is an overall influence of depth on the allocation of attention as measured by the UFOV, even if it is not completely consistent in how the influence occurs. Specific limitations to each experiment were discussed in their respective sections, but general limitations that are applicable to both studies, as well as areas of future research, are discussed below.

For both experiments, only three depths were used in relation to the central target: near, same, and far. The depths were chosen because they were the largest disparities that could be reliably fused (i.e., the largest depths), and therefore had the most likely change of finding an effect. Additionally, only three depths were chosen to keep each experimental session a reasonable duration. Because of this, it is not known what the effects of any intermediate depths are on attention. Future research should address this knowledge gap, as it would be particularly useful to understand at what depth there is a decrease in processing abilities or if there is a linear effect (e.g., processing ability gets worse the farther away). In addition to answering basic questions, this could help in determining how to configure future versions of the 3D UFOV. If an effect can be found at less extreme depths, this may reduce any type of floor effect found—particularly with selective attention (i.e., more difficult) tasks.
Similarly, the depth of the distractors always appeared at the same depth as the peripheral target. This simplified the experiment and the interpretation significantly, but it is unknown what effect there would be if there were more than one depth plane used relative to the central target. This could offer interesting lines of future research for a number of reasons. First, it would be worthwhile to see how adding different depths affects how attention is divided. Is it more costly to divide attention between three different depth planes than it is for two, or is the effect only present when there is more than one? This could have large applications, because natural scenes require depth to be divided between objects of many depths. A second avenue is looking for how well the target pops out or how well distractors can be filtered. There is evidence that depth is a feature that guides attention (Wolfe & Horowitz, 2004), and this UFOV paradigm would make it relatively easy to see how the depth is affected by the depth of distractors.

Another limitation of this experiment is the use of stereopsis as the only cue with which observer’s perceived depth. The advantage is that other depth cues may create confounds (e.g., differences in size or luminance), whereas stereopsis keeps all of these other aspects constant across depths. However, as mentioned previously, using stereopsis limits how much depth can be presented because of the fusion limits. Adapting the recommendations by Patterson and Martin (1992) to the UFOV through pilot data, this limited the depths displays to 21 arcmin relative to the central target.

A way to increase the simulated depths would be to use other depth cues. One cue in particular, the ground plane or surface extending from the viewing position to objects in a scene, seems to have a strong effect on the perception of depth (Bian & Andersen, 2010; Bian, Braunstein, & Andersen, 2005; Thompson, Dilda, & Creem-Regehr, 2007). Presenting the
UFOV with that might enable larger depths to be tested. For instance, creating a 3D space in which the UFOV appears with a ground plane—almost a “room”—could easily show different depths. This 3D “room” has been used to investigate the effect of depth in the multiple object tracking paradigm (Weber & Ni, 2015) and could easily be applied to the UFOV. In addition to increasing the possible depth, using other cues would reduce other restraints of using stereopsis (e.g., required duration, stereoblindness) and would therefore make the 3D UFOV more adaptable to different research questions and a larger population.

One of the primary reasons for testing the effect of depth was because depth occurs in the real world, and the UFOV is attempting to capture vision as it applies to everyday tasks. With the results supporting the general hypothesis that depth has an effect, it is important to understand how that aspect of attention actually relates to everyday life. This can be done in two ways: by continuing to add to the ecological validity of the UFOV and by relating it to actual real-world tasks.

In terms of creating a more ecologically valid test, adding depth to the UFOV test was important because depth occurs in the real world and has an effect on attention. But depth does not just occur with stereopsis, as the other depth cues are used to understand the 3D layout of scenes. In fact, it has been suggested that the specific functional use of stereopsis is limited to specific types of task, and that most tasks—even complex ones like driving (McKnight, Shinar, & Hilburn, 1991; Troutbeck & Wood, 1994) and flying (Snyder & Lezotte, 1993)—can be performed without the use of binocular depth cues (see Fielder & Moseley, 1996 for detailed discussion). Further understanding in how depth is allocated as it processed naturally would increase the ecological validity of the test. As mentioned above, this could be done by using other depth cues.
The second important area of future research is to address what aspects of the real-world actually involve the depth component that was captured in the 3D UFOV. In the application of driving, which is the most common application for the UFOV, depth plays an important role. For instance, a roadway scene is complex and objects in it appear at many different depths, such as hazardous objects that must be processed and responded to very quickly in order to avoid a crash. The ability to allocate attention across different depths (as in Experiment 1A) and respond to objects at unexpected depths (as in Experiment 1B). Experiment 2 (Chapter 5) assesses whether the divided attention measure used in Experiment 1C can predict performance on a video-based hazard perception test.

Beyond the 3D UFOV’s relationship with hazard perception addressed in this dissertation, there are likely other applications to real-world tasks. A recent meta-analysis has found that that the UFOV is related to many cognitive functions (Woutersen et al., 2017), and some of these areas, such as spatial ability and general attention may have considerable depth components. Because of these many facets of the UFOV, it could likely be applied to any number of real-world tasks beyond driving that have important depth components. It would be worthwhile to assess the relationship between the effect of depth in those tasks and that in the UFOV, to see if the 3D UFOV is capturing something specific and unique.

