

VISUO-SPATIAL ABILITIES IN REMOTE PERCEPTION:  
A META-ANALYSIS OF EMPIRICAL WORK

by

Thomas D. Fincannon  
B.A. University of South Florida, 2003  
M.S. University of Central Florida, 2010

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the Department of Psychology  
in the College of Sciences  
at the University of Central Florida  
Orlando, Florida

Spring Term  
2013

Major Professor: Florian G. Jentsch

© 2013 by Thomas D. Fincannon

## ABSTRACT

Meta-analysis was used to investigate the relationship between visuo-spatial ability and performance in remote environments. In order to be included, each study needed to examine the relationship between the use of an ego-centric perspective and various dimensions of performance (i.e., identification, localization, navigation, and mission completion time). The moderator analysis investigated relationships involving: (a) visuo-spatial construct with an emphasis on Carroll's (1993) *visualization* (VZ) factor; (b) performance outcome (i.e., identification, localization, navigation, and mission completion time); (c) autonomy to support mission performance; (d) task type (i.e., navigation vs. reconnaissance); and (e) experimental testbed (i.e., physical vs. virtual environments). The process of searching and screening for published and unpublished analyses identified 81 works of interest that were found to represent 50 unique datasets. 518 effects were extracted from these datasets for analyses.

Analyses of aggregated effects (Hunter & Schmidt, 2004) found that visuo-spatial abilities were significantly associated with each construct, such that effect sizes ranged from weak ( $r = .235$ ) to moderately strong ( $r = .371$ ). For meta-regression (Borenstein, Hedges, Figgins, & Rothstein, 2009; Kalaian & Raudenbush, 1996; Tabachnick & Fidell, 2007), moderation by visuo-spatial construct (i.e., focusing on visualization) was consistently supported for all outcomes. For at least one of the outcomes, support was found for moderation by test, the reliability coefficient of a test, autonomy (i.e. to support identification, localization, and navigation), testbed (i.e., physical vs. virtual environment), intended domain of application, and gender.

These findings illustrate that majority of what researchers refer to as "spatial ability" actually uses measures that load onto Carroll's (1993) *visualization* (VZ) factor. The associations between this predictor and all performance outcomes were significant, but the significant

variation across moderators highlight important issues for the design of unmanned systems and the external validity of findings across domains. For example, *higher* levels of autonomy for supporting navigation *decreased* the association between *visualization (VZ)* and performance. In contrast, *higher* levels of autonomy for supporting identification and localization *increased* the association between *visualization (VZ)* and performance. Furthermore, moderation by testbed, intended domain of application, and gender challenged the degree to which findings can be expected to generalize across domains and sets of participants.

## ACKNOWLEDGMENTS

I would like to begin by thanking my chair and advisor, Dr. Florian Jentsch. While working on this project and under you at the Team Performance Laboratory (TPL), I have learned to develop the critical skills that are necessary to design, analyze, and disseminate research. Your support has been integral to this process. I would also like to thank Dr. Clint Bowers, Dr. Valerie Sims, and Dr. Jessie Chen, who served as members of my dissertation committee. Your feedback regarding methods, perspectives of cognitive ability, and applied operation of unmanned systems strengthened this work.

As meta-analysis is built off of the work of other researchers, their work should be acknowledged. In addition to the work that is highlighted in the reference section, I would like to thank researchers that responded to the call for work to be used in this dissertation.

Alphabetically, these people include Ellen Bass, John Blich, Daniel Cassenti, Jessie Chen, Shih-Yi (James) Chien, Lisa Fern, Joshua Gomer, Michael Goodrich, Leo Gugerty, Robert Hutton, Frank Lacson, Patricia McDermott, Betty Murphy, Robin Murphy, Tal Oron-Gilad, Lauren Reinerman-Jones, Harvey Smallman, Grant Taylor, Christopher Wickens, and Eric Wiebe. Your support and comments were greatly appreciated.

Next, I would like to thank my colleagues at TPL. To Sherri Rehfeld, Reagan Hoeft, Bill Evans, and Michelle Harper-Sciarini, you were my seniors who paved the way for me to work in this area. To Mike Curtis, Joe Keebler, Scott Ososky, Dave Schuster, Elizabeth Phillips, Javier Rivera, Brittany Sellers, and Andrew Talone, I have grown with you over the years, and your feedback has helped me to become a stronger researcher. Furthermore, each of you worked on research that was included in this meta-analysis, and as it is partially a reflection of research, I would like to thank you for your contribution.

Finally, I would like to thank the philanthropists and members of the theater and arts community in Orlando. When I was aggravated, you provided a necessary diversion. By working with you to support various causes and functions within the community, you helped me to better understand myself and become a more balanced person. I am truly grateful to have known each and every one you.

## TABLE OF CONTENTS

LIST OF FIGURES .....	ix
LIST OF TABLES.....	x
LIST OF ABBREVIATIONS .....	xiii
CHAPTER 1: INTRODUCTION .....	1
Statement of the Problem.....	1
Purpose.....	4
CHAPTER 2: LITERATURE REVIEW .....	6
Moderation by Ability.....	6
Moderation by Outcome .....	15
Moderation by Other Constructs.....	20
Statement of Hypotheses.....	24
CHAPTER 3: METHOD .....	28
Search & Inclusion Criteria .....	28
Coding of Studies.....	29
Analysis .....	33
CHAPTER 4: RESULTS.....	36
Data Screening.....	36
Main Effects.....	37
Moderator Analysis.....	42

CHAPTER 5: DISCUSSION.....	50
Theoretical Implications .....	<b>Error! Bookmark not defined.</b>
Practical Implications.....	61
Limitations .....	62
Future Research .....	65
Concluding Remarks.....	68
APPENDIX A: LIST OF SEARCH TERMS .....	70
APPENDIX B: LETTER TO REQUEST ACCESS TO UNPUBLISHED DATA.....	72
APPENDIX C: SAMPLE LETTER TO REQUEST CLARIFICATION OF PUBLISHED DATA .....	74
APPENDIX D: DESCRIPTIONS OF CONSTRUCTS FOR SUBJECTIVE RATINGS .....	76
APPENDIX E: LIST OF FORMULAS .....	81
APPENDIX F: TABLES FOR STEM & LEAF PLOTS, NORMALITY ANALYSIS, AND POWER ANALYSES.....	92
APPENDIX G: MATRIX OF CORRELATIONS FOR MODERATOR VARIABLES .....	97
APPENDIX F: FLOWCHART FOR OBTAINING AND SCREENING STUDIES FOR INCLUSION.....	99
REFERENCES .....	101



## LIST OF FIGURES

Figure 1. Visualization (VZ) impacting non-visual performance through a visual performance mediator .....	18
Figure 2. The impact of visualization (VZ) on non-visual performance being negated by two visual mediators .....	19
Figure 3. Summary of hypotheses .....	24

## LIST OF TABLES

Table 1. Common categories of tests that load onto Carroll's (1993) visualization (VZ) factor....	9
Table 2. Tests that load onto Carroll's (1993) spatial relations (SR) factor .....	10
Table 3. Categories of tests that load onto Carroll's (1993) closure speed (CS) factor .....	12
Table 4. Categories of tests that load onto Carroll's (1993) closure flexibility (CF) factor .....	13
Table 5. Categories of tests that load onto Carroll's (1993) perceptual speed (P) factor .....	14
Table 6. Categories of tests that load onto Carroll's (1993) visual memory (MV) factor .....	14
Table 7. Description of visual performance outcomes for coding .....	31
Table 8. Levels of autonomy to support identification (Kappa = .84) .....	32
Table 9. Levels of autonomy to support localization (Kappa = .79).....	32
Table 10. Levels of autonomy to support navigation (Kappa = .92) .....	32
Table 11. Description of visual performance outcomes for coding (k = number of studies).....	36
Table 12. Main effects of all visuo-spatial measures on performance.....	41
Table 13. Model for testing the moderating effects of construct on mission completion time (k = 18 studies).....	42
Table 14. Model for testing the moderating effects of construct on identification (k = 29 studies) .....	43
Table 15. Model for testing the moderating effects of construct on localization (k = 15 studies)	43
Table 16. Model for testing the moderating effects of construct on navigation (k = 18 studies) .	44
Table 17. Main effects of visualization (VZ) on measures on performance.....	45
Table 18. Model for variables moderating the effect of visualization (VZ) on identification (k = 25 studies).....	47
Table 19. Model for variables moderating the effect of visualization (VZ) on mission completion time (k = 18 studies) .....	47

Table 20. Model for variables moderating the effect of visualization (VZ) on localization (k = 15 studies).....	48
Table 21. Model for variables moderating the effect of visualization (VZ) on navigation (k = 18 studies).....	48
Table 22. Summary of results per hypothesis .....	51
Table 23. Stem & leaf plot for the uncorrected association between identification and visuo-spatial ability.....	93
Table 24. Stem & leaf plot for the uncorrected association between localization and visuo-spatial ability .....	93
Table 25. Stem & leaf plot for the uncorrected association between navigation and visuo-spatial ability .....	93
Table 26. Stem & leaf plot for the uncorrected association between mission time and visuo-spatial ability.....	94
Table 27. Stem & leaf plot for the uncorrected association between identification and visualization (VZ).....	94
Table 28. Stem & leaf plot for the uncorrected association between localization and visualization (VZ) .....	94
Table 29. Stem & leaf plot for the uncorrected association between navigation and visualization (VZ) .....	95
Table 30. Stem & leaf plot for the uncorrected association between mission time and visualization (VZ).....	95
Table 31. Normality analyses for Tables 12, 24, 25, 26, and 27.....	95
Table 32. Normality analyses for Tables 17, 28, 29, 30, and 31.....	96
Table 33. Power analyses for Table 12.....	96

Table 34. Power analyses for Table 17 .....	96
Table 35. Matrix for correlation between moderators .....	98

## LIST OF ABBREVIATIONS

---

CF	Closure Flexibility
CS	Closure Speed
HR	Human-Robot
HRI	Human-Robot Interaction
MV	Visual Memory
P	Perceptual Speed
RSTA	Reconnaissance, Surveillance, and Target Acquisition
SR	Spatial Relations
UV	Unmanned Vehicle
VZ	Spatial Visualization

---

# CHAPTER 1: INTRODUCTION

## Statement of the Problem

Human operation of unmanned vehicles (UVs) represents a segment of human-robot interaction (HRI) research that has grown considerably over recent years. Within this domain, one of the constructs that have been studied involves spatial ability and its relation to performance in the HRI domain (Billings & Durlach, 2008, 2009; Chen, 2010; Chen, Durlach, Sloan, & Bowens, 2008; Chen, & Terrence, 2008, 2009; Fincannon, Evans, Jentsch, & Keebler, 2010a; Fincannon, Ososky, Keebler, & Jentsch, 2010b; Lathan & Tracey, 2002). While many of these studies have demonstrated an association between spatial ability and performance (Chen, 2010; Chen et al., 2008; Chen, & Terrence, 2008, 2009; Fincannon et al., 2010a, 2010b; Lathan & Tracey, 2002), these findings are not consistent across all research (Billings & Durlach, 2008, 2009). This inconsistency demonstrates a need to review the literature, and the goal of this work was to conduct a quantitative meta-analysis that examined these associations with spatial ability in UV operation.

There are many potential benefits to using UVs in applied environments (e.g., UVs remove operators from hostile/hazardous environments; UVs enable operators to search areas that are physically inaccessible to humans), but research has found that the use of this technology can be more difficult than expected. For example, a long-term goal of UV research involves minimizing the number of humans and maximizing the number of robots that are necessary to complete a task, but empirical research does not always support this assumption (Burke & Murphy, 2004; Rehfeld, Jentsch, Curtis, & Fincannon, 2005). Specifically, researchers have found that: (a) individual operators experience difficulties with controlling more than one UV (Chadwick, 2005, 2006), and (b) increasing the number of human operators will generally improve performance (Burke & Murphy, 2004; Fincannon, Evans, Phillips, Jentsch, & Keebler,

2009; Rehfeld et al., 2005). As difficulties limit the degree to which expected benefits are realized, researchers must develop a better understanding of underlying issues and relationships.

Many of the problems with UV operation are generally attributed to the difficulties of “remote perception” (Burke & Murphy, 2004; Burke, Murphy, Covert, & Riddle, 2004; Casper & Murphy, 2003; Fincannon, Keebler, Jentsch, & Curtis, in press-b; Woods, Tittle, Feil, & Roesler, 2004). As “remote perception” is a vague term, researchers need to specify relevant constructs for these perceptual outcomes and discuss the differences accordingly. For example, team research has used empirical methods to highlight how identification, localization, and navigation tasks are distributed across different operator roles (Fincannon, Keebler, Jentsch, Phillips, & Evans, 2011; McDermott, Luck, Allender, & Fisher, 2005) for reconnaissance, surveillance, and target acquisition (RSTA) tasks. In contrast, other tasks, such as manipulation of a robotic arm in laparoscopic surgery (Keehner et al., 2004), incorporate additional, non-visual elements (e.g., fine motor control) into performance outcomes. As each task requires different cognitive processes, researchers must understand which measure of “remote perception” they are interested in studying (e.g., identification, localization, navigation, mission completion time), and defining the construct deserves attention.

Another key to understanding perception from a UV lies in developing an understanding of the components that make this task difficult. A discussion of this issue by Woods and colleagues (2004) asserted that problems with perception can be attributed to issues that involve the physical separation of people from the environment that they intend to study. One of these factors involves the *limited field of view* from remote cameras. For example, UVs that were used for search and rescue at the World Trade Center provided a 53° field of view (Casper & Murphy, 2003), and this lack of peripheral information can hinder an operator’s ability to direct attention to meaningful information (Woods et al., 2004). A second issue was termed *scale ambiguity*,

which involved determining the scale of a robot in relation to its environment. Specifically, operators have a limited representation of the environment (e.g., limited field of view) that hinders their ability to understand how a UV relates to objects in the environment, and this can create problems moving between objects and avoiding obstacles (Casper & Murphy, 2003). A third issue involves the requirement for UV operators to resolve *conflicts between perceptual cues*. For example, Woods and colleagues (2004) noted that vestibular feedback typically accompanies physical movement through an environment, but since operators are not moving in conjuncture with their UV(s), there can be a conflict between visual and vestibular information that must be resolved. As a result of these issues, UV operation presents a context with a variety of obstacles.