Overall, the results of Experiment 1 suggest that depth is an important aspect on how attention is allocated in the UFOV. This experiment serves an important role of increasing the ecological validity of the test, and lays the groundwork for many future areas of research.
CHAPTER 5
EXPERIMENT 2

5.1 Purpose and Hypotheses

The results of Experiment 1 suggest that depth has an effect on the allocation of attention in the UFOV. The natural extension of that experiment is to assess how the depth component of the UFOV relates to attention in the real-world tasks. Applying the UFOV to the real-world tasks is important for two primary reasons. First, the UFOV test was created in order to fill the gap of visual tests that were not sensitive to everyday activities (Ball & Owsley, 1993); the main goal of developing this different type of visual attention measure was to capture how attention was used for common activities. Second, since real-work scenes have depth inherent to them, capturing that aspect in the UFOV test should make it more ecologically valid, and therefore more appropriate in understanding attention in the real world. The everyday task that the UFOV is most often associated with is driving. In order to address the applicability of the 3D UFOV, Experiment 2 evaluates the ability of the test to predict an important aspect of driving: hazard perception. This experiment is very exploratory in nature, and the results lay the groundwork for future work relating visual attention functions (like the UFOV) to driving performance.

In Experiment 1, five different UFOV stimuli were created and tested; two for each Experiment 1A and 1B, and one for 1C. The divided attention test when the depth of peripheral targets was unknown before the stimulus display (1B and 1C) showed the largest effect of depth and variance. If the effects replicate in a simulated driving task, this version would have a greater chance of showing the importance of depth when applying it. Specifically, the stimulus used in Experiment 1C was simplified in that it only presented the peripheral targets at one retinal eccentricity and was displayed on a monitor. This simplification meant that the test took
less time to administer and was more versatile in where it could be displayed; for these reasons it was chosen to be used as the 3D UFOV test for Experiment 2. Even though Experiment 2 had a much larger sample size, it was predicted that the effect of depth found in Experiment 1C would be replicated (H1).

A video-based hazard perception (HP) test developed by Burge and Chaparro (2018) was used to evaluate the application of the 3D UFOV to HP. This test uses fuzzy signal detection theory (FSDT) to evaluate the ability of participants to correctly identify hazards in natural scenes in city and highway environments. The use of FSDT allows for the hazards to be classified along a continuum instead of as discrete, binary states; this is more representative of how they appear in the real world (Wallis & Horswill, 2007). See Section 2.2.4 for more information on the use of FSDT in HP tests. In addition to the two traditional SDT metrics (i.e., sensitivity and response criterion), the HP also gives a hazard response lag, which is the time between participant response and the onset of the hazard, as a third variable.

The HP videos were non-staged driving scenes filmed from the driver’s point of view, and ended before the driver would reach a hazardous event. Each video was viewed by driving instructors to evaluate the potential of a collision after the video is ended if the driver continued on the current curse (i.e., no evasive action was taken). This hazard potentiality was scored on a continuum (0% to 100%) to score the likelihood of that collision, with high inter-rated reliability (ICC = .97, .71, and .77). Using experts (i.e., trained driver education instructors) to rate the scenes establishes some validity of the test. Since a major goal of this study is to understand the relationship between UFOV and crash rates, using a validated test is very important and was part of the rationale for choosing this method of testing HP.
These three HP scores (sensitivity, response criterion, and response lag) can be compared to the UFOV scores to determine if there is a relationship. Specifically, it was predicted that 2D UFOV would relate to the HP scores (H2), with the 3D components accounting for some additional variance (H3). It is likely that sensitivity would be the most related to UFOV scores, since sensitivity generally measures a bottom-up process, whereas response criterion represents a top-down process (Wickens et al., 2004). It is also predicted that those with a lower display duration threshold in the 3D UFOV will be able to respond to the hazards more quickly, and therefore have a shorter response lag.

Finally, because of the large sample, two additional components can be tested. First participants with a wide range of driving experience, as measured by years having a driver’s license, were recruited. It is predicted that driving experience will predict HP scores (H4). Specifically, it is hypothesized that response lag will be shorter for those with more driving experience and that they will exhibit a more liberal response bias, which would replicate earlier work on HP (e.g., Horswill & McKenna, 2004; Wallis & Horswill, 2007). Comparing the score of experts and novices is a common method of validating a HP test, where more experienced drivers should do better at that task (Horswill & McKenna, 2004). Finding the hypothesized relationship would add to the overall validity of the Burge and Chaparro’s (2018) test. The different hypotheses are summarized in Table 8.
Table 8: Summary of Hypotheses for Experiment 2.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Display duration thresholds on the depth conditions of the 3D UFOV will be higher than the no-depth condition.</td>
</tr>
<tr>
<td>H2</td>
<td>2D UFOV will predict HP scores, where higher thresholds UFOV (i.e., worse performance) being related to lower HP scores. Specifically, it is hypothesized that the 2D UFOV should be related to both sensitivity and response lag.</td>
</tr>
<tr>
<td>H3</td>
<td>3D UFOV will predict HP scores, where higher thresholds UFOV (i.e., worse performance) being related to lower HP scores. Specifically, it is hypothesized that the 3D UFOV should be related to both sensitivity and response lag, and add some predictive power to the 2D UFOV.</td>
</tr>
<tr>
<td>H4</td>
<td>Driving experience will predict HP scores, where those who have been driving longer will score better on the HP test. Specifically, more experienced drivers will have a more liberal response bias and shorter response lag.</td>
</tr>
</tbody>
</table>

5.2 Method

5.2.1 Participants

Because of the medium to high effect size related UFOV to HP and SA in the literature (Chaparro et al., 1999; Kaber et al., 2016), $r = .4$ was used in a power analysis to calculate the required number of participants. Forty-six participants were required for the study to have sufficient power. Fifty-five participants were recruited for this study from Wichita State University or the greater Wichita community, and were either granted psychology course credit or, if not a student, $15 for their time. Before volunteering, they were informed of the eligibility requirements and were tested for visual function prior to the experiment. All participants had normal or corrected-to-normal visual acuity, with all acuity lower than 25/30 as measured by a Snellen acuity chart (Lighthouse Low Vision Product, Long Island City, NY). Eight participants did not pass the stereoacuity test of at least 50 seconds of an arc, as measured by the Randot Stereo Test (Stereo Optical Company, Inc., Chicago, IL) and were therefore ineligible to continue with the study. After the screening tests, 48 participants (19 males; 29 female) age 18 to 40 ($M = 24.8; SD = 6.3$) completed the study. These participants had a wide range of driving
experience, from less than one year to 24 years (\( M = 8.7; \ SD = 6.4 \)). All of these participants completed the entire study.

5.2.2 Design

The study had three predictor variables and three criterion variables. The three predictor variables were the display duration thresholds at the three depths of the peripheral targets in the 3D UFOV. The stimulus for the UFOV was identical to that used in Experiment 1C; specific information can be found in Section 4.4.

The three criterion variables represent the three hazard perception scores used by Burge and Chaparro (2018). The first two represent fuzzy signal detection metrics of sensitivity and response criterion (see 2.5.4 for description) that compared expert ratings of hazard potentiality of each clip to the participants’ ratings. In addition to the FSDT metrics, a hazard response buffer was used as the third criterion variable, which represents the time from the end of the clip (i.e., the onset of a hazard) from the participant’s response. The original study divided the clips by environment (city vs. highway), and that delineation was kept in this study.

5.2.3 Apparatus and Stimuli

The 3D UFOV test was the same as used in Experiment 1C. See Section 4.4 for specific information.

The hazard perception test was adapted from Burge and Chaparro (2018). Using their database of expert-rated videos, 20 highway and 20 city videos were selected. Because this study had significantly different research questions than the Burge and Chaparro (2018) study, these 40 videos had to be carefully selected to have a similar distribution of expert-rated hazard potentiality and length as the original study to ensure the results could be comparable (R. J.
Burge, personal communication, October 25, 2018). A sample screenshot of one of the videos can be seen at Figure 32.

![Figure 32: Screenshot of a city environment from the HP test. The top shows a highway environment and the bottom shows a city environment.](image)

The videos were displayed through Experiment Builder (SR Research); that program also accepted user input throughout the study. The stimuli were presented on a 64-inch LG OLED
65B7A 4k Ultra HD TV at a resolution of 1920 × 1080 at a refresh rate of 60 Hz. The monitor was driven by an NVIDIA GeForce GTX 1080 Ti graphics card and controlled by a Windows 10 Enterprise operating system. Participants sat 140 cm away from the display.

5.2.4 Procedure

The experimental session lasted approximately 60 minutes. Participants first completed the consent form and had any questions addressed by the researcher. They were led into the experiment room where they performed both the Snellen acuity test and the Randot stereo test. Those who passed the screening procedures were shown to the 3D UFOV test. At this time, the experimenter turned off the lights so participants could begin dark adaption while they adjusted their seat for maximum comfort and were given overall instructions about the session. All participants first completed the 3D UFOV test; first with a demo, practice, and then experimental session. They were instructed to take breaks when needed to avoid visual discomfort or eye strain. After they completed the experimental version, the experimenter turned on the lights and participants moved to the HP test position.

Participants were first given a verbal overview of the study and then followed a series of onscreen instructions. They were instructed to watch the videos as if they were driving in the real world and respond to any traffic conflict by pressing the space bar. A traffic conflict was defined as a situation in which a collision with another road use would occur unless some time of evasive action (e.g., slowing down, steering) was taken by the driver. The instructions emphasized that participants should respond as soon as they saw a traffic conflict.

After responding to a traffic conflict, the video would end and the participant would rate their confidence that a collision would occur without intervention on a scale of 0 to 100%. If they did not respond to a trial, the video would end and the participant would rate their
confidence that a response would not have been required. Participants were also asked to identify the traffic conflict to which they were responding, if a response were made. These post-trial questions were administered by the computer program. See Appendix B for the post-trial questions.

After completing the instructions, participants went through a practice trial with the experimenter to ensure it was understood. The 40 videos were presented in a random order, and each trial was initiated by the participant so he or she could take brakes when appropriate. After the HP was completed, participants completed a driving history questionnaire. This questionnaire, which was adapted from a variety of driving-related questions (Morris, Cooper, Ton, Plummer, & Easterlund, 2016; Owsley, Stalvey, Wells, & Sloane, 1999; Reimer et al., 2005), can be seen in Appendix C. After completing the questionnaire, the participants were thanked for their time and the session ended.