One strategy for improving perceptual performance involves visuo-spatial abilities. Research by my colleagues and me (Fincannon et al., 2010a) has noted that published work has examined multiple abilities across a variety of studies (Billings & Durlach, 2008, 2009; Chen, 2010; Chen, & Terrence, 2008, 2009; Fincannon et al., 2010b; Lathan & Tracey, 2002). The problem with this was best described by Carroll (1993), who noted that there are “contradictory and confusing conclusions about exactly what abilities exist and how they should be defined and measured” (p. 304). This ambiguity has resulted in researchers using a general “spatial ability” term to describe a variety of measures that may ultimately assess different visuo-spatial constructs. Since poor usage of the “spatial ability” term can ultimately lead to misunderstandings in research with unmanned systems, there was a need for a quantitative meta-analysis to classify visuo-spatial constructs that have been studied and examine how they impact operational effectiveness.

An empirical meta-analysis in HRI must account for the diversity of systems that appear in UV research. For example, researchers have studied operator performance in systems ranging



from tele-operation (Burke & Murphy, 2004; Burke, Murphy, Coover, & Riddle, 2004; Casper & Murphy, 2003) to semi-autonomous control (Rehfeld, Jentsch, Curtis, & Fincannon, 2005), and recent iterations have extended this by focusing on supervisory control (Chen, Barnes, and Harper-Sciarini, 2010) and human-robot teaming (Phillips, Ososky, & Grove, 2011; Schuster, Ososky, Jentsch, Phillips, & Evans, 2011; Wang & Lewis, 2007). This diversity also extends to performance outcomes, which have been defined and measured in a variety of different ways (Fincannon, Jentsch, Sellers, & Keebler, 2012; Fincannon, Ososky, Jentsch, Phillips, & Keebler, 2010c; McDermott et al., 2005; Sellers, Fincannon, & Jentsch, 2011; Steinfeld et al., 2006). Since this variation can lead to inconsistent findings across published work, hypotheses need to consider how these findings differ (Fincannon, Keebler, & Jentsch, in press-a), and a quantitative meta-analysis can be used to conduct this analysis.

In summary, there are a number of potential benefits to using UVs in applied contexts, but difficulties with UV operation often minimize the degree to which operators can utilize these strengths. While visuo-spatial abilities can help operators overcome these difficulties, differences across visuo-spatial abilities, outcomes, and context can lead to confusion surrounding the associations between constructs, and a meta-analytic review was required to assess various aspects of moderation in UV operations. The goal of this research was to conduct a quantitative meta-analysis for examining these issues.

### Purpose

The purpose of this dissertation was to conduct a study that examines the following questions:

1. In the domain of UV operation, what visuo-spatial constructs are discussed under the heading of “spatial ability”?
2. Which visuo-spatial constructs are associated with operational effectiveness?

3. What outcomes are associated with a specific visuo-spatial construct?
4. When an association exists between a visuo-spatial construct and perceptual outcome, what factors moderate this relationship?

The goal of the analysis in this dissertation was to identify contexts in which visuo-spatial abilities become important and measures that are most likely to report unique information.

There are four sections to this dissertation. The first section discusses relevant literature. This section is sub-divided into sections that address visuo-spatial constructs, performance outcomes, variations across studies, and a summary of hypotheses. The second section discusses experimental methodologies that are associated with obtaining, coding, and analyzing the body of published and unpublished work. The third section tests hypotheses and presents supplemental analysis. The fourth section discusses findings within the context of the larger body of literature. Relevant appendices are included at the end of the document.

## CHAPTER 2: LITERATURE REVIEW

The goal of this work was to conduct a quantitative meta-analysis of UV operation, and the current section presents a review of relevant literature. It begins with a discussion of visuo-spatial constructs from Carroll's (1993) model of intelligence that also appear in current HRI literature. This next section discusses different outcomes that have been empirically studied in UV operation. The third section builds on this foundation by discussing constructs that moderate the relationship between spatial ability and performance. The final section summarizes hypotheses that were identified through the course of the literature review.

### Moderation by Ability

Spatial ability has been extensively studied in UV literature (Billings & Durlach, 2008, 2009; Chen, 2010; Chen et al., 2008; Chen, & Terrence, 2008, 2009; Fincannon et al., 2010a, 2010b; Lathan & Tracey, 2002), but research indicates that the construct is quite complex. Carroll (1993) noted that confusion surrounding the nature of the construct has caused researchers to rely on a general "spatial ability" term, which can create several problems. First, researchers can use the same term to describe two or more constructs, and this would result in findings across studies that possibly contradict one another. Second, researchers can use different terminology to describe the same construct, which can lead researchers to expect differences that do not exist. Consequently, visuo-spatial constructs should be sufficiently reviewed before conducting a meta-analysis.

An examination of literature in the domain of visuo-spatial ability reveals a long history of research (Bowers, Milham, & Price, 1998; Carroll, 1993; Ekstrom, French, Harman, Dermen, 1976; Guilford, 1985; Guilford & Hoepfner, 1971; Horn & Cattell, 1966; Lohman, 1996; Lohman, Pellegrino, Alderton, & Regian, 1987; McGee, 1979; McGrew, 2005, 2009; Michael, Zimmerman, & Guilford, 1951; Thurstone, 1938; Vogel, Bowers, Vogel, 2003) that contains

many criticisms (Cronbach & Snow, 1977; Guilford, 1980; Haris, 1967; Horn & Knapp, 1973, 1974; Horn & Noll, 1997; Lohman, 1989), rebuttals (Carroll, 2003; Guilford, 1974; Horn & Cattell, 1982), and calls for continued research (Hegarty & Waller, 2005; McGrew, 2009) regarding the classification of these terms. In order to address issues surrounding the complexity of this confusing debate, Carroll (1993) reanalyzed over 460 datasets. The result of this meta-analysis included the creation of a three stratum model of intelligence that has been acknowledged as one of the more significant contributions to modern theory (Cucina, Gast, & Su, 2012; Hegarty & Waller, 2005; Horn, 1998; McGrew, 2005, 2009). As a result of this foundation, Carroll's model served as the primary foundation for discussing visuo-spatial constructs.

Within Carroll's (1993) model, an important distinction exists between *level* (i.e., level of mastery) and *speeded* factors. Specifically, *level* factors are marked by their complexity and difficulty (i.e., the factor is defined by an individual's ability to complete a range of easy to difficult items under untimed conditions), whereas *speeded* factors are defined by a rate of perceiving relatively simple items (i.e., the factor is marked by an individual's speed in responding to items that are easy enough to be completed by any respondent). If this distinction holds true with UV operation, one might expect speeded factors to be associated with UV performance that relies on perception of stimuli within a short period of time, whereas level factors might be associated with perception of stimuli that are difficult to observe. In context of visuo-spatial factors, this issue is probably most important for distinguishing *visualization* (VZ) from *spatial relations* (SR), which both require a spatial transformation (i.e., as level and speeded factors respectively). Both of these constructs are addressed in more detail within the following subsections.

The following subsections provide a construct-driven review of the major visuo-spatial factors that were identified in Carroll's (1993) research. The majority of the factors were found to load onto the *broad visual perception (2V)* factor, which includes spatial visualization, spatial relations, closure speed, closure flexibility, and perceptual speed. Due to its appearance across several studies (Keebler, 2011; Lathan & Tracey, 2002, Lohman et al., 1987), this review also extends to visual memory, which loaded onto Carroll's factor for *memory and learning (2Y)*. Each of the following factors will be discussed in more detail below.

#### *Spatial Visualization (VZ)*

*Visualization (VZ)* represents a commonly recognized factor in the visuo-spatial domain, and Carroll (1993) defined it as the "ability in manipulating visual patterns, as indicated by level of difficulty and complexity in visual stimulus material that can be handled successfully" (p. 362). Given that the emphasis on the difficulty and complexity of a mental transformation, it is not surprising to find that researchers have discussed the impact of this construct on operator performance (Chen, 2010; Chen et al., 2008; Chen, & Terrence, 2008, 2009; Fincannon et al., 2008, 2010a, 2010b, 2010c; Schuster et al., 2008; Sellers et al., 2011). As a result, *visualization (VZ)* must be closely examined within the domain of UV operation.

One of the more important details that researchers have noted from Carroll's (1993) model involves the diversity of measures (see Table 1) that represent this factor (Fincannon et al., 2010a; Lohman, 1996). Specifically, Carroll's meta-analysis of visuo-spatial abilities worked from a framework that was established by Lohman and colleagues (1987), and although the discussion of a *visualization (VZ)* factor was very similar across both works, Lohman and colleagues argued for a spatial orientation factor that was not supported (i.e., perspective taking markers of spatial orientation were found to load onto the *VZ* construct) in Carroll's model. As a result of this, researchers (Fincannon et al., 2010a; Lohman, 1996) have recommended that a

variety of measures be used to assess the *visualization (VZ)* construct. This recommendation implied a hypothesis where the effects of *visualization (VZ)* on operational effectiveness would be moderated by the test that is used to assess this relationship.

Table 1. Common categories of tests that load onto Carroll's (1993) *visualization (VZ)* factor

<b>Category</b>	<b>Task Description</b>	<b>Sample Tests</b>
Paper Formboard & Assembly Tasks	Various components of a figure are mentally combined to complete a whole image	<ul style="list-style-type: none"> <li>• Form Board Test</li> <li>• Fitting Shapes Test</li> </ul>
Block Rotation Tasks	One or more rotated images are compared to a reference image to determine whether they are same or different	<ul style="list-style-type: none"> <li>• Guilford Zimmerman Test of Spatial Visualization</li> <li>• Cube Comparison Test</li> <li>• Vandenberg &amp; Kuse Test of Mental Rotation</li> </ul>
Paper Folding Tasks	After viewing an image that illustrates how a piece of paper has been folded, holes are punched into the folded piece of paper, and participants must determine how the arrangement of holes will appear in the unfolded piece of paper	<ul style="list-style-type: none"> <li>• Paper Folding Test</li> <li>• Punch Holes Test</li> </ul>
Surface Development Tasks	Subjects must determine how a pattern should be rolled up into, or broken down from, a three-dimensional figure	<ul style="list-style-type: none"> <li>• Surface Development Test</li> <li>• Pattern Development</li> </ul>
Perspective Tasks	Objects or reference points from a picture are mentally aligned to make judgments about differing viewpoints	<ul style="list-style-type: none"> <li>• Guilford-Zimmerman Test of Spatial Orientation</li> <li>• Eliot-Price Perspectives Test</li> </ul>

### *Spatial Relations (SR)*

Carroll (1993) defined *spatial relations (SR)* as the “speed in manipulating relatively simple visual patterns” (p. 363), and the ability is typically marked by simple measures of mental rotation (see Table 2). While markers of this construct (e.g., card rotation) have traditionally been associated with spatial orientation (Ekstrom et al., 1976; Horn & Cattell, 1966; McGee, 1979; Michael, Guilford, Frutchter, & Zimmerman, 1957), relatively modern meta-analyses have drawn a distinction between complex rotation of the *visualization (VZ)* factor (i.e., the most difficult transformation that a person can perform) and speeded rotation of the *spatial relations (SR)* factor (i.e., the time to perform simple transformations that can be completed by everyone). As measures of this construct appears in studies of UV operation (Keebler, 2011; Rehfeld, 2006), there was a need to consider its impact on operational effectiveness.

Table 2. Tests that load onto Carroll’s (1993) *spatial relations (SR)* factor

<b>Task Description</b>	<b>Sample Tests</b>
Simple manipulation of a visual form by rotation and/or reflection	<ul style="list-style-type: none"><li>• Card Rotation Test</li><li>• Flags Test</li></ul>

Given that *spatial relations (SR)* marks a separate construct, one could expect differences in how this construct predicts performance with UV systems. Specifically, the *spatial relations (SR)* factor is marked by simple transformations (Carroll, 1993; Lohman et al., 1987). This is in contrast to perception from a UV, which has been described as a difficult task (Casper & Murphy, 2003; Woods et al., 2004). For example, Casper and Murphy (2003) noted that victims in search and rescue can be obscured by rubble and soot, making target identification difficult. A recent study (Fincannon et al., in press-b) not only demonstrated how the introduction of these elements hinders identification, but it also showed how obscurity alters the impact of cognitive ability on

performance. If the *spatial relations (SR)* factor is a measure of simple transformation, it should be less predictive of complex visual performance than *visualization (VZ)* in UV operation.

This does not mean that the *spatial relations (SR)* factor is irrelevant to UV operation. Since *spatial relations (SR)* represents a speeded factor (i.e., high-ability individuals are faster at completing mental transformations), one might expect this construct to be associated with dimensions of operator performance that require fast reaction times. Recent research by my colleagues and me (Fincannon et al., 2012) found that some measures of operational effectiveness (e.g., stopping a UV before it runs into a person or object) formed stronger associations with speeded (i.e. closure speed) than level (i.e., spatial visualization) factors of visual perception. Since the measure of performance in question relied on an operator's ability to respond to a visual stimulus under speeded conditions, this finding suggests that speeded factors of visual perception may be uniquely associated with UV tasks that rely on speeded perception, and one could expect *spatial relations (SR)* to be important for these dimensions of performance.

#### *Closure Speed (CS)*

*Closure speed (CS)*, which has also been termed “gestalt perception,” is defined as “apprehending and identifying a visual pattern, without knowing in advance what the pattern is, when the pattern is disguised or obscured in some way” (Carroll, 1993, p. 363). Given that markers of the *closure speed (CS)* construct (see Table 3) emphasize the recognition of a stimulus that is obscured (e.g., camouflage, partial obstruction, etc.) and stored in long-term memory, there is a somewhat obvious rationale for hypothesizing an association between this construct and perception. Furthermore, the emphasis on the partial obstruction of stimuli also implies the use of unique cognitive processes, such as edge completion from Biederman's (1987) *Theory for Recognition by Components*. Given that *closure speed (CS)* has also appeared in relatively early



discussions of UV navigation (Lathan & Tracey, 2002), the *closure speed (CS)* factor needs to be considered as a unique construct.

In the context of Carroll’s (1993) work, there are two conceptualizations of *CS* that need to be considered. First, the *closure speed (CS)* construct was formally classified as a speeded factor of intelligence (Carroll, 1993), and as with *spatial relations (SR)*, *closure speed (CS)* would be expected to be associated with speeded dimensions of visual performance. The second conceptualization of *closure speed (CS)* involved some apprehension by Carroll regarding the exclusivity of *closure speed (CS)* as a factor of cognitive speed. Specifically, Carroll (1993) stated that the factor could be “classified as a level factor in the sense that it is, in effect, a measure of the degree of obscuration or degradation that an individual can tolerate” (p. 465). This implies that *closure speed (CS)* is similar to *visualization (VZ)* in that both may be predictive of complex visual perception. As research has reported findings with speeded (Fincannon et al., 2012) and complex (Phillips et al., 2010b) dimensions of visual performance, *closure speed (CS)* was considered as a unique construct.

Table 3. Categories of tests that load onto Carroll’s (1993) *closure speed (CS)* factor

Category	Task Description	Sample Tests
Representation of Images	A familiar representation of an obscured object must be named	<ul style="list-style-type: none"> <li>• Gestalt Completion Test</li> <li>• Snowy Pictures Test</li> </ul>
Representation of Words	A familiar representation of an obscured word must be named	<ul style="list-style-type: none"> <li>• Concealed Words Test</li> <li>• Disguised Words Test</li> </ul>

#### *Closure Flexibility (CF)*

According to Carroll (1993), *closure flexibility (CF)* can be defined as the “speed of detecting and disembedding a known stimulus array from a more complex array” (p. 341) or the “speed in finding, apprehending, and identifying a visual pattern, knowing in advance what is to

be apprehended, when the pattern is disguised or obscured in some way” (p. 363). While *closure flexibility (CF)* does appear to be similar to the *closure speed (CS)* factor (i.e., both constructs require the recognition of obscured stimuli), markers of the *closure flexibility (CF)* factor use stimuli that are more abstract and unfamiliar, which removes the long-term memory component from *closure speed (CS)*. Markers of the *closure flexibility (CF)* construct (see Table 4) appear in several studies of UV operation (Billings & Durlach, 2008, 2009; Chen & Terrance, 2008, 2009), and as Carroll (1993) described it as a speeded factor, *closure flexibility (CF)* would be expected to be uniquely related to speeded measures of visual performance.

Table 4. Categories of tests that load onto Carroll’s (1993) *closure flexibility (CF)* factor

Task Description	Sample Tests
A simple geometric shape is presented in an unobstructed format to a person, who must then detect this target in a more complex array of geometric noise/camouflage	<ul style="list-style-type: none"> <li>• Hidden Figures</li> <li>• Hidden Patterns</li> <li>• Group Embedded Figures</li> </ul>

#### *Perceptual Speed (P)*

*Perceptual speed (P)* has been commonly recognized as a factor, which can be defined as the “speed in finding a known visual pattern, or in accurately comparing one or more patterns in a visual field, such that the patterns are not disguised or obscured” (Carroll, 1993, p. 363).

Research has presented this factor as a construct that is related to working memory (Carroll, 1993; McGrew, 2005). Recent research (Fincannon, Jentsch, Sellers, & Keebler, 2012; Sellers, Fincannon, & Jentsch, in press) has considered *perceptual speed (P)* in HRI, and findings appear to indicate that there may be an association between this construct and workload. Therefore, *perceptual speed (P)* was considered as a factor of interest in UV operation.

Table 5. Categories of tests that load onto Carroll’s (1993) *perceptual speed (P)* factor

<b>Category</b>	<b>Task Description</b>	<b>Sample Tests</b>
Locating Stimuli	One or more symbols/patterns must be located in a visual field	<ul style="list-style-type: none"> <li>• Finding A’s Test</li> </ul>
Comparing Stimuli	Two stimuli must be compared to determine whether or not they are identical to each other	<ul style="list-style-type: none"> <li>• Identical Pictures Test</li> </ul>

*Visual Memory (MV)*

Carroll (1993) defined *visual memory (MV)* as the ability to “form, in a study phase, a mental representation (or possibly an image) of visual material that is presented . . . and to use that representation in responding in a test phase by recognition or recall” (p. 302). While this construct appeared in a discussion of visuo-spatial abilities by Lohman and colleagues (1987), Carroll’s (1993) analysis, which considered a more diverse range of constructs, found that the *MV* factor was more strongly correlated with measures of memory and learning. Under this framework, it might be possible to expect *visual memory (MV)* to be associated with learning aspects of a remote environment that are typically novel to operators (e.g., search and rescue in a rubble pile or collapsed building). As this construct appears in research addressing navigation (Lathan & Tracey, 2002) and target identification (Keebler, Sciarini, Jentsch, Nicholson, & Fincannon, 2010), there was a need to examine *visual memory (MV)* in UV operation.

Table 6. Categories of tests that load onto Carroll’s (1993) *visual memory (MV)* factor

<b>Category</b>	<b>Task Description</b>	<b>Sample Tests</b>
Reproduction	Immediately after exposure to a flash card, an examinee must reproduce (draw) that design	<ul style="list-style-type: none"> <li>• Reproduction of Visual Designs Test</li> </ul>
Recognition	Following a study period, an examinee must select items that correctly match the stimulus	<ul style="list-style-type: none"> <li>• Map Memory Test</li> </ul>

### *Summary*

In summary, confusion in the domain of visuo-spatial constructs has resulted in a variety of tests falling under the “spatial ability” heading (Carroll, 1993), and this section has briefly reviewed how these issues permeate research on UV operation. In the context of Carroll’s (1993) model, there are arguably six major abilities that appear in the published body of work: (a) *visualization (VZ)*; (b) *spatial relations (SR)*; (c) *closure speed (CS)*; (d) *closure flexibility (CF)*; (e) *perceptual speed (P)*; and (f) *visual memory (MV)*. These are abilities that have the potential to produce a different relationship with outcomes of operational effectiveness, and the use of a singular term has probably created problems with the generalization of findings across studies (Fincannon et al., in press-a). As a result of this issue, a quantitative meta-analysis of literature was needed to provide a preliminary classification of studies by the visuo-spatial construct(s) that were assessed.

Many constructs have been assessed in studies of UV operation, but research has argued that the majority of these studies actually use tests that load onto the *visualization (VZ)* factor (Fincannon et al., 2010a). While relationships of other visuo-spatial constructs are informative, the primary purpose in this meta-analysis was to focus analyses on the *visualization (VZ)* factor. This was particularly important for framing studies that aggregated tests across multiple visuo-spatial constructs to provide a single assessment of “spatial ability” (Chen & Terrance, 2008, 2009; Lathan & Tracey, 2002). As a result, a goal of this meta-analysis was to identify where *visualization (VZ)* has an impact on operational effectiveness and develop a more detailed analysis accordingly.

### Moderation by Outcome

After developing an understanding of visuo-spatial constructs in UV operation, the next step was to direct attention to performance outcomes. Existing research has found that perception

of remote environments was one of the more difficult aspects of UV operation (Burke et al., 2004; Burke & Murphy, 2004; Casper & Murphy, 2003; Woods et al., 2004), which suggests that emphasis should be placed on this dimension of performance. Research has extended this concept of perceptual performance by demonstrating how different aspects of perception (e.g., identification, localization, navigation) manifest themselves in UV operations (Fincannon et al., 2009, 2011, 2012; McDermott et al., 2005; Steinfeld et al., 2006). Furthermore, some outcomes (e.g., time to complete a mission) are not inherently visual in nature and need to be addressed separately. Differences between visual and non-visual dimensions of performance represent an important component for the development of the hypotheses in the following subsections.

In addition to drawing distinctions between types of performance (i.e., visual vs. non-visual), it should be noted that this review was designed to focus on a subset of robotic tasks. Specifically, this domain involves search (e.g., RSTA for reporting information) and navigation (e.g., transportation of equipment) tasks for UV operation. As a result of this, other tasks (e.g., laparoscopic surgery, manipulation of a robotic arm for bomb disposal, etc.) are probably also associated with visuo-spatial abilities, but beyond the scope of this meta-analysis. Furthermore, the tasks that were excluded from this analysis are likely to include other dimensions of performance (e.g., fine motor control) that extend beyond the emphasis on visual perception. The following sections will provide a review of outcomes for determining operational effectiveness with unmanned systems in search and navigation tasks.

### *Visual Performance*

Empirical research by my colleagues and me (Fincannon et al., 2010a, 2011) has observed three distinct dimensions of visual performance in UV operation. The first dimension involves the identification of objects in remote environments. The second dimension focuses on localization in a remote environment. In discussing localization, it should be noted that it not

only involves understanding the location of an operator's UV, but it also requires understanding of the location of targets and objectives in the environment. A third dimension of interest involves vehicle navigation. While navigation includes processes that are related to localization of components of a remote environment (e.g., UV, targets, future objectives, route obstructions), it goes further by requiring the operator to use that understanding of a UV in relation to a variety of other elements in the environment to develop a future-oriented strategy for moving through that environment. Each of these three dimensions are used in UV operation, and empirical findings have shown that the role within a multi-operator, multi-system team predicts the type of visual information that will be sought and/or provided to facilitate team effectiveness (Fincannon et al., 2011; McDermott et al., 2005). Therefore, these constructs will be considered in separate analyses.

Dimensions of visual performance are not always neatly divided into easily distinguishable dimensions of identification, localization, and navigation. Research by my colleagues and me (Fincannon et al., 2010a) has argued that RSTA tasks are unique in that they require proficiency across multiple measures of visual performance. Operators are not only responsible for knowing what a target is, but when the task involves two or more objectives, operators must also determine the location of each target in the environment to report that information correctly for each objective. A failure to do so might result in operators mistaking enemy units for friendly units or civilians. In addition to this, operators are also responsible for the planning of routes (e.g., to avoid route obstructions) and developing search strategies, which are necessary for a UV to navigate to objectives. Given this multidimensional nature, RSTA tasks deserve special consideration.

### *Non-Visual Performance*

There was also a need to pay attention to measures of performance that are not inherently visual. Specifically, mission completion time appears in several studies (Billings & Durlach, 2008, 2009; Chen et al., 2008). Reaction time can involve visual processes (e.g., *speeded* factors of visuo-spatial ability), but performance on visual tasks (e.g. navigation and identification) can also impact the amount of time that is necessary for an operator to complete a mission. For example, an operator who can quickly avoid obstructions in a route should be faster when navigating through an environment. In the context of cognitive abilities, the effects of a visuo-spatial construct on a non-visual measure of performance should be mediated by performance on the visual tasks that are required to complete a mission (see Figure 1). This mediation serves as the foundation to hypothesize an effect of visuo-spatial constructs on non-visual performance.

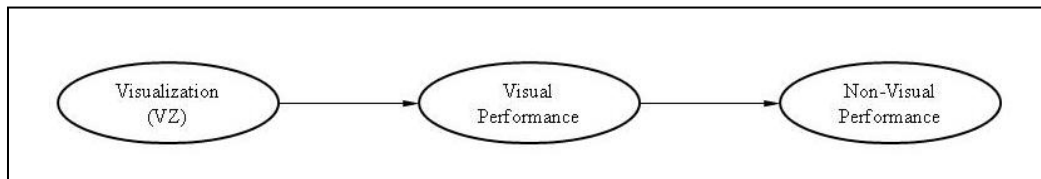


Figure 1. *Visualization (VZ)* impacting non-visual performance through a visual performance mediator

### *Mission Time by Task*

Mediation can lead to the hypothesis of *visualization (VZ)* influencing non-visual performance, but it may not impact all measures of visual performance in the same way. For example, high-ability operators appear to be faster with tele-operation tasks (Long, Gomer, More, & Pagano, 2011). Additionally, research has also demonstrated that high-ability operators are more likely to reach objectives on a reconnaissance task (Chen et al., 2011; Schuster et al., 2008).

Since operators must take time to survey the terrain for targets, including more objectives would be expected to increase mission time. Navigation skills are required to complete both tasks, and *visualization (VZ)* should universally make operators faster at moving between locations in an environment. Since reconnaissance tasks are unique in that they require operators to stop at objectives to report information, an association between *visualization (VZ)* and a higher frequency of stopping should increase mission time and counteract the positive association between *visualization (VZ)* and mission time via navigational speed (see Figure 2). As a result of this issue, there was a need to distinguish between simple navigation tasks and tasks that require operators to learn from the environment. Specifically, tasks that require an operator to learn from the environment are likely to involve multiple mediators that reduce the impact of *visualization (VZ)* on mission time.

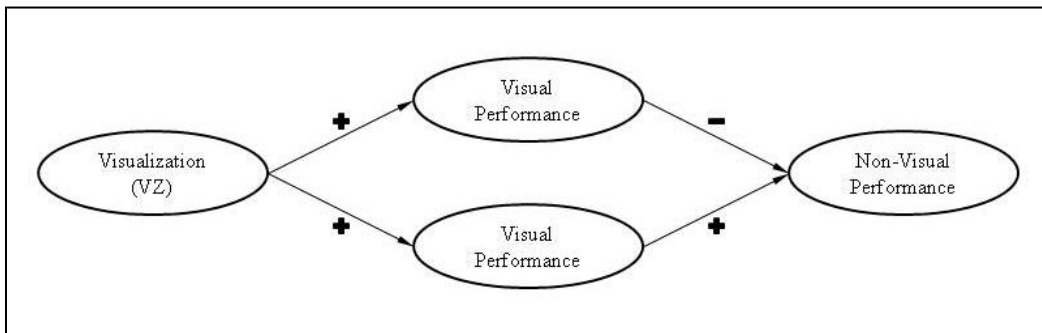


Figure 2. The impact of *visualization (VZ)* on non-visual performance being negated by two visual mediators

*Summary*

In summary, associations between *visualization (VZ)* and measures of visual and non-visual measures of performance deserve attention in UV operation. Visual measures include identification, localization, and navigation, and each of these are expected to be related to



*visualization (VZ)*. While non-visual measures of performance (i.e., time) may not be directly impacted by *visualization (VZ)*, they are impacted by visual dimensions of UV tasks, and *visualization (VZ)* would be expected to impact non-visual performance through its associations with visual performance. As a result of this, *visualization (VZ)* was expected to form associations with visual and non-visual dimensions of performance, such that higher ability should be associated with better performance (see Figure 2).

### Moderation by Other Constructs

As the previous sections have developed a foundation for hypothesizing associations between *visualization (VZ)* and performance, the current section builds on this by highlighting additional factors that are expected to moderate this relationship. Research by my colleagues and me (Fincannon et al., in press-a) has noted that findings will not always generalize across studies. When these variations emerge, we argued that differences can be attributed to underlying differences across unconsidered constructs, and an analysis of these issues can support the development of generalizable theory in HRI. The following subsections will review constructs that potentially moderate the impact of *visualization (VZ)* on performance.

#### *Autonomy in UV Operation*

Existing research has argued that the implementation of automation into a task can change how the human user works to accomplish that task (Riley & Parasuraman, 1997), and the implementation of autonomy in UV operation would be expected to produce the same pattern of effects. In the context of the relationships that are described in earlier sections, *visualization (VZ)* should improve visual and non-visual performance with unmanned systems. If the autonomy takes on the role of executing different visual tasks (see Fincannon et al., 2011, for an example with human teaming), the impact of the human operator on these tasks, and by extension *visualization (VZ)*, should diminish. This leads to the expectation that the impact of *visualization*

(VZ) on visual and non-visual performance should be moderated by the implementation of robotic autonomy.