5.3 Results

5.3.1 Useful Field of View

The UFOV display duration thresholds were the best PEST procedure (Lieberman & Pentland, 1982; Pentland, 1980), and analyzed with a between-subjects ANOVA. The effect of depth was significant, $F(2,94) = 155.9, p < .001, \eta_p^2 = .77$. Post hoc tests showed that both the near and far depth planes were significantly different than the same depth plane ($p < .001$); no difference was found between near and far ($p = .922$). Figure 33 shows the means and standard errors.
5.3.2 Hazard Perception Scores

The Burge and Chaparro (2018) HP test gives three metrics. The first was hazard response lag. To meet the assumption of normalcy, the log10 of the response lag was used. The FSDT metrics were used to evaluate the specific hazard response behavior, where the expert ratings represented the state of the world and the participants’ confidence ratings represented the response. Membership to the SDT metrics (i.e., hit, miss, false alarm, and correct rejection) was calculated using formulae proposed by Parasuraman et al. (2000). These membership values were then summed to return standard measures of hits and false alarms. Two measures, sensitivity, and response criterion can be calculated from this data. Sensitivity refers to how good a participant was at discriminating a signal from the background noise. Higher sensitivity means that there are more correct responses and fewer errors. Response criterion refers to a bias the participant had at selecting there was a signal instead of no signal. Positive response criterion values reflect a more conservative bias and negative reflect a more liberal bias (Wickens et al., 2004). Because of the relatively low number of trials, nonparametric values of
sensitivity \( (A') \) and response criterion \( (B'') \) were used given the relatively low number of trials (See et al., 1997). Appendix D shows the calculations of these scores. The assumption of normal distribution was assessed using the Kolmogorov-Smirnov statistic for the log10 response buffer, sensitivity, and response criterion; this assumption was met for all three HP metrics \( (p > .05) \).

The first statistical test was to determine if there were differences between city and highway clips in the three metrics. To determine this, three paired-samples \( t \) tests were used to evaluate the differences between environments. There was a significant difference in response lag, \( t(47) = 6.5, p < .001 \), indicating that there was a longer response lag in the highway environment; Figure 34 show the means and standard errors. There was also a significant difference in response criterion, \( t(47) = -5.9, p < .001 \), showing that there was a more liberal bias in highway environment; Figure 35 shows this. There was not a difference in sensitivity \( t(47) = -0.03, p = .97 \). In light of these findings, the subsequent analyses treat the city and highway environment differently for response lag and response criterion, whereas for sensitivity an overall category is used. This leads to five separate HP variables: city response lag, highway response lag, city response criterion, highway response criterion, and overall sensitivity.
5.3.3 Driving Experience Predicting Hazard Perception

To build the validity of the HP test, driving experience was evaluated in terms of predicting HP performance. In previous work, categorical variables (e.g., “expert” vs “novice” drivers) have been used for this. In the present study, continuous variables of experience, as
measured by years with a driver’s license, were used to reduce the problems with artificially creating categorical variables (Young, 2016).

Pearson’s correlation coefficients were calculated for the response buffer, finding significant correlations between driving experience and city environment ($r = .39$, $p = .007$), but not highway environment ($r = .21$, $p = .158$). A linear regression was conducted to evaluate the prediction of HP response lag in city environment from driving experience. The regression equation is: $Response\ Buffer = (0.012)\*Driving\ Years + .554$. Approximately 15% of variance of the HP scores was accounted for by its linear relationship with driving experience. Figure 36 shows a scatterplot with the regression line.

![Figure 36: Relationship between driving experience and response buffer. The dotted line represents the regression line.](image)

No significant correlation between driving experience and the FSDT metrics were found. Table 9 shows these results.
Table 9: Pearson Correlation Coefficients between Driving Experience and FSDT Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>$r$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (Overall)</td>
<td>.021</td>
<td>.89</td>
</tr>
<tr>
<td>Response Criterion (City)</td>
<td>−118</td>
<td>.43</td>
</tr>
<tr>
<td>Response Criterion (Highway)</td>
<td>.108</td>
<td>.47</td>
</tr>
</tbody>
</table>

5.3.4 UFOV Predicting Hazard Perception

Correlations were calculated between the three UFOV threshold measures (near, same, and far) and the five HP scores. No significant correlations were found. Table 10 shows the correlations and significant test results.

Table 10: Pearson Correlation Coefficient and Significance values for the UFOV Thresholds and HP scores

<table>
<thead>
<tr>
<th>UFOV Statistic</th>
<th>Hazard Buffer (City)</th>
<th>Hazard Buffer (Highway)</th>
<th>Sensitivity (Overall)</th>
<th>Response Criterion (City)</th>
<th>Response Criterion (Highway)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Near $r$</td>
<td>−.12</td>
<td>−.13</td>
<td>−.002</td>
<td>.22</td>
<td>.16</td>
</tr>
<tr>
<td>$p$</td>
<td>.43</td>
<td>.37</td>
<td>.92</td>
<td>.13</td>
<td>.29</td>
</tr>
<tr>
<td>Same $r$</td>
<td>.17</td>
<td>.05</td>
<td>−.12</td>
<td>.16</td>
<td>.07</td>
</tr>
<tr>
<td>$p$</td>
<td>.25</td>
<td>.72</td>
<td>.46</td>
<td>.29</td>
<td>.64</td>
</tr>
<tr>
<td>Far $r$</td>
<td>−.15</td>
<td>−.22</td>
<td>−.19</td>
<td>.27</td>
<td>.15</td>
</tr>
<tr>
<td>$p$</td>
<td>.31</td>
<td>.12</td>
<td>.20</td>
<td>.06</td>
<td>.28</td>
</tr>
</tbody>
</table>

5.4 Discussion

5.4.1 Differences in City and Highway HP Scores

In both the response buffer and the response criterion, there were significant differences between the city and the highway environment. The differences in response criterion replicates Burge and Chaparro (2018) and can be explained by the differences in highway and city driving. These differences are both structural differences in the roadways themselves but also in the amount of information presented to the driver.