When considering robotic autonomy, systems are complex, and this complexity can lead to differences in the types of autonomy that are implemented across UV research. For example, autonomy to support navigation (Rehfeld, 2006) can be very different from autonomy that is designed to cue operators to salient stimuli and support target identification (Chen & Terrance, 2008). This variation in autonomy creates a need to consider different types of autonomy (i.e., to support identification, localization, and/or navigation) across different dimensions of visual and non-visual performance.

#### *The Testbed in UV Operation*

Researchers that study UV operation have drawn a distinction between research from the laboratory and the field. This distinction has been extended to pursue theoretically driven hypotheses that comment on generalizability in HRI (Fincannon et al., in press-a). An example of this issue can be drawn by examining the distinction between “real” and simulated UVs that are used in research. Some researchers experiment with UVs that are commercially available as a final product that would be used in applied settings (Burke et al., 2004; Burke & Murphy, 2004; Casper & Murphy, 2003), whereas other researchers use virtual environments (Chen et al., 2008) or reversed engineered materials in scaled physical environments (Fincannon et al., 2011) to simulate UVs in applied settings. There are not only differences between end-user UVs and simulation of these systems, but there are also differences between physical and virtual environments (Huthmann, 2009). As a result of these differences, the testbed that was used to examine relationships in UV operation was examined as a factor that potentially moderates the association between *visualization (VZ)* and performance with UV systems.

### *Other Moderators*

Thus far, the review of literature has highlighted several constructs that potentially moderated the relationship between *visualization (VZ)* and operator performance with UVs. As these variables are not the only constructs that can potentially moderate these relationships, there was also a need to consider additional moderators that could be used as covariates in an empirical analysis. This section reviews relevant constructs that were used in this analysis.

An analysis regarding the impact of autonomy creates a need to simultaneously consider the further impact of reliability. Researchers have reported that realistic applications of autonomy are marked by reliability that can be below expected levels (Woods et al., 2004), and this has led researchers to manipulate the reliability of autonomy in UV operation (Chen, Barnes, & Kenny, 2011). As a failure of autonomy can create a need for human intervention, the human operator must still supervise autonomy in its execution of various tasks. If visuo-spatial abilities help an operator with visual supervision, the impact of *visualization (VZ)* on performance should vary in accordance with the need for an operator to intervene. Specifically, the impact of *visualization (VZ)* on UV performance should be moderated by the reliability of autonomy to perform a visual task, such that the effect size will be greatest under conditions of low reliability.

Prior research within our group has noted that the effects of visuo-spatial ability on performance can diminish with an increase in team size (Sellers et al., 2011). In order to control for this, studies were coded in accordance with the number of operators that were used to complete the task.

One of the goals of UV research has been to maximize the number of vehicles that an operator can control at a single time. Given that perception from a single UV has been found to be difficult (Fincannon et al., in press-b; Woods et al., 2004), increasing the number of UVs should subsequently increase the difficulty of perceiving remote environments. Since

*visualization* (VZ) should be associated with complex visual performance, an increase in visual complexity that is associated with attending to multiple UVs should place a greater emphasis on the need for this ability to achieve higher levels of performance. Otherwise stated, one might expect that the relationship between *visualization* (VZ) and performance should increase as the number of UVs increases.

A testbed (i.e., physical vs. virtual environment) can be used to examine issues surrounding the external validity of research (Fincannon et al., in press-a), but this was not the only approach that could be used to study this relationship. For example, researchers in physical environments incorporate end-user vehicles (Long et al., 2011), whereas other studies simulate end-user vehicles (Fincannon et al., 2011). As a result of this, the moderator analysis also considered simulated, as opposed to end-user, vehicles as a covariate.

Discussions of UV operation have noted that differences in the perspective (e.g., ground vs. air) may moderate the relationship between *visualization* (VZ) and an outcome of interest (Diaz & Sims 2003; Fincannon et al., 2010a). In order to control for this, the perspective that is used to provide real-time imagery to the operator (i.e., air, ground, or both) was also recorded.

A final covariate for moderation involved gender. Specifically, researchers have found consistent gender difference across measures of spatial ability (Hegarty & Waller, 2005; Voyer, Voyer, & Brydent, 1995). Furthermore, differences in gender appear to include differences in spatial processing (Vogel, Bowers, & Vogel, 2003) that alters the predictive ability of tests across gender (Fatolitis, 2008). If the predictive nature of a test is truly moderated by gender, the percentage of women or men within a study should be considered as a covariate.

### *Summary*

In summary, there are several constructs that are likely to moderate the impact of *visualization* (VZ) on performance with UVs in remote environments. These moderators include

levels of autonomy, the reliability of autonomy, the number of operators, the number of UVs, the testbed, the use of vehicle simulation (i.e., end-user vehicle vs. simulation of end-user vehicle), vehicle type (i.e., UAV, UGV, or both), and gender. All of these relationships provide a general commentary on the theory of how *visualization (VZ)* forms associations with measures of visual performance.

### Statement of Hypotheses

The previous sections reviewed literature in the domain of UV operation to support several hypothesized relationships. These hypotheses are summarized in Figure 3 and the following subsections.

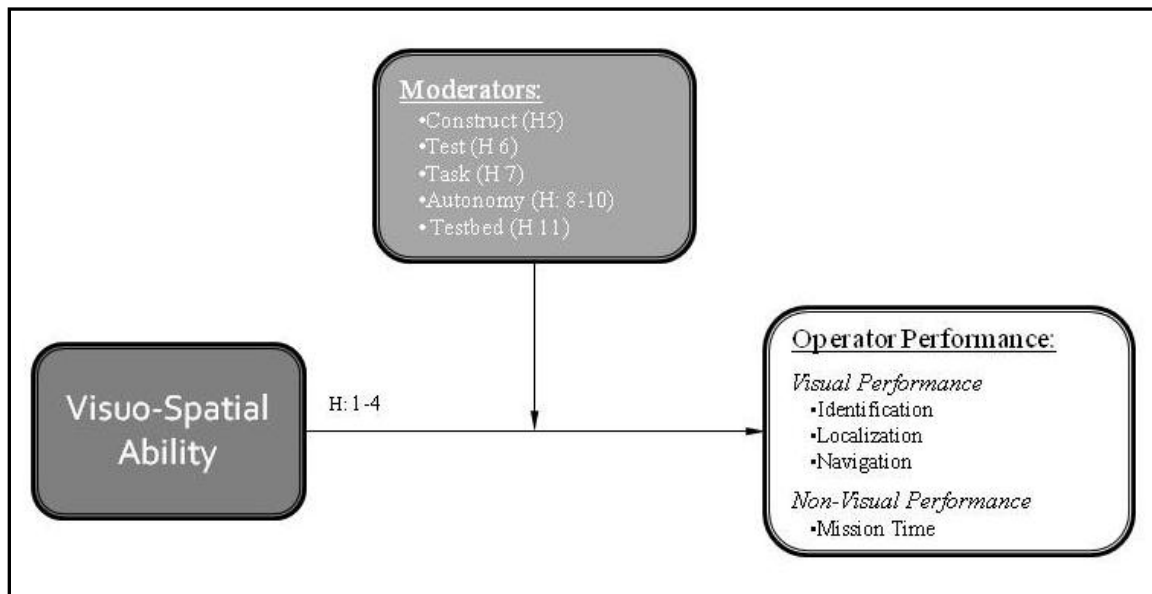


Figure 3. Summary of hypotheses

### *Main Effects*

The review of the literature highlighted six major factors of visuo-spatial ability and five major dimensions of performance. Each of these abilities would be expected to form a positive association with each of these outcomes. The resulting hypotheses are summarized as follows.

*H (1 to 4): Visuo-spatial ability is negatively associated with mission completion time (1) and positively associated with identification (2), localization (3), and navigation (4).*

### *Interaction by Construct*

Research has noted that the majority of published material uses measures that load onto the *visualization (VZ)* factor (Fincannon et al., 2010a). Furthermore, the review of literature that has been presented in this work suggests that the *visualization (VZ)* factor is more important for dimensions of operator performance that commonly appear in published work. This led to the following hypothesis.

*H 5: The effect of visuo-spatial ability on performance is moderated by the visuo-spatial construct, such that measures of visualization (VZ) form the strongest association with performance outcomes.*

### *Interaction by Test*

Carroll's (1993) *visualization (VZ)* factor represents a broad construct that includes measures that have been associated with different factors in past research. For example, the cube comparison test (Ekstrom et al., 1976) and perspective taking tests (Lohman et al., 1987) have

been cited as representing different spatial orientation factors. As a result of this, I hypothesized that:

*H 6: The effect of visualization (VZ) on an outcome is moderated by the measure that is used to assess the VZ construct.*

#### *Interaction by Task*

As discussed in the literature review, the association between *visualization (VZ)* and mission completion time is expected to be mediated by visual performance. However, visual mediators are expected to vary across UV tasks, and I hypothesized that:

*H 7: The effect of visualization (VZ) on a mission completion time is moderated by mission task, such that the association is strongest with tasks that only require navigation and weakest with tasks that require an operator learn (i.e., identification and/or localization in addition to navigation) from an environment.*

#### *Interaction by Autonomy*

In the review of literature, we stated that the implementation of autonomy would be expected to reduce the impact of *visualization (VZ)* on performance. Since autonomy can be implemented across a variety of different performance outcomes, I formulated the following hypotheses:

*H 8: The effect of visualization (VZ) on identification is moderated by autonomy that is designed to support identification, such that the association will be strongest under the lowest levels of autonomy.*

*H 9: The effect of visualization (VZ) on localization is moderated by autonomy that is designed to support localization, such that the association will be strongest under the lowest levels of autonomy.*

*H 10: The effect of visualization (VZ) on navigation is moderated by autonomy that is designed to support navigation, such that the association will be strongest under the lowest levels of autonomy.*

#### *Interaction by Testbed*

Researchers have noted that findings do not necessarily generalize across studies (Fincannon et al., in press-a). In order to examine this issue, I hypothesized that:

*H 11: The effect of visualization (VZ) on an outcome is moderated by the testbed that is used to assess the relationship.*



## CHAPTER 3: METHOD

### Search & Inclusion Criteria

A literature search was conducted in order to find studies that were relevant for inclusion in this meta-analysis. This review included databases such as: (a) PsychInfo; (b) IEEE; (c) Google Scholar; (d) HFES Proceedings; and (e) Dissertation Abstracts. A variety of search terms (see Appendix A for list) were used to find relevant material. In order to avoid the file drawer effect, a search was conducted to locate unpublished material. This included sending out a letter (see Appendix B) via the HFES and HRI list-serves. Publications and unpublished datasets were then screened for inclusion within the meta-analysis.

Inclusion criteria were defined by three major components. First, the study had to include at least one of the six visuo-spatial constructs (i.e., visualization, spatial relations, closure speed, closure flexibility, perceptual speed, and visual memory). While these may appear under one of the aforementioned headings, measures also appear under the heading of spatial ability. Second, the study was required to measure at least one of the four performance outcomes that were presented in the literature review (i.e., identification, localization, navigation, and mission time). Third, the association between predictor and outcome needed to be assessed from an egocentric perspective within a remote environment. Data was not excluded on the basis of sample size, participant populations, or year of publication. Upon reviewing these criteria, 85 studies were found to be relevant for inclusion.

In several instances, screening found that results from the same dataset were presented across multiple publications (e.g., dissertation and journal publication). In order to avoid double coding of results, data was coded to reflect the number of unique datasets (i.e.,  $K$ ). Of the 85 studies of published and unpublished data, 54 dataset appeared to be unique for this meta-analysis (i.e.  $K=54$ ).

## Coding of Studies

In order to conduct a meta-analysis, quantitative data was extracted from each study. This began with recording of effect sizes and continued with the coding of moderating variables. The following sub-sections provide details regarding the steps of this process.

### *Data Extraction*

Two types of data were extracted from the effects that were recorded from each dataset. The first involved the use of correlations as a measure of effect size for primary hypothesis testing. When correlations were unavailable, statistics that were provided in the analysis (e.g.,  $F$  statistic, mean, standard deviation, etc.) were transformed into correlations. As discussed by Hunter and Schmidt (2004), there are various factors (e.g., measurement error, dichotomization) that impact the relationships between measures of a construct; this analysis used corrections that were developed by Hunter and Schmidt to attenuate the observed correlations accordingly. The purpose of this approach was to increase power to observe significant effects in a moderator analysis (Borenstein, Hedges, Figgins, & Rothstein, 2009).

The second type of data for extraction included *p-values*. Since exact values are not typically provided by authors (e.g.,  $p < .05$ ), this value was determined by using the statistic that was provided by the authors to obtain an exact *p-value*. The purpose of obtaining this information was to conduct a file drawer analysis (i.e., the number of non-significant effects that are needed to negate the findings of the meta-analysis).

Based on the reporting of data and communication with authors, effect sizes could not be obtained for four studies (i.e., 4 unique datasets). This served as further exclusion criteria that reduced the number of studies and datasets. The final analysis was conducted on 81 studies that represented 50 unique datasets.

### *Rater Information & Training*

Two raters coded variables that were used as moderators in this analysis. For training, these raters were provided with descriptions of relevant material (see Appendix D) and 12 unique datasets to assess variables. When there were disagreements in coding, the specific disagreements were reviewed with each rater, and they were asked to code variables. Once coding was found to be sufficiently reliable, each rater coded studies independently for the analysis.

*Kappa* was used to assess interrater reliability. As illustrated in Appendix E, this metric provides an estimate of observed agreement divided by all possible agreement. Since a *Kappa* of .70 was considered to be a high level of agreement for this metric (Landis & Koch, 1977), a *Kappa* of .70 was used as the standard for training before raters were allowed to proceed with coding. The following subsections report reliability from this training.

### *Predictors*

Predictors were coded in three different ways. First, studies were coded by the exact test that was used to assess a visuo-spatial ability (*Kappa* = 1.0). Second, tests were coded by the construct listed in Tables 1 through 6. Third, tests of the *visualization* (VZ) construct were coded by the set of sub-dimensions that were described in Table 1. Coding of data to H5 reflected:

- Visualization (1) vs. Other (0)
- Spatial Relations (1) vs. Other (0)

Coding to test H6 reflected:

- Cube Comparison Test (1) vs. Visualization (0)
- Guilford-Zimmerman Test of Spatial Orientation (1) vs. Visualization (0)
- Gugerty Spatial Orientation Test (1) vs. Visualization (0)
- Hegarty Object Perspective Test (1) vs. Visualization (0)

### *Outcomes*

As discussed in the introduction, this analysis examined four different dimensions of performance (i.e., identification, localization, navigation, and mission time). Mission time was a relatively simple outcome to code, but as noted in the introduction, certain tasks (e.g., reconnaissance) require a combination of different dimensions of visual performance. In order to address this, coding of visual performance outcomes was given special attention, and a procedure for coding this outcome (see Table 7) was developed from prior work (Fincannon et al., 2011). Raters received descriptions of this material (see Appendix D) for coding.