City driving and highway driving is very different in terms of the physical environment. Olson, Dewar, and Farber (2010) list features of roadways, such as roadway width and
intersections, which can cause significant differences in driving behavior. In highway driving
(as defined in the videos), all were in a divided highway, meaning that all cars were going the
same direction and at roughly the same speed. Further, videos in the highway environment
offered no intersections (only merging lanes). A disproportionate number of fatal crashes occur
at intersections (National Highway Transportation Safety Administration, 2018), and these types
of environments are very common in the video clips used for this study. Additionally, there is
generally a different speed traveled in these environments, and therefore there are different
appropriate and comfortable headway distances that are used (Taieb-Maimon & Shinar, 2001).
These differences may have created different environments that lead to the differences in
responses by participants.

In addition to specific aspects of the roadway design that may lead to different driving
behavior, the highway scenes were less visually complex than the city scenes, and therefore may
have had a much lower perceptual load. According to Lavie (2005), it may be more difficult for
individuals to selectively attend to targets and filter out distractors with high perceptual load.
Complexity has been shown to have an effect on driving performance (Chaparro & Alton, 2000)
and search efficiency (Ho, Scialfa, Caird, & Graw, 2001) during the driving task. These
important environmental differences could change the way hazards are perceived.

The results also suggest that there are differences in how drivers perceive hazards with
city and highway environments based at least partially on experience. It has been suggested that
a reason why experienced drivers perform better on HP tasks is because they have a more
sophisticated and accurate mental model of the roadway environment (Horswill & McKenna,
2004). Research by Underwood and colleagues (Chapman & Underwood, 1998; Underwood,
Chapman, Bowden, & Crundall, 2002) suggests that an impoverished mental model is why there
are differences in visual scanning patterns between experts and novices while they view HP video clips. Of particular importance to the present study is that scan patterns of novices did not change between the complexity of the roadway (urban, suburban, and rural), whereas the scan patterns do change for experts. Taken together with the literature mentioned above, this study suggests that there are differences in HP ability between city and highway environments, and that these differences interact with driving experience. The results shed important light on how drivers interact with roadway hazards in different types of environments.

5.4.2 Driving Experience Predicted City Driving Response Lag

A particularly important aspect of the city and highway differences is that driving experience only predicted HP performance in the city environment. This partially supports H4. Based on the differences discussed above, it may be that the aspects of city driving are what drivers learn as they become more experienced. More experienced drivers may be able to more quickly identify hazardous objects under high perceptual load conditions, where any difference between experts and novices was negligible in the sparse highway environment.

It is surprising that only response lag was predicted by driving experience. Using a video-based HP test, Wallis and Horswill (2007) found that experienced drivers did have a significantly more liberal response bias than inexperienced drivers. Though the current study used the method to categorize experienced drivers (over ten years of driving), Wallis and Horswill (2007) did have a higher average driving experience for that group ($M = 23.29, SD = 9.74$) compared to the present study ($M = 15.05, SD = 4.38$). Despite H4 only being partially supported, the differences between experts and novices in the response lag in city environments adds to the general validity of this test.
5.4.3 3D UFOV and HP Scores

The 3D UFOV replicated the effect of Experiment 1C, and found that when the peripheral targets are presented at a different depth than the central target, threshold scores are higher and therefore supported H1. This is important because the sample for Experiment 2 was larger and more diverse. It is surprising that the 3D UFOV did not predict any aspect of the HP score, and this does not support H2 or H3. Possible explanations are discussed below, and these address issues with the 3D UFOV test itself and its relation to the HP test.

It was initially hypothesized (H2) that the 2D UFOV, which is conceptually similar to the commercial versions, would predict more variance than the depth-related thresholds. Conceptually, the test measures the UFOV construct and there was an effect observed for the effect of retinal eccentricity, which is expected in a UFOV test (see experiments 1A and 1B). One possible reason for the lack of relationship is that the divided attention test is too easy and did not have enough variance. A meaningful way of comparing this to commercial UFOVS is through the normative data that has been collected where duration threshold was the dependent measure. Such normative data has been from 135 individuals aged 10 to 25 (Bennett et al., 2009) and 2759 individuals aged 56-94 (Edwards et al., 2006); these values can serve as a benchmark. Figure 37 shows the means and standard deviations for the pertinent tests as well as the 2D component of the 3D UFOV test used in this study. The 3D UFOV appears to be similar to the younger adults’ scores in terms of both central tendency and variance, but very different from older adults. This is important because a majority of the studies that have shown a relationship between the UFOV and crash risk have been with older adults. Not having comparable variance may have made this test less sensitive, and therefore substantially reduce the ability for it to predict HP scores.
The thresholds for conditions where the peripheral targets occur in depth do have substantially more variance. Therefore, even though the H2 was not supported, H3 could have been supported on its own. A reason for the lack of a relationship could be related to the method and magnitude of the depths that were measured in the UFOV. First, the magnitude of depth tested was not very large, only being displayed 9 cm in front and 15 cm behind the central target. This limit is because of the fusion limits of stereopsis. That is nowhere near the distances that are important in the driving scenario, with differences in the depths of roadway objects being in terms of meters instead of centimeters.

Similarly is the way that depth was simulated. In the 3D UFOV test, the only depth cue given was stereopsis. The HP was displayed on a traditional 2D screen (i.e., not stereoscopic display) and only contained monocular depth cues. Because depth was tested two different
ways, it is possible the effect measured through only binocular cues does not transfer to tasks with only monocular cues. Though stereopsis can be used to discriminate depths at great distances (Allison, Gillam, & Vecellio, 2009; Palmisano, Gillam, Govan, Allison, & Harris, 2010), it seems to provide the most useful depth information at very close distances (Cutting & Vishton, 1995; Surdick et al., 1994). Further, there is research suggesting that stereopsis does not play an important role in driving, as comparisons between binocular and monocular drivers has found no difference in various metrics of driving performance (McKnight et al., 1991; Troutbeck & Wood, 1994). For further discussion of the functional importance of stereopsis, see Fielder and Moseley (1996).

Another aspect that could explain the lack of relationship is what the UFOV was testing. The 3D UFOV measured divided attention, where participants were required to attend to central and peripheral targets simultaneously. The HP test may not really be assessing divided attention, and more selective attention as it relates to visual search. Although this is what the UFOV is in general attempting to measure (Ball & Owsley, 1993), the subtest of the UFOV used was specifically measuring divided attention. Experiment 1C suggested that the cost of dividing attention is what is contributing to the effect of depth, and if that attentional function is not occurring during the HP, it is unlikely there would be a relationship.

Finally, it is important to address what the HP test is actually measuring and how it is being measured. HP was chosen because it is a form of situation awareness (Horswill & McKenna, 2004) and is one of the few tasks that can predict crashes (Horswill et al., 2010; Horswill et al., 2015). This therefore may likely be a link between an attention function like UFOV and crash risk. However, it may be that much of the HP construct relies on level 3 SA, which involves predicting future events based on the current situation. An integral part of the
HP task is to accurately predict future events (i.e., will the current situation lead to a crash?). If that is what the HP test is actually tapping into, it is likely that the lower levels of SA do not play as large of a role. Previous research has suggested that the UFOV only relates to level 1 SA (Kaber et al., 2016), and therefore the effect might be small. If the effect were smaller than hypothesized, this study would lack the power to find it.

In addition to concerns that HP may not be an appropriate variable to test, there are concerns with how it was measured. HP was measured with a video-based measure, which is the most common method to assess the construct (Horswill, 2016b). This method essentially isolates HP from other aspects of driving. Underwood, Crundall, and Chapman (2011) have criticized these tests because they do not put realistic demands on the visual system that is required by driving. Many of these demands are related to vehicular control and navigation. For example, drivers tend to fixate on a tangent point when steering around curves (Land & Lee, 1994); the need to fixate on that is completely absent in the video-based tests. It is possible that the aspects of attention that are being captured by the UFOV would be more appropriately measured while performing a HP test concurrently with other driving tasks (e.g., car following).

In this study, HP was measured using FSDT. This method was chosen because the fuzzy nature of hazardous event makes it a more appropriate measure than fitting crisp, discrete categories of non-hazard and hazard onto a scenario. Following the method used by Burge and Chaparro (2018), participants viewed the video clips until they responded, at which time they rated the likelihood of a crash occurring. When looking at the participant’s responses, it appears that they were very confident in their responses, and perhaps treated less continuously than they should have. Of all of the clips, over 49% of the responses were in the top or bottom 5% of crash likelihood, compared to the 12.5% of the expert rating; over 31% rated it at either 100% or
0%, compared to the 2.5% of the expert rating. Figure 38 shows the histogram of the expert and participant ratings of the responses. Given the surprisingly homogeneity of the responses, it could be that the SDT measures were not sensitivity enough to find any relationship with the UFOV scores.

Figure 38: Distribution of responses for the experts and participants.

5.4.4 Future Directions

Though there were many unanticipated results in Experiment 2, the findings lay the foundation for subsequent research. Specifically, there are areas in the UFOV development, HP test development, and the relationship between the UFOV and crash risk.

As mentioned above, one potential problem with the 3D UFOV is the relative lack of variance compared with that found in older adults (the population in which the UFOV is often found to predict crash rates). Work in the future of the 3D UFOV, in addition to those discussed in Section 4.5 should relate work to increase the variability and the validity of the test. Because Gabor patches were used as the peripheral targets, many changes could be made to increase or decrease the difficulty, as many of the features of the Gabor patch (e.g., orientation, motion,
frequency) have been shown to guide attention (Wolfe & Horowitz, 2004). This should not only be done with the divided attention, but also with a 3D UFOV that measures selective attention and these could be correlated with already-established measures of the UFOV construct.

One of the goals for developing the 3D UFOV test was to increase the ecological validity of the test. This approach should be continued, and evaluation of other aspects of natural scenes that could be incorporated into it. One such feature could be motion, where participants must divide attention across multiple areas, where one or more of the targets are moving. This would be similar to the multiple-object tracking paradigm, and would likely increase the ecological validity of the test. In terms of applying it to the real word, most hazards are moving, and being able to track movements (and predict a trajectory of those hazards) may be very important to avoid hazards successful.

The results of this study also lay the groundwork for learning more about HP, as measured by the videos created by Burge and Chaparro (2018). The current results add to the validity the original article provided, but only in the city conditions. Future research should attempt to understand the difference in environments, and could be used to refine the HP test. Further, validity in using this HP test to predict actual driving metrics (e.g., crashes) would be incredibly important in linking this task with real-world performance.

In the current version of the HP test, participants end the video when they perceive a hazard to which they must respond. In the city environment, more experienced drivers responded quicker (i.e., had a greater response lag) then those who had less experience. In order to further understand expert-novice differences, it may be fruitful to incorporate aspects of the temporal occlusion paradigm into the HP test. This paradigm is often used in expertise research in sports psychology (Abernethy & Russell, 1987; Hagemann, Schorer, Cañal-Bruland, Lotz, &
Strauss, 2010), and in it participants view a dynamic scene in which the scene is occluded during specific points in the action sequence. After the occlusions, participants are asked to predict what will happen next. This is very similar to the video-based HP test. Instead of the end of the video being initiated by the participant, the videos could be stopped and occluded at various points to understand when or with what information experienced drivers can accurately recognize a hazard. Assessing the confidence that a crash would occur at different times may also lead to more variability with responses, which was discussed as an issue above.