Table 7. Description of visual performance outcomes for coding

<b>Outcome Description (<i>Kappa</i>)</b>	<b>Task Description</b>
Identification (.91)	The outcome requires a participant, or autonomy, to recognize and report an object or target by name
Localization (.93)	The outcome requires a participant, or autonomy, to determine where a vehicle, target, or object is located within an environment
Navigation (1.0)	The outcome requires a participant, or autonomy, to move an entity (i.e., physical UV or virtual representation) through a remote environment
Mission Time (1.0)	The amount of time that was taken to complete a navigation or reconnaissance task

### *Moderators*

This section reviews the coding of variables that were hypothesized to moderate the impact of *visualization (VZ)* on performance. These variables included: (a) distinguishing between physical and virtual environments (*Kappa* = 1.0); (b) recording the number of UVs (*Kappa* = 1.0); (c) reliability (*Kappa* = 1.0); and (d) level of autonomy (see Tables 8, 9, and 10 for *Kappa*). The testbed environment, number of UVs, and reliability were recorded as objective numbers from each study, and there was no disagreement (i.e., a *Kappa* of 1.0 was obtained for

each metric). In contrast, the level of perceptual autonomy was more subjective, and a coding scheme was created for assessment.

Table 8. Levels of autonomy to support identification (*Kappa* = .84)

<b>Degree of Autonomy</b>	<b>Outcome Description</b>	<b>Task Description</b>
High	Fully Autonomous Identification	The autonomy identifies targets and objects that are encountered for the participant
	Detection Support	The participant is informed of the presence of targets and objects to identify
Low	Operator Identification	The participant is not provided with any autonomous support for the identification of targets and objects

Table 9. Levels of autonomy to support localization (*Kappa* = .79)

<b>Degree of Autonomy</b>	<b>Outcome Description</b>	<b>Task Description</b>
High	Autonomous Localization	The autonomy provides a participant with information about location (e.g., coordinates) in a remote environment
Low	Operator Localization	The participant must perform a localization task without autonomous support

Table 10. Levels of autonomy to support navigation (*Kappa* = .92)

<b>Degree of Autonomy</b>	<b>Outcome Description</b>	<b>Task Description</b>
High	Fully Autonomous Navigation	The autonomy navigates through an environment without feedback from the participant
	Supervisory Control of UV Navigation	The autonomy alters the route of a UV or avatar, and the participant can review and alter this decision
	Semi-Autonomous Navigation	A participant uses waypoints to plan a route through a remote environment that is executed by a UV or avatar
Low	Teleoperation	The participant must manually maneuver a UV or avatar through a remote environment

While H8, H9, and H10 were based on research with human teams, the concept of varying levels of autonomy across multiple types of autonomy appears in other work (Parasuraman, Sheridan, & Wickens, 2000). The ten levels of automation that were described by Parasuraman and colleagues served as a reference for the development of a coding scheme to describe perceptual autonomy as a continuous variable. Pilot testing of this frame led to the categories listed in Tables 8, 9, and 10.

#### *Control Variables*

In addition to the primary constructs of interest, this meta-analysis considered control variables that could also impact the outcomes of interest in this analysis. These variables include:

- The reliability of autonomy [*Kappa* = 1.0]
- Operator number [*Kappa* = 1.0]
- Vehicle number [*Kappa* = 1.0]
- The use of simulated vehicles [*Kappa* = .77]
- Domain of interest (i.e., generic task vs. designed to study HRI) [*Kappa* = 1.0]
- Vehicle type (i.e., UAV, UGV, or both) [*Kappa* = .78]
- Gender (i.e., recorded as percentage of women) [*Kappa* = 1.0]

#### Analysis

Two different procedures were used to assess the hypotheses. Each of these processes began with the extraction of correlations, which included the transformation of other statistics into correlation coefficients (see Appendix E for formulas). As different procedures were used to assess these different hypotheses, separate subsections are provided below to address each of these methods.

### *Main Effect Analysis*

Procedures that were described by Hunter and Schmidt (2004) served as a primary framework for assessing the main effects. After correlations had been attenuated, they were aggregated to provide an assessment of H1 through H4. This included: (a) the total number of participants per estimate; (b) the total number of studies per metric; (c) a mean estimate from uncorrected correlations with a confidence interval; (d) a mean estimate from corrected correlations with a confidence interval; (e) a failsafe number; and (f) a statistic to comment on unexpected moderation.

Excel was used to calculate all effects for H1 through H4 (see Tables 12 and 17). The specific procedure of obtaining this data involved the following:

1. Raw correlation coefficients were obtained for each study.
2. Reliability coefficients were used to attenuate the correlation coefficients and provide a corrected correlation coefficient. If a reliability coefficient was not available, source material (Ekstrom et al., 1976) was used to provide estimates of reliability. If source material did not provide estimates of reliability (see Gugerty & Brooks, 2004), an average across all measured reliability coefficients was used as a substitute.
3. For each study, an aggregate was created for the corrected and uncorrected correlation coefficients within each study.
4. Once an aggregate was obtained for each study, the corrected and uncorrected correlation coefficients at the level of each study were used to determine the confidence intervals for H1 through H4.
5. *P-values* for uncorrected, study-level correlations were used to provide estimates for the file drawer effect and *Q*-statistic.

6. In order to be consistent with the American Psychological Association's (APA, 2010) Meta-Analysis Reporting Standards (MARS), data was screened for normality and power. This information was provided for every estimate of a main effect.

### *Multivariate Analysis*

In addition to mean estimates, the meta-analysis conducted meta-regression to test H5 through H11. Since the variance of a correlation depends on the strength of the correlation (Borenstein et al., 2009), corrected correlations were transformed into *Fisher's Z* (see Appendix E for formulas). *Fisher's Z* was then weighted according to the inverse of the sampling variance for final analysis. As this method did not control for the reliability of the cognitive ability test, this coefficient was used as a covariate in the moderator analysis (Borenstein et al., 2009).

In addition to using *Fisher's Z* to control for non-random variation of the standard deviation, it was common for individual studies (e.g., Rehfeld, 2006) to examine multiple constructs (e.g., visuo-spatial ability and levels of autonomy). Since the hypotheses involved assessing these differences, there was a need to compare measures within a study, which were not independent of each other (Hunter & Schmidt, 2004). While the method of obtaining a study-level aggregate that was described above could have been used, it was not optimal for performing the desired comparison. In order to control for the non-independence of within-study assessment, multi-level linear modeling was used instead (Kalaian & Raudenbush, 1996; Tabachnick & Fidell, 2007). Specifically, each study was used as a categorical variable to control for the impact of nesting, and a moderator variable was coded for every effect within each study. Once the coding had been completed in this manner, statistical analyses were performed with SPSS software.



## CHAPTER 4: RESULTS

The results of the meta-analysis are presented in this section and consist of three subsections. The first subsection presents analyses that were necessary for data screening. The second subsection presents analyses that were relevant to H1 through H4, which used techniques by Hunter and Schmidt (2004) to examine the main effects. The third subsection present analyses that were relevant to H5 through H11, which used multi-level linear modeling (Kalaian & Raudenbush, 1996; Tabachnick & Fidell, 2007) to test for moderation.

### Data Screening

Data screening began with a review of studies to ensure that different publications were presenting analyses from unique datasets. From the 81 studies that were identified, 50 datasets were found to be unique. From this data, 518 associations were extracted to assess relevant hypotheses. A breakdown of the number of datasets and associations by outcome are illustrated in Table 11.

Table 11. Description of visual performance outcomes for coding ( $k$  = number of studies)

<b>Outcome</b>	<b><i>k</i></b>	<b># of Associations</b>	<b>Mean # of Associations Per Study</b>	<b>Median # of Associations Per Study</b>
Identification	29	195	6.72	5
Localization	15	119	7.93	3
Navigation	18	120	6.67	4.5
Mission Time	18	84	4.67	3

At first glance, the total number of associations may appear to be somewhat large, but a closer inspection of the data revealed that this was primarily attributed to factors that were considered in the moderator analysis. For example, studies that use multiple predictors (see H5

and H6) or manipulates autonomy (see H8 through H 10) would easily provide three to five associations (see Medians of Table 11). In some cases, multiple moderators varied within a single study, and in one of these studies, this variation (i.e., across 4 predictors, 3 levels of autonomy, and 2 levels of reliability) produced 24 associations per outcome. Although the number of associations per study may appear to be large, this level of extraction was necessary for testing hypotheses.

Another source of variation emerged with studies that used multiple outcomes that were designed to assess the same construct. For example, one study by Rehfeld (2006) used two measures of localization to test hypotheses about performance outcomes. In order to minimize bias within this meta-analysis that might be associated with selecting one measure over another, both associations were extracted and aggregated to represent a single relationship. Therefore, the second reason for a high number of associations per study was attributed to a methodology that was intended to improve the design of this meta-analysis.

### Main Effects

This section presents findings with respect to main effects of all visuo-spatial abilities on performance, which are presented in H1 through H4. Findings for these hypotheses are summarized in Table 12.

H1 stated that visuo-spatial ability would be expected to form a negative association with mission completion in remote environments. Both the corrected ( $\rho = -.287$ ) and uncorrected ( $r = -.318$ ) correlation coefficients were moderately strong. As indicated by the confidence interval around each estimate (Table 12), H1 was supported.

For moderation, the significance of the  $Q$  statistic ( $Q = 40.9$ ) and variance explained by the attributes ( $\%VA = 45\%$ ) indicated that moderators are likely to change the relationship between *visualization* ( $VZ$ ) and mission completion time. As indicated by the credibility

intervals, this estimate ranged from -.10 to -.54 for corrected correlation coefficients and -.02 to -.56 for uncorrected correlation coefficients.

H2 stated that visuo-spatial ability was expected to form a positive association with target identification in remote environments. Both the corrected ( $\rho = .265$ ) and uncorrected ( $r = .235$ ) correlation coefficients were weak. As indicated by the confidence interval around each estimate (Table 12), these differences were significantly greater than zero and supported H2. The lack of variation around the corrected correlation was attributed to a standard deviation of zero, which may point to second-order sampling error.

For moderation, the significance of the  $Q$  statistic ( $Q = 20.8$ ) indicated that moderators are likely to change the relationship between *visualization* (VZ) and identification. As indicated by the credibility intervals, this estimate ranged from .07 to .40 for uncorrected correlation coefficients. As with the confidence interval, problems with that standard deviation carried over to the credibility interval and estimates of variance explained by the attributes. If second-order sampling error is present, it would also bias these statistics.

H3 stated that visuo-spatial ability would be expected to form a positive association with localization in remote environments. Both the corrected ( $\rho = .373$ ) and uncorrected ( $r = .341$ ) correlation coefficients were moderately strong and greater than zero. The confidence interval around each estimate (Table 12) was significant and supported H3. As with target identification, lack of variation in the confidence interval for the corrected correlation was attributed to a standard deviation of zero, which again indicated that second-order sampling error may be present.

The significance of the  $Q$  statistic ( $Q = 11.5$ ) indicated that moderators are likely to change the relationship between *visualization* (VZ) and localization. As indicated by the credibility intervals, this estimate ranged from .22 to .47 for uncorrected correlation coefficients.

As with the confidence interval, problems with that standard deviation carried over to the credibility interval and estimates of variance explained by the attributes. This could also be attributed to second-order sampling error.

H4 stated that visuo-spatial ability would be expected to form a positive association with navigation performance in remote environments. Both the corrected ( $\rho = .304$ ) and uncorrected ( $r = .277$ ) correlation coefficients were moderately greater than zero, and the confidence interval around each estimate (Table 12) supported H4.

For moderation, the significance of the  $Q$  statistic ( $Q = 32.6$ ) and variance explained by the attributes ( $\%VA = 56\%$ ) indicated that moderators are likely to change the relationship between *visualization* ( $VZ$ ) and navigation. As indicated by the credibility intervals, this estimate ranged from .14 to .46 for corrected correlation coefficients and .06 to .50 for uncorrected correlation coefficients.

Appendix F presents the analyses for normality and power. These results found that the power to detect an effect size of .20 was strong (i.e., .98 to 1.00). Therefore, the estimates of these effects are expected to be stable. Data was found to be normal for estimates of localization, navigation, and mission completion time. However, statistics for identification indicated that the data had a positive skew and leptokurtic distribution.

In summary, the corrected estimates of association between visuo-spatial ability and performance were significant for all outcomes. Some of these associations were strong, but this was not unusual in that visuo-spatial abilities have a long history of success in predicting performance with piloting and technical training (Hegarty & Waller, 2005). The results that are presented in Table 12 appeared to indicate that these relationships were consistent with existing literature, which formed the basis for H1 through H4. While the non-normality of the distribution

for identification outcomes had the potential to be problematic, the proposed procedures and re-analyses of the following section appeared to resolve the issue.

Table 12. Main effects of all visuo-spatial measures on performance

Outcome	<i>N</i>	<i>K</i>	<i>r</i>	<i>SD<sub>r</sub></i>	<i>CI<sub>r</sub></i> 95%		<i>CrI<sub>r</sub></i> 80%		<i>ρ</i>	<i>SD<sub>ρ</sub></i>	<i>CI<sub>ρ</sub></i> 95%		<i>CrI<sub>ρ</sub></i> 80%		%VA	<i>fde</i>	<i>Q</i>
Identification	1154	29	.24	.13	.19	.28	.07	.40	.27	0	.27	.27	.27	.27	100%	80	20.8*
Localization	1054	15	.34	.10	.29	.39	.22	.47	.37	0	.37	.37	.37	.37	100%	6	11.5*
Navigation	971	18	.28	.17	.20	.36	.06	.50	.30	.12	.25	.36	.14	.46	56%	22	32.6*
Mission Time	799	18	-.29	.21	-.19	-.38	-.02	-.56	-.32	.17	-.24	-.40	-.10	-.54	45%	21	40.9*

*Note:* *N* = total number of participants across studies; *k* = number of studies; *r* = mean observed correlation (uncorrected); *SD<sub>r</sub>* = standard deviation for *r*; *CI<sub>r</sub>* 95% = confidence interval for *r*; *CrI<sub>r</sub>* 80% = credibility interval for *r*; *ρ* = mean observed correlation (corrected); *SD<sub>ρ</sub>* = standard deviation for *ρ*; *CI<sub>ρ</sub>* 95% = confidence interval for *ρ*; *CrI<sub>ρ</sub>* 80% = credibility interval for *ρ*; %VA = percent of the variance explained by the attributes; *fde* = number of non-significant effects to negate the significance of *r*; *Q* = moderator statistic, \* *p* < .05

### Moderator Analysis

H5 stated that the association between visuo-spatial ability and performance would be moderated by visuo-spatial construct, such that the strongest associations would form with the *visualization (VZ)* factor of Carroll's (1993) model. In addition to the *visualization (VZ)* factor, the regression analyses in this section also considered unique variation that was attributed to Carroll's (1993) *spatial relations (SR)* factor. The final models of the analyses for each outcome (i.e., mission completion time, identification, localization, and navigation) are presented in Tables 13, 14, 15, and 16.

For mission completion time, *visualization (VZ)* was found to moderate the association between ability on performance,  $F(1, 67) = 7.58, p < .01$ . In a manner that was consistent with hypotheses, associations were found to be stronger with measures of the *visualization (VZ)* factor (Table 13). No unique variance was attributed to *SR*.

Table 13. Model for testing the moderating effects of construct on mission completion time ( $k = 18$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	-0.50	0.43	1.16
Visualization	0.30	0.06	5.34*
Spatial Relations	0.08	0.09	0.92

*Note: Associations reverse coded, such that positive coefficients reflect stronger relationships; for visualization, 1 = VZ & 0 = other; for spatial relations, 1 = SR & 0 = other. \*  $p < .05$*

Next, *visualization (VZ)* was found to moderate the association between ability and target identification,  $F(1, 188) = 6.00, p < .05$ . As illustrated in Table 14, the direction of this association was consistent with hypothesis 5, such that relationships were stronger for measures of visualization. *SR* was not a significant moderator.

Table 14. Model for testing the moderating effects of construct on identification ( $k = 29$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.47	0.05	1.34
Visualization	0.12	0.35	2.45*
Spatial Relations	0.00	0.07	0.06

Note: For visualization,  $1 = VZ$  &  $0 = other$ ; for spatial relations,  $1 = SR$  &  $0 = other$ . \*  $p < .05$

With measures of localization, *visualization* ( $VZ$ ) was found to be a significant moderator,  $F(1, 115) = 4.65, p < .05$ . However, closer inspection of Table 15 indicated that the direction of this association was opposite to H5. Specifically, associations were found to be weaker for measures of visualization. The non-visualization constructs included *SR* (i.e., card rotation test), *closure speed [CS]* (i.e., the embedded figures test), and *visual memory [MV]* (i.e., building memory test), and correlations for these constructs were consistent with this trend. Unique consideration of *SR* did not account for any additional variance. The number of studies per construct was not sufficient to permit hypothesis testing beyond what was presented in Table 15.

Table 15. Model for testing the moderating effects of construct on localization ( $k = 15$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.55	0.13	4.25*
Visualization	-0.07	0.03	2.16*
Spatial Relations	-0.04	0.06	0.59

Note: For visualization,  $1 = VZ$  &  $0 = other$ ; for spatial relations,  $1 = SR$  &  $0 = other$ . \*  $p < .05$

For measures of vehicle navigation (Table 16), *visualization* ( $VZ$ ) was found to moderate the association between ability and performance in the hypothesized direction,  $F(1, 107) = 7.95, p < .01$ . As with all of the other metrics, *SR* was not a significant predictor in the model.



Table 16. Model for testing the moderating effects of construct on navigation ( $k = 18$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.15	0.32	0.46
Visualization	0.11	0.04	2.82*
Spatial Relations	0.08	0.06	1.45

*Note: For visualization, 1 = VZ & 0 = other; for spatial relations, 1 = SR & 0 = other. \*  $p < .05$*

Since measures of the visualization construct were consistently found to moderate the association between visuo-spatial ability on performance, data from Table 12 was re-aggregated to only include these metrics, and the results are presented in Table 17. The new aggregate for measures of target identification was calculated across 25 studies. Re-aggregation did not alter the number of studies for any of the other outcomes. As the analyses from Table 13 through 16 illustrate, the new means were weaker for measures of localization and stronger for measures of identification, navigation, and mission completion time. Analyses of the confidence intervals indicated that all of these associations were still significant. Because confidence intervals for the corrected correlation with measures of identification and localization showed no variation, problems with second-order sampling error still appeared to be present. The  $Q$  statistics were still significant for each outcome, which indicated the likelihood of additional moderators.

Appendix F presents the analyses for normality and power. These results found that the power to detect an effect size of .20 were still strong (i.e., 1.00 for all outcomes). Furthermore, the re-aggregated data was found to be normal for estimates of all outcomes. As a result, this readjustment of data to focus on the *visualization* (VZ) factor appeared to improve the distributions of data, such that they were more in line with assumptions for tests of significance.

Table 17. Main effects of *visualization (VZ)* on measures on performance

<b>Outcome</b>	<b><i>N</i></b>	<b><i>K</i></b>	<b><i>r</i></b>	<b><i>SD<sub>r</sub></i></b>	<b><i>CI<sub>r</sub></i> <b>95%</b></b>	<b><i>CrI<sub>r</sub></i> <b>80%</b></b>	<b><i>ρ</i></b>	<b><i>SD<sub>ρ</sub></i></b>	<b><i>CI<sub>ρ</sub></i> <b>95%</b></b>	<b><i>CrI<sub>ρ</sub></i> <b>80%</b></b>	<b>%VA</b>	<b><i>fde</i></b>	<b><i>Q</i></b>
Identification	971	25	.28	.14	.22 .34	.10 .46	.31	0	.31 .31	.31 .31	100%	51	20.4*
Localization	1054	15	.33	.11	.28 .39	.19 .47	.37	0	.37 .37	.37 .37	100%	11	13.6*
Navigation	971	18	.29	.16	.22 .37	.08 .50	.32	.11	.26 .37	.17 .46	59%	19	30.4*
Mission Time	799	18	-.32	.19	-.23 -.41	-.08 -.56	-.35	.15	-.28 -.42	-.16 -.54	50%	19	35.9*

*Note:* *N* = total number of participants across studies; *k* = number of studies; *r* = mean observed correlation (uncorrected); *SD<sub>r</sub>* = standard deviation for *r*; *CI<sub>r</sub>* 95% = confidence interval for *r*; *CrI<sub>r</sub>* 80% = credibility interval for *r*; *ρ* = mean observed correlation (corrected); *SD<sub>ρ</sub>* = standard deviation for *ρ*; *CI<sub>ρ</sub>* 95% = confidence interval for *ρ*; *CrI<sub>ρ</sub>* 80% = credibility interval for *ρ*; %VA = percent of the variance explained by the attributes; *fde* = number of non-significant effects to negate the significance of *r*; *Q* = moderator statistic, \* *p* < .05

### *Visualization Hypotheses*

Testing for H5 clearly indicated that the strength of associations was moderated by the use of tests that load onto Carroll's (1993) visualization construct. Due to the low frequency across studies and the likelihood of confounding further hypothesis testing with the visualization construct, markers of other constructs (i.e. spatial relations, closure speed, closure flexibility, perceptual speed, and visual memory) were excluded from further analyses. Therefore, testing of H6 through H11 used the datasets that are presented in Table 17, and the correlation matrix for the moderator variables with this data was presented in Appendix G.

H6 stated that the association between *visualization (VZ)* and performance would be moderated by the tests that are used to examine these relationships. As illustrated in Table 18, this hypothesis was only supported with respect to the spatial orientation test (Gugerty & Brookes, 2004) for measures of target identification,  $F(1, 53) = 9.27, p < .01$ . Specifically, associations were found to be stronger with the test of spatial orientation than traditional markers of visualization. Analyses with other tests (i.e., Guilford-Zimmerman Test of Spatial Orientation, Cube Comparison Test, and Object Perspective Test) and outcomes were not found to be significant. H8 stated that the association between *visualization (VZ)* and target identification would be moderated by autonomy that is designed to support identification, but Table 18 indicates that H8 was not supported

H7 stated that the association between *visualization (VZ)* and mission completion time would be moderated by task type. Table 19 illustrates task type was a significant moderator,  $F(1, 39) = 9.74, p < .005$ . As indicated by the coefficient, the associations were weaker for reconnaissance tasks, and H7 was supported. Supplemental analyses indicated that gender was also a significant moderator,  $F(1, 39) = 9.03, p < .005$ , such that associations between *visualization (VZ)* and mission completion time were stronger for men than women.

Table 18. Model for variables moderating the effect of *visualization* (VZ) on identification ( $k = 25$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.55	0.55	0.26
GB SpO	0.22	0.07	3.05*
Identification Autonomy	-0.06	0.07	0.82

Note: For SpO,  $1 =$  Gugerty & Brooks SpO Test &  $0 =$  VZ; \*  $p < .05$

Table 19. Model for variables moderating the effect of *visualization* (VZ) on mission completion time ( $k = 18$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	-0.83	0.60	1.39
% Women	-1.30	0.43	3.00*
Task Type	-0.21	0.07	3.12*

Note: Associations reverse coded, such that positive coefficients reflect stronger relationships; For Task Type,  $1 =$  Reconnaissance &  $0 =$  Navigation; \*  $p < .05$

H9 stated that the association between *visualization* (VZ) and localization would be moderated by autonomy that is designed to support localization. As illustrated in Table 20, autonomy to support localization was found to be a significant moderator,  $F(1, 63) = 14.60, p < .001$ , but it was not in the hypothesized direction. Instead of decreasing the strength of the associations, providing more information about the location of objects in the remote environment was found to increase the association between visualization and localization performance. Table 20 also indicates that gender [ $F(1, 63) = 7.76, p < .01$ ], intended application [ $F(1, 63) = 7.93, p < .01$ ], autonomy to support identification [ $F(1, 63) = 6.57, p < .05$ ] and autonomy to support navigation [ $F(1, 63) = 7.03, p < .05$ ] moderated the association between visualization and localization performance. Specifically, associations between *visualization* (VZ) and localization

performance were stronger for men, for research that was intended to study HRI, for higher levels of identification support, and for lower levels of navigation support.

Table 20. Model for variables moderating the effect of *visualization* (VZ) on localization ( $k = 15$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.28	0.13	2.10*
% Women	-0.55	0.20	2.79*
HRI Intent	0.15	0.05	2.82*
Environment	0.15	0.06	2.49*
Identification Autonomy	0.19	0.07	2.56*
Navigation Autonomy	-0.05	0.02	2.65*
Localization Autonomy	0.15	0.04	3.82*

Note: For HRI Intent, 1 = Indented for HRI & 0 = other; For Environment, 1 = Physical & 0 = Virtual; \*  $p < .05$

H10 stated that the association between *visualization* (VZ) and navigation performance would be moderated by autonomy that is designed to support vehicle navigation. Navigation autonomy was found to be a significant moderator,  $F(1, 27) = 6.23, p < .05$ , and Table 21 indicates that this effect was in the hypothesized direction.

Table 21. Model for variables moderating the effect of *visualization* (VZ) on navigation ( $k = 18$  studies)

<b>Variable</b>	<b>Coefficient (<math>\beta</math>)</b>	<b>SE</b>	<b>t-value</b>
Reliability	0.03	0.46	0.06
Navigation Autonomy	-0.11	0.04	2.50*

Note: \*  $p < .05$

H11 stated that the association between *visualization (VZ)* and performance would be moderated by the testbed that was used to test this relationship. For outcomes with measures of localization, the testbed was found to be a significant moderator  $F(1, 63) = 6.18, p < .05$ . As illustrated in Table 20, associations were found to be stronger in physical environments than virtual environments. No additional support was found for the other three outcomes (i.e., identification, navigation, and mission completion time).

## CHAPTER 5: DISCUSSION

This meta-analysis illustrated a variety of meaningful relationships, and Table 22 provides a summary of results for the hypotheses that were tested in the previous section. For the main effects of visuo-spatial ability on performance, H1 (i.e., negative association with mission completion time), H2 (i.e., positive association with identification), H3 (i.e., positive association with localization), and H4 (i.e., positive association with navigation) were all supported as hypothesized. When visuo-spatial ability was distinguished according to construct, effects were generally found to be stronger for Carroll's (1993) *visualization (VZ)* factor (i.e., H5 was supported), and although one effect was significant in opposite direction (i.e., association with *VZ* was weaker for localization outcome), *visualization (VZ)* consistently emerged as a unique construct. For moderation by test (i.e. H6), associations with identification were found to be stronger for the spatial orientation test by Gugerty and Brooks (2004), but no other effects were found for any other metrics or outcomes (i.e., weak support). Findings for H7 (i.e., moderation by task type for mission completion time), H10 (i.e., moderation by navigation support for navigation performance), and H11 (i.e., moderation by testbed for localization performance) indicated that all of the hypotheses were supported as originally stated. While moderation by localization support for localization performance was significant (i.e., associations strengthened with higher levels of support), the effect was opposite of what was expected in H9 (i.e., associations were expected to weaken with increased support). No significant relationship was found for H8. The following subsections discuss the implications of these findings and supplemental analyses in further detail.

Table 22. Summary of results per hypothesis

Hypothesis	Result
<i>H1: Visuo-spatial ability will be negatively associated with mission completion time</i>	Hypothesis supported.
<i>H2: Visuo-spatial ability will be positively associated with identification</i>	Hypothesis supported.
<i>H3: Visuo-spatial ability will be positively associated with localization</i>	Hypothesis supported.
<i>H4: Visuo-spatial ability will be positively associated with navigation</i>	Hypothesis supported.
<i>H5: The effect of visuo-spatial ability on performance will be moderated by the visuo-spatial construct, such that measures of visualization (VZ) form the strongest association with performance outcomes</i>	Hypothesis supported for identification, navigation, and mission completion time. For localization, differences were significant, but contrary to H5.
<i>H6: The effect of visualization (VZ) on an outcome will be moderated by the measure that is used to assess the VZ construct</i>	Hypothesis supported for the Gugerty & Brooks test of Spatial Orientation with identification outcomes. Across all other metrics and outcomes, spatial orientation tests were indistinguishable from VZ.
<i>H7: The effect of visualization (VZ) on a mission time will be moderated by mission task, such that the association is strongest with tasks that only require navigation and weakest with tasks that require an operator learn from an environment</i>	Hypothesis supported.
<i>H8: The effect of visualization (VZ) on identification will be moderated by autonomy that is designed to support identification, such that the association will be strongest under the lowest levels of autonomy</i>	Hypothesis not supported.
<i>H9: The effect of visualization (VZ) on localization will be moderated by autonomy that is designed to support localization, such that the association will be strongest under the lowest levels of autonomy</i>	Hypothesis not supported. The effect was significant, but opposite to what was hypothesized.
<i>H10: The effect of visualization (VZ) on navigation will be moderated by autonomy that is designed to support navigation, such that the association will be strongest under the lowest levels of autonomy</i>	Hypothesis supported.
<i>H11: The effect of visualization (VZ) on an outcome will be moderated by the testbed that is used to assess the relationship</i>	Hypothesis supported for outcomes involving localization. For outcomes involving identification, navigation, and mission completion time, the testbed did not change the strength of the associations.