The most surprising finding was the lack of relationship between the 3D UFOV and HP test. As mentioned above, part of this could be due to this HP test focusing on higher level SA than the UFOV is actually predicting. Future research should attempt to understand a possible relationship better, by focusing on a lower level of SA in driving.

Some examples of this could be specifically looking at peripheral hazards, and hazards that are quick-onset. Dichotomizing hazards like this has been done in the literature, where the onset or the availability of hazard information is present before the hazard actually occurs (Crundall et al., 2012; Vlakveld et al., 2011). It is likely that when information about the hazard is available before its onset, higher level of SA would be used. If the hazard has a sudden onset (i.e., without any warning or environmental cue) that may better reflect what is captured by the UFOV, particularly if the hazard originates in the periphery. Further, testing HP in a way that separates responses into levels of SA, such as the “What Happens Next” paradigm (Jackson, Chapman, & Crundall, 2009), may better relate the variance in HP to that of UFOV.

In addition to a standard, video-based version of the HP test, relating the UFOV to other driving measures may be fruitful. It has been proposed that video-based measures do not adequately capture visual demands placed on the driver (Underwood et al., 2011). In response to
this, simulator studies have been used that evaluate aspects of HP while the participant is driving (Crundall et al., 2012). Adding realistic demands of driving, such as following a lead vehicle, could more adequately represent the UFOV in the real world. However, considerable work creating realistic and confound-free simulator environments would be needed, as well as validation work relating those simulated tasks to actual driving tasks.
CHAPTER 6

CONCLUSIONS

The present dissertation was conducted in general to contribute to the understanding of how attention is used in everyday life. Specifically, the studies attempted to bridge the gap between a attention measure (UFOV) and an important, complex, and safety-critical task (driving). This was done in two different ways: to increase the ecological validity of the UFOV task and to understand how that task is related to the driving task. Understanding the relationship between applied and basic constructs is exceptionally important in the field of human factors. This project should lay the foundation for future studies to continue to address those goals.

The first aspect was to increase the ecological validity of the UFOV test. This was done by developing a version of the test that captures difference in attention due to the depth of the targets in the scene, not just the 2D representation of the display. The results of Experiment 1 suggest that there is an effect of depth, and that it can significantly impact how well areas and objects of the scene are processed. Future research should further understand the role of stereoscopic and non-stereopsis depth on the UFOV, as well as incorporating more real-world features.

The second aspect of the study was to apply this measure of attention to real-world tasks. Unlike laboratory-based measures where experimenters can control almost every component of a stimulus and isolate specific aspects, real-world tasks are complex and often involve the combination of cognitive processes to successfully complete them. Given previous research, it seemed likely that the UFOV, which measures spatial attention and has been related to crash risk, would be able to predict HP. The results of Experiment 2 did not support this overall
hypothesis. Because of this, future research should continue to address this relationship with the goal of understanding what aspects of driving the UFOV actually does relate to and predict. Additionally, it is important to understand what added value the 3D component of the UFOV test might be able to predict driving performance. Further, the specific video-based HP test was evaluated, and there are some adjustments that could make it more suitable to measure individual differences in how attention relates to real-world tasks.

As mentioned in the introduction, the discipline of psychological science, particularly perception and cognition, has often worked to apply findings and understand how various psychological principles are related to real-world tasks (Helton et al., 2015; Hoffman & Deffenbacher, 1992). Attention, a form of information processing, lies at the “heart of human performance” (Wickens & Carswell, 2012, p. 117), and understanding its limitations is an important part of understanding how it might impact our ability to complete tasks and an important part of the field of human factors psychology (Stone, Chaparro, Keebler, Chaparro, & McConnell, 2018).

Despite not supporting all the hypotheses, the current study lays groundwork for future researchers to further understand the role of attention in everyday life. The UFOV presents an interesting concept and challenge for researchers. As it seems to capture many aspects of attention (Woutersen et al., 2017) and has shown to be related to a critically-important real-world metric, crash risk (Goode et al., 1998), there is promise that with continuing our understanding of that task and how it is applied, we will continue to improve understanding of our visual functions in everyday tasks. This was the goal of the UFOV test when it was created two decades ago (Ball & Owsley, 1993), and that goal should motivate researchers to continue to understand the psychological processes behind the safety of human performance.
REFERENCES
REFERENCES


Burge, R. J. (2015). *Using fuzzy signal detection theory to assess the impact of text messaging on drivers’ hazard perception ability.* (Doctoral Dissertation), Wichita State University, Wichita, KS.


Pierce, R. S., & Andersen, G. J. (2014). The effects of age and workload on 3D spatial attention in dual-task driving. *Accident Analysis & Prevention, 67*, 96-104.


APPENDICES
A: Stereoscopic Discomfort Survey

This questionnaire was adapted from Hoffman et al. (2008) and was administered via an online survey at the end of the 3D Ufov test in experiments 1A and 1B. The text of the questions is reprinted below:

For each of the following symptoms, select a description that best represents the severity of that symptom at this moment.