## Theoretical Implications

Findings from this analysis indicate that visuo-spatial abilities are important for predicting performance in remote environments. Support for H1 through H4 indicates that the associations between ability and performance are consistently significant across a variety of outcomes. Given the strength of these associations, it was not surprising to see that these metrics are commonly used as predictors of performance in applied settings (Hegarty & Waller, 2005). As the following subsections discuss, these findings do not appear to be absolute across all circumstances, and researchers will need to be aware of relevant differences.

### *Moderation by Visuo-Spatial Construct*

The first important finding involves moderation by the visuo-spatial construct (i.e., support for H5). As discussed in the review of literature, there are multiple perspectives regarding visuo-spatial factors. This diversity creates confusion with respect to understanding “spatial ability,” and in the domain of UV operation, this confusion was reflected in the variety of metrics that are discussed as measures of “spatial ability.” In spite of this variety, the majority of researchers used measures of Carroll’s (1993) *visualization (VZ)* construct, and the use of this predictor typically elicits the strongest associations with performance. Therefore, regardless of how researchers have labeled the construct, associations between “spatial ability” and performance typically involve metrics that load onto Carroll’s (1993) *visualization (VZ)* factor, and Carroll’s model emerged as one of the better guides for examining relationships with visuo-spatial factors.

With respect to construct validity and the *visualization (VZ)* factor, findings from the meta-analysis illustrated a problem with using multiple predictors. Specifically, when all of these metrics load onto the *visualization (VZ)* factor, there are no meaningful differences between using one metric or another. Otherwise stated, two or more measures of same construct essentially

illustrate the same relationship, and if the goal of a researcher is to illustrate meaningfully different relationships, this may not be the most effective methodology. Instead, research with multiple metrics would most likely benefit from assessing visuo-spatial constructs that are more distinguishable.

Another issue with using multiple predictors involves aggregating metrics across factors (see Lathan & Tracey, 2002). Specifically, this method confounds constructs, and as indicated via support for H5, aggregation of metrics across constructs typically reduced the relationship between ability and performance. Instead of aggregation, research with multiple constructs might be best served by incorporating multivariate methods (see Fincannon et al., 2012; Gugerty, Brooks, & Treadaway, 2004). As this approach considers the unique variance that is associated with each construct, it should be better suited to assess relevant differences.

With respect to specific visuo-spatial constructs, problems with the spatial orientation factor were also observed. Specifically, the majority of metrics that are designed to assess spatial orientation did not appear to illustrate relationships that were distinguishable from Carroll's (1993) *visualization (VZ)* factor. This was especially true for traditional metrics, such as the Guilford-Zimmerman Test of Spatial Orientation (Guilford, 1975, 1985) and the Cube Comparison Test (Ekstrom et al., 1976). As a result of this inability to draw a distinction, it was not surprising to see researchers aggregate these metrics (Wong, 2009). Given that extensive re-analyses of literature have challenged the distinction between the visualization and spatial orientation factors (Carroll, 1993), researchers need to be aware of this issue.

While separating *visualization (VZ)* from *spatial orientation* is difficult, research in the domain of psychometrics has continued (Hegarty & Waller, 2004). Of these metrics, limited support for the unique contribution of a spatial orientation test by Gugerty and Brooks (2004) did emerge. While one might expect spatial orientation to be associated with measures of

localization, this was not found for the Gugerty and Brooks test. Instead, the Gugerty and Brooks test was found to elicit the strongest association with target identification. The non-intuitive relationship appears to be problematic for the Gugerty and Brooks test.

Considering past issues that have emerged with the study of visuo-spatial ability (Carroll, 1993), it might be worth considering whether the spatial orientation task by Gugerty and Brooks (2004) is a measure of one cognitive ability or whether it samples from multiple cognitive constructs. Specifically, Chen and colleagues (Chen & Terrence, 2009) have used this metric as a predictor of performance in UV operation, whereas Gugerty and Colleagues (Gugerty & Brooks, 2004) have used this metric as an outcome to discuss navigational training and interface design. In one such study, Gugerty and colleagues (Gugerty et al., 2004) performed a regression analysis, which found that measures of crystallized intelligence, fluid intelligence, technical knowledge, and mental rotation were all significant predictors that accounted for unique variance in the spatial orientation metric. If multiple cognitive abilities are uniquely associated with this metric as an outcome, it raises questions about the psychometric purity of the test (e.g., is this a pure measure of spatial orientation? Is it a metric of *visualization [VZ]* that is confounded by other cognitive abilities?). Typically, factor analysis has been used to assess this question, but the factor analytic research that has been used to discuss recent distinctions between mental rotation and other tests of spatial orientation (Hegarty & Waller, 2005) appears to be absent from research by Gugerty and Brooks (2004). If this metric of orientation was designed as an outcome that is as multidimensional as applied RSTA tasks are, the metric is likely to be confounded by assessing other constructs, which is problematic for this meta-analysis.

In the study of how individual differences form associations with performance, analyses for H5 highlighted an important issue that is not commonly considered. Specifically, effects sizes for *SR*, *CF*, and *MV* appeared to be stronger than what was observed for traditional metrics of

*visualization (VZ)*. While the number of studies with these constructs may be limited, research by my colleagues and me (Fincannon et al., 2010a, 2012) have argued for the consideration of unique contributions by different visuo-spatial constructs. The results reported in Table 15 are uncommon in that they highlight the strength of an association with these constructs over the metrics of *visualization (VZ)* and spatial orientation that were considered in this study. With respect to *closure flexibility (CF)*, it should be noted that McGee's (1979) discussion of spatial orientation used measures that load onto Carroll's (1993) *closure flexibility (CF)* factor (i.e., embedded figures test), and as this may have confounded McGee's discussion of spatial orientation, there may be unexplored potential for examining the relationships between *closure flexibility (CF)* and localization. Further research will be needed to examine these relationships.

In summation, there are many contradictions regarding models visuo-spatial ability (Carroll, 1993; Ekstrom et al., 1976; Guilford 1985; Lohman et al., 1987; McGee, 1979), and this makes it important for researchers to understand these differences and how they apply to research. Findings from this meta-analysis were most consistent with Carroll's (1993) discussion of visuo-spatial ability, such that the majority of research used measures of Carroll's *visualization (VZ)* factor. In spite of the intuitive argument for spatial orientation, traditional measures of this construct were not uniquely distinguishable from *visualization (VZ)*. There may be support for unique contributions of new spatial orientation tests and other visuo-spatial abilities, but further research will be needed to fully understand these relationships.

#### *Moderation by Autonomy*

In the domain of UV operation, systems should be designed such that they can be used effectively by a wide variety of operators. Since *visualization (VZ)* was typically associated with better operator performance, an effective increase to the level of autonomy should decrease the strength of this association. The intentions of H8 through H10 were to test this relationship. Of

the levels of autonomy that were considered, navigation support was most effective at decreasing the associations between *visualization (VZ)* and navigational outcomes (see Table 21). While this appears to be good (i.e., autonomy moderates impact of ability, such that *higher* levels of autonomy result in *higher* levels of performance for all operators), this meta-analysis only considered the strength of associations, and assumed that higher levels of autonomy consistently improved performance. As some of the analyses from this meta-analysis illustrate this relationship (Fincannon et al., 2012), there appears to be some support for ultimately using autonomy to decrease the need for the selection of operators in UV operation.

The counter-argument to decreasing the association between *visualization (VZ)* and performance involves a moderator's potential to decrease performance. An obvious example of this might involve a fully autonomous system (i.e., the operator is unable to intervene) that is unreliable in its execution of a task. Another issue involves high levels of autonomy that reduce *human-in-the-loop* processing, which can be exemplified through the negative impact of auto-piloting systems on pilot performance and situation awareness (Endsley & Kiris, 1995; Parasuraman, Molloy, & Indramani, 1993). In the context of Table 20, navigation support was found to decrease the association between *visualization (VZ)* and localization performance. However, this analysis did not provide any indication as to whether this relationship is desirable. If navigation support is associated with a decrease in an operator's ability to localize objects in a remote environment, this meta-analysis might have highlighted a troubling relationship. Future research will be needed to ensure that navigation support does not actually hinder an operator's ability to localize objects.

In spite of the success in finding that higher levels of autonomy decreased the association between *visualization (VZ)* and performance, the exact opposite was also found to be true for other levels of autonomy. Contrary to expectations, *higher* levels of autonomy that supported

identification and localization were found to *increase* the association between *visualization (VZ)* and localization performance. One possible explanation for this may lie in understanding the differences between types of autonomy that were considered in this meta-analysis. In the context of the model that was presented by Parasuraman and colleagues (2000), autonomy for identification and localization are best categorized as levels of automation for acquiring information, whereas autonomy for navigation goes further to support analysis and decision making. For localization tasks, autonomy for acquisition simply provided information, but since the autonomy did not interpret anything, *visualization (VZ)* supported an operator's ability to further process this information. In contrast, navigation autonomy processed information, and by removing the need for an operator to attend to localizing or navigating within the environment, the association between *visualization (VZ)* and these dimensions of performance were diminished. If future technology is intended to increase performance for low-ability operators, findings from this analysis suggest that this technology will need to support information processing across a variety of performance outcomes, and until the technology is capable of accomplishing this task, visuo-spatial ability will continue to be an important tool for selection.

In summary, moderation by autonomy (i.e., H8 through H10) could reduce the association between *visualization (VZ)* and performance (i.e., *higher* levels of navigation support *minimized* the relationship with navigation performance), but this was not the case in all circumstances (i.e., *higher* levels of identification and localization support *maximized* the relationship with localization performance). Previous research has argued that automation can have unintended consequences (Parasuraman & Riley, 1997), and findings from this meta-analysis support this assertion. In the case of UV design, it is desirable to create a system that can be used by the most diverse population, but some of these design interventions simply increase

the need for selecting high-ability operators. As roboticists intend to further develop UV autonomy, individual differences will require their attention.

#### *Moderation by Task*

The analyses in this study indicated that unique associations with *visualization (VZ)* were very dependent upon the outcome that was assessed in the study. This was not only evident in the mixed support for H5, but it was also present in the analysis of variables that moderate the association between *visualization (VZ)* and performance (see Tables 18 through 21). From a theoretical perspective, these findings illustrate how differences across outcomes change the generalizability of specific relationships (Fincannon et al., in press-a), and researchers need to attend to relevant issues.

Support for H7 (i.e., moderation by task type) points to a couple of important issues with mission completion time. First, combined support for H1, H2, H3, H4, and H7 strengthens the assertion that the impact of *visualization (VZ)* on a mission completion time is mediated by multiple dimensions of visual performance. Second, since mediators appeared to vary from task to task, faster performance may not always be associated with high-ability operators. Otherwise stated, successful completion of some visual tasks appeared to require more mission time, and high-ability operators were more adept at spending the *appropriate* amount of time on a task. As a result of this finding, the interpretability of mission completion time does not appear to be consistent, and researchers need to understand that faster is sometimes, but not always, better.

In conclusion, it is important for researchers to understand the task that operators are required to perform. Specifically, relationships between *visualization (VZ)* and performance changed across different dimensions of performance. Furthermore, the interpretability of specific dimensions of performance (i.e., mission completion time) appeared to change across different operator tasks. Therefore, the task is arguably as important as the visuo-spatial construct.

### *Moderation by Environment*

Recent research has highlighted the importance of considering external validity in UV operation (Fincannon et al., in press-a). In the context of this argument, H11 was presented to explore differences that were associated with experimentation in virtual and physical environments. Results indicated that associations between *visualization (VZ)* and localization (see Table 20) were stronger in physical environments. Since end-user applications involve physical systems, these findings indicate that estimates of a relationship with a virtual environment are likely to underestimate the true importance of *visualization (VZ)* for predicting performance in the field.

Another interesting finding with the external validity issue was illustrated by how the intended application influenced the strength of the associations between *visualization (VZ)* and performance. As illustrated by Table 20, associations with performance were found to be stronger for research that was intended to apply to an HRI domain, which challenges the external validity of theoretical work that is not domain-specific. Part of the reason for this may lie in unique aspects of UV operation, such as variations in the level of autonomy for H9 and H10 that were significant as moderators. While this meta-analysis attempted to control some of these factors, there are likely to be other constructs that vary across domains of application, and focusing on the intended application illustrates that these moderators are likely to be present. As a result of this, the simple act of designing research to study a specific domain will likely increase the degree to which the results of that research generalize to that domain.

In summary, findings do not necessarily generalize as one might expect. Some of this moderation was attributed to hypothesized methods of simulation (i.e., research conducted in physical vs. virtual environments). As moderation was also attributed to the intended domain of application (i.e., HRI vs. general theory), research that is designed to study UV operation is likely



to create variation in other moderating variables (i.e., beyond task, autonomy, testbed, and other factors considered in this meta-analysis) that alter the relationship between *visualization* (VZ) and performance. Therefore, research questions with visuo-spatial abilities may still need to be examined across multiple settings to maximize our theoretical understanding of external validity.

#### *Moderation by Gender*

As discussed earlier in the paper, there are a variety of psychometric confounds that make it difficult to difficult to assess visuo-spatial ability. One of these focuses on how different test takers use different strategies to complete the same test (Carroll, 1993). In the domain of gender differences, men have been found to demonstrate higher levels of spatial ability than women (Voyer et al., 1995). If low spatial ability is associated with the degree to which non-spatial strategies are used to complete spatial test, test of spatial ability should be less effective at assessing the spatial ability of women. Furthermore, if spatial test were poor at assessing the spatial ability of women, performance by women on these spatial tests should be less predictive of performance, which was confirmed in this meta-analysis (see Tables 19 & 20). Therefore, this meta-analysis has highlighted a psychometric flaw in the assessment of *visualization* (VZ).

Moderation by gender primarily illustrates a problem with external validity. Specifically, research by my colleagues and me (Fincannon et al., in press-a) has noted findings that are not necessarily consistent across groups of participants. Within the context of this meta-analysis, measures of *visualization* (VZ) are less predictive of performance for women than men. As a result of this psychometric flaw, the external validity of *visualization* (VZ) metrics has been challenged and future research will need to develop metrics that are not moderated by these factors.

## Practical Implications

Across a variety of outcomes, *visualization (VZ)* emerged as one of the stronger predictors of performance. This is illustrated by support for H5 (i.e., moderation by construct) across three outcomes and the stability of associations that were reported in Table 17 for all four outcomes. In the context of practical applications, this means that the single best metric for predicting operator performance will likely load onto Carroll's (1993) *visualization (VZ)* factor.