How tired are your eyes?
- Very Fresh
- OK
- Mildly Tired
- Moderately Tired
- Very Tired

How clear is your vision?
- Very Clear
- OK
- Mild Blur
- Moderate Blur
- Much Blur

How tired and sore are your neck and back?
- Very Fresh
- OK
- Mild Ache
- Moderate Ache
- Severe Ache

How do your eyes feel?
- Very Fresh
- OK
- Mild Strain
- Moderate Strain
- Severe Strain

How does your head feel?
- Very Fresh
- OK
• Mild Ache
• Moderate Ache
• Severe Ache
**B: Post-Trial Hazard Ratings**

These questions were administered after each HP clip. It was adapted from (Burge, 2015; Burge & Chaparro, 2018) and was displayed via the computerized HP program.

Please give a percentage of confidence as to whether a response would have been required to avoid a collision in this scenario (0% = not confident at all; 100% = full confident).

What were you responding to?
C: HP Post-Session Questionnaire

This questionnaire was adapted from various sources (N. L. Morris et al., 2016; Owsley et al., 1999; Reimer et al., 2005) and was administered via an online survey at the end of the HP test in Experiment 2. The text of the questions is reprinted below:

Driving Experience
1. Do you have a valid driver’s license?
2. How many years have you had your driver’s license?
3. Approximately how many hours do you drive each week?
4. Approximately how many miles do you drive each week?
5. Where do you do a majority of your driving? (1 = never; 2 = hardly ever; 3 = sometimes; 4 = most days; 5 = every day)
   a. Rural road
   b. Rural interstate/highway/freeway
   c. Urban road/surface street
   d. Urban interstate/highway/freeway

Crashes and Citations
6. During the last three years, how many minor road accidents have you been involved in and were at fault? A minor accident is one in which no-one required medical treatment AND cost of damage to vehicles and property were less than $1,000.
7. During the last three years, how many major road accidents have you been involved in where you were at fault? A major accident is one in which either someone required medical treatment, or cost of damage to vehicles and property were greater than $1,000, or both.
8. During the last three years, have you been pulled over for (Yes or No):
   a. Speeding
   b. Careless or dangerous driving
   c. Driving under the influence of alcohol/drugs

Shortened American Driving Behavior Questionnaire (0 = never; 5 = nearly all the time)
9. Try to pass another car that is signaling a left turn
10. Select the wrong turn lane when approaching an intersection
11. Fail to “Stop” or “Yield” at a sign, almost hitting a car that has the right of way
12. Misread signs and miss your exit
13. Fail to notice pedestrians crossing when turning onto a side street
14. Drive very close to a car in front of you as a signal that they should go faster or get out of the way
15. Forget where you parked your car in a parking lot
16. When preparing to turn from a side road onto a main road, you pay too much attention to the traffic on the main road so that you nearly hit the car in front of you.

17. When you back up, you hit something that you did not observe before but was there.

18. Pass through an intersection even though you know that the traffic light has turned yellow and may go red.

19. When making a turn, you almost hit a cyclist or pedestrian who has come up on your right side.

20. Ignore speed limits late at night or very early in the morning.

21. Forget that your lights are on high beam until another driver flashes his headlights at you.

22. Fail to check your rear-view mirror before pulling out and changing lanes.

23. Have a strong dislike of a particular type of driver, and indicate your dislike by any means that you can.

24. Become impatient with a slow driver in the left lane and pass on the right.

25. Underestimate the speed of an oncoming vehicle when passing.

26. Switch on one thing, for example, the headlights, when you meant to switch on something else, for example, the windshield wipers.

27. Brake too quickly on a slippery road, or turn your steering wheel in the wrong direction while skidding.

28. You intend to drive to destination A, but you ‘wake up’ to find yourself on the road to destination B, perhaps because B is your more usual destination.

29. Drive even though you realize that your blood alcohol may be over the legal limit.

30. Get involved in spontaneous, or spur-of-the-moment, races with other drivers.

31. Realize that you cannot clearly remember the road you were just driving on.

32. You get angry at the behavior of another driver and you chase that driver so that you can give him/her a piece of your mind.
D: Fuzzy Signal Detection Calculations

Continuous values base on mapping functions (Parasuraman et al., 2000)

- **Hits:** $H = \min (s, r)$
- **Misses:** $M = \max (s - r, 0)$
- **False Alarms:** $FA = \max (r - s, 0)$
- **Correct Rejections:** $CR = \min (1 - s, 1 - r)$

Where $s$ is the continuous signal value and $r$ is the continuous response value

Equations for SDT rates (Parasuraman et al., 2000)

- **Hit Rate:** $\frac{\sum (\min (s_i, r_i))}{\sum (s_i)}$ for $i = 1$ to $N$
- **Miss Rate:** $\frac{\sum (\max (s_i - r_i, 0))}{\sum (s_i)}$ for $i = 1$ to $N$
- **False Alarm Rate:** $\frac{\sum (\max (r_i - s_i, 0))}{\sum (1 - s_i)}$ for $i = 1$ to $N$
- **Correct Rejection Rate:** $\frac{\sum (\min (1 - s_i, 1 - r_i))}{\sum (1 - s_i)}$ for $i = 1$ to $N$

Equations for Sensitivity, $A'$, and Response Criterion, $B''$ (See et al., 1997)

$$A' = \frac{1}{2} + \frac{(H - FA)(1 + H - FA)}{4H(1 - FA)}$$

$$B'' = \frac{H(l - H) - FA(l - FA)}{H(l - H) + FA(l - FA)}; \text{when } H \geq FA$$

$$B'' = \frac{FA(l - FA) - H(l - H)}{FA(l - FA) + H(l - H)}; \text{when } H \leq FA$$

Where $H$ is hit rate and $FA$ is false alarm rate