It is important to understand that different types of autonomy did not have the same moderating impact. Specifically, higher levels of autonomy for supporting navigation *decreased* the association between *visualization (VZ)* and performance, but higher levels of autonomy for supporting identification and localization *increased* the association between *visualization (VZ)* and performance. If managers are interested in using autonomy to eliminate the need for *visualization (VZ)* as a selection tool, these relationships need to be taken into consideration.

Another implication for practice involves the importance of a measure's reliability in assessing a construct. Specifically, Tables 15 and 20 illustrated that the reliability coefficients of cognitive ability tests were positively associated with the strength of the association between that ability and localization performance. If a test is more likely to elicit a strong association, it should be easier to use this metric to select operators that perform well with remote systems. Therefore, a good test for selection should have a high reliability coefficient (e.g., 0.95 coefficient of the Guilford-Zimmerman test of Spatial Visualization).

Moderation by environment and intended domain of the application has significant implications for practice. Specifically, a lot of work might be published on a given subject, but this meta-analysis indicated that all of this research may not be relevant to a specific area of practice. As the results illustrated, an easy approach for increasing the degree to which findings generalize to a domain simply involved designing a task to mimic that domain (i.e., UV

experimentation for UV application). Therefore, challenges to external validity exist, but relatively simple interventions can be used to alleviate this issue.

### Limitations

Many of the hypotheses for this meta-analysis were supported, but this does not imply that this is a perfect analysis. The following subsections highlight issues that may limit the implications of the findings.

#### *Theoretical Scope*

As discussed in the introduction, there are many disagreements regarding the theoretical underpinnings of visuo-spatial abilities. For example, some researchers have argued that working memory is an important component of one or more visuo-spatial abilities (Kaufman, 2007; Miyake, Friedman Rettinger, Shah, & Hegarty, 2001). In contrast to this, other researchers have argued that working memory involves a higher level of functioning that is more strongly associated with general intelligence, memory span, and/or cognitive speed (Carroll, 1993; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Kane, & Engle, 2003; McGrew, 2005). This meta-analysis was designed to examine relationships regarding the general grouping of factors, but it was not designed to examine these theoretical relationships. Future research will be needed to explore these issues.

#### *Source Material*

Many of the problems that emerged throughout the course of conducting this analysis involved data extraction. Specifically, this involved incomplete reporting of significant and non-significant findings within the published material, which had the potential to be problematic for 21 of the 85 studies that were found to be relevant for this analysis. I had access to the raw data for 12 of these datasets, and personal correspondence with other authors supplied data for 5 more studies. Of the remaining 4 studies, one author simply restated non-significance with respect to

*p*-values (i.e., even though correlations were requested, the author stated that *p*-values were less than 0.05), and the remaining authors were never reached. Therefore these four studies were excluded, and of the 85 studies that were relevant to this meta-analysis, only 81 were included.

Of the excluded studies, two had a non-significant relationship with a measure of *closure flexibility (CF)*, but the exclusion of this material may not be overly problematic. Specifically, H5 (i.e., moderation by construct) stated that findings with this construct should be weaker than effects for the *visualization (VZ)* factor, and the insignificance of *closure flexibility (CF)* is consistent with this hypothesis. Furthermore, the re-analysis of main effects (Table 17) and testing of H6 through H11 only used measures of the *visualization (VZ)* construct, which would have excluded these metrics.

Another problem with the associations that were recorded from the source material involved aggregation of predictors, which was observed for 20 of 81 studies for this analysis. With respect to H5 and H6, this had the potential to eliminate data points and reduce power. With respect to H7 through H11, aggregation across constructs would confound *visualization (VZ)* with other visuo-spatial factors, making it more difficult to assess relevant relationships. Through personal correspondence with the authors of these materials, findings for the individual predictors were obtained for 17 of these studies. Data for the remaining 3 studies were excluded from the testing of H5 and H6, but were included for the testing of H7 through H11. In order to minimize the effect of confounding *visualization (VZ)* with other visuo-spatial factors, aggregation was considered as a covariate (i.e., 1 = aggregation across constructs; 0 = other), but it was not found to be significant. Therefore, the effect of aggregation was probably not strong enough to be observed in this analysis.

### *Method of Analysis*

While this meta-analysis did identify a number of interesting associations with *visualization (VZ)*, there are still limitations. Specifically, the meta-regression techniques that were used have been found to suffer from a lack of power (Hunter & Schmidt, 2004), and this can be exacerbated by minimal usage of tests (e.g., low occurrence of object perspective test). As a result of these problems, this analysis may have only identified the strongest moderators.

A second problem with this analysis is that it was designed to assess the relative strength of one construct over another. As a result, two unique constructs (e.g., closure flexibility and visual memory) may elicit the same effect size (e.g.,  $r = .45$ ), but since there is no difference in the relative strength of the effect, the technique that was used in this meta-analysis would not detect a difference. Future research may benefit from considering alternative techniques (e.g., regression with a meta-correlation matrix) to assess unique variance of visuo-spatial constructs.

### *Additional Concerns*

As mentioned earlier in the analysis (see Tables 12 & 17), there was no variance around the corrected correlation coefficients for outcomes with identification and localization, which points to an issue with second order sampling error. While the number of participants across studies was sufficient to provide a stable assessment of the mean correlation (see Appendix F), variation around this aggregate correlation is attributed the number of unique studies. In the presence of second-order sampling error, one can under estimate (i.e., what was observed in Tables 12 & 17) or overestimate variance. This challenges the validity of the confidence intervals that have been assessed in this analysis. As additional data becomes available, it may be necessary to reevaluate this statistic.

As an alternative explanation, the lack of variance might also be attributed the method of obtaining correlation coefficients in Table 12 and 17. Specifically, the majority of study-level

effects were obtained by aggregating the specific coefficients that were obtained from each study (see Table 11), and when examining a distribution of aggregates, the central limit theorem dictates that increasing the number of observation per aggregate will narrow the distribution of these aggregates (Mendenhall & Sincich, 2007). In the context of this rationale, Table 11 indicated that identification and localization outcomes had the higher number of effects per study. As such, the issues that were observed with the distributions might be expected.

### Future Research

This meta-analysis illustrated the importance of defining constructs, and this need concept needs to be incorporated into future research. In the context of this issue, “spatial ability” and “performance” are generic terms that can create confusion, and researchers must use specific terms to describe their constructs. When researchers define their constructs, the model that is used to guide this process is incredibly important, but since these models typically disagree on various aspects of visuo-spatial ability, it is important to find a model that helps a researcher identify meaningfully different relationships. The results of this meta-analysis indicated that Carroll’s (1993) description of visuo-spatial abilities is the best global model to guide future research.

In highlighting weak dissociations with spatial orientation and moderation by gender, there is clearly a need for further psychometric development. For example, a recent meta-analysis aggregated brain imaging metrics for measures of mental rotation (Zacks, 2008). If activation within a certain area of the brain is found to be associated with specific ability, this technique may be better at determining mental processes of that ability than a written test. Future research should consider the use of these applications to provide a better assessment of cognitive abilities.

In terms of practical applications, this meta-analysis is probably most effective in highlighting the importance of *visualization (VZ)* in predicting operator performance with unmanned systems. Analysis with localization outcomes did show potential application for other constructs (Table 15), but since the “comparison construct” spanned multiple factors (i.e., with a small number of studies per factor), it is difficult to use the current analysis to provide strong recommendations for any *non-visualization (VZ)* construct. There appears to be a benefit to assessing additional constructs (e.g., *spatial orientation, closure flexibility, visual memory*), but further research will be needed to strengthen our understanding of these constructs and their association with various outcomes of interest in UV operation.

Moderation environment (i.e., physical vs. virtual) and the intended domain of application (i.e., HRI research to study UV operation vs. other) raised more questions than it answered. In a theoretical context, these variables can be vague (e.g., HRI research vs. other) or contain a great deal of variation within categories (e.g., the design of virtual environment), and as described by my colleagues and me (Fincannon et al., in press-a), a generalized theory of causal inference will require further research to identify the underlying mediator(s) behind these relationships. For example, Huthmann (2009) found that the influence of environment type (i.e., physical vs. virtual) on UV operator performance was mediated by presence. Therefore, presence could be an important factor. If differences are attributed to physical environments providing more information (e.g., trees, plants, clouds, depth cues, etc.), one or more manipulations to make a virtual environment more realistic might ameliorate the effects that were observed in this analysis. Alternatively, recent research by my colleagues and me (Fincannon et al., in press-b) found that familiarity and obstruction of a stimulus could change the relationship between *visualization (VZ)* and identification, and these differences could simply be related to the type of

stimulus and how it is placed in an environment. Many of these questions are beyond the scope of this meta-analysis, and future research will need to explore these relationships.

This analysis was successful for illustrating important relationships regarding *visualization (VZ)* and variables that moderate its relationship with various dimensions of performance (i.e., identification, localization, navigation, and mission completion time), but the analysis can be taken further. The moderator analysis in this paper focused on two-way interactions, but higher order interactions need attention. As an example of this, recent research by my colleagues and me (Fincannon et al., in press-b) has illustrated how familiarity, obstruction, and *visualization (VZ)* interact to predict identification, and similar questions and techniques can be applied to variables that are presented in this meta-analysis. Future research needs to continue examining these relationships.

In addition to moderation, this meta-analysis also illustrated a need to focus on mediation. The theoretical foundation for H7 involved untested assumptions about variables that mediate the relationship between *visualization (VZ)* and mission completion time. Since findings support this hypothesis, mediation is likely to be present, but future research needs to take this finding further by illustrating relevant relationships. Furthermore, it was important to illustrate that the environment (i.e., physical vs. virtual) and intended domain of application (i.e., HRI vs. other) change relationships with *visualization (VZ)* and performance, but as moderators, these variables are extremely vague. For example, a virtual environment can be changed to make it more, or less, like the real world. If these changes alter the relationship between *visualization (VZ)* and an outcome of interest, these design elements are the true variables of interest. Future research will not only need to assess basic mediation (e.g., a dimension of visual performance mediating the relationship between ability and time), but it will also need to use mediation to better understand moderation.



Finally, future research will need to consider other outcomes of interest. Specifically, other dimensions of performance (e.g., manipulation of a robotic arm) were excluded from this analysis, and these outcomes are likely to form unique associations with one or more abilities (e.g., some combination of visual and non-visual abilities). Furthermore, this meta-analysis also excluded non-performance outcomes. As recent research has shown that *perceptual speed (PS)*, as opposed to *visualization (VZ)*, formed a strong association with the NASA TLX (Fincannon et al., 2012), workload may serve as useful construct for further analysis. Therefore, this meta-analysis has provided an initial examination of specific performance outcomes, but future research will be needed to explore relationships with other dimensions of performance and non-performance outcomes.

### Concluding Remarks

In conclusion, this meta-analysis illustrated several points, which can be summarized as follows:

- Of the models discussing visuo-spatial ability, findings were most consistent Carroll's (1993) presentation of constructs
- The majority of the research with visuo-spatial ability used metrics that load onto Carroll's (1993) *visualization (VZ)* factor
- When studies use metric that load onto Carroll's (1993) *visualization (VZ)* factor, associations with performance are typically stronger
- Since findings are not consistent across all outcomes, unique hypotheses must be developed for each dimension of UV performance
- Variations across tasks can change the interpretability of mission completion time

- As moderating variables, the levels of autonomy (i.e., support for identification, localization, or navigation) had mixed effects, such that higher levels of autonomy would both increase and decrease the association between *visualization (VZ)* and performance
- With respect to external validity, differences were attributed to the testing environment (i.e., physical vs. virtual), intended application (i.e., HRI vs. general research), and gender

As a result of these findings, *visualization (VZ)* is clearly an important predictor of performance, and metrics that assess this construct should be useful for the selection of high-performing operators in UV tasks. Based on the moderator analysis, future research should consider the unique contributions of different visuo-spatial factors, different types of UV technology, and methods that might improve, or illustrate further problems with, external validity.

## **APPENDIX A: LIST OF SEARCH TERMS**

- Spatial ability
- Spatial visualization
- Spatial orientation
- Spatial relations
- Mental rotation
- Closure speed
- Closure flexibility
- Gestalt perception
- Closure flexibility
- Perceptual speed
- Visual memory
- Robot
- Unmanned vehicle
- Unmanned ground vehicle
- Unmanned aerial vehicle
- Reconnaissance
- Search and rescue

**APPENDIX B: LETTER TO REQUEST ACCESS TO UNPUBLISHED  
DATA**

To Whom This May Concern:

My name is Thomas Fincannon, and I am writing this letter in regard to a meta-analysis involving visuo-spatial constructs and remote perception. Specifically, I am looking to examine the impact of cognitive abilities that include:

- Spatial visualization
- Spatial orientation
- Spatial relations
- Closure speed
- Closure flexibility
- Perceptual speed
- Visual memory

If you have conducted a study using one or more of these constructs in the domain of unmanned vehicle operation (i.e., published or unpublished), please respond to this message, and I will send out a form requesting more information.

Respectfully,

Thomas Fincannon

**APPENDIX C: SAMPLE LETTER TO REQUEST CLARIFICATION  
OF PUBLISHED DATA**

Author:

My name is Thomas Fincannon, and I am writing this letter to request information about a paper entitled *[paper title]* that was published in *[source]* on *[year]*. Specifically, I am conducting a meta-analysis involving visuo-spatial constructs and would like to request more detail about the effects within your dataset.

In your paper, you reported *[list relevant relationships]*. I would like to request information about *[describe unanswered question]*. Specifically, could you please provide me with effects:

- *Statistic for effect A*
- *Statistic for effect B*
- *Statistic for effect C*

Any information that you can provide would be of value, and your time is greatly appreciated. If you have questions or problems with providing the requested information, please let me know.

Respectfully,

Thomas Fincannon



**APPENDIX D: DESCRIPTIONS OF CONTRACTS FOR  
SUBJECTIVE RATINGS**

## **PERFORMANCE CLASSIFICATION**

1. Is the measure reported in units that require one to correctly indentify or classify an object in the environment? Specific examples include target identification, targets classification, and reconnaissance.
  - If the measure is primarily associated with identification/classification, is performance associated with recognizing that a specific target is located at a specific objective or point on a map?
  - Is one required to perform the assessment from a camera or virtual image that moves through a remote environment?
2. Is the measure reported in units that require one to know where something is located in the remote environment? Specific examples include determining location on a map or orientation/distance of an object in relation to its vehicle or point of reference.
  - If the measure is primarily associated the location of something in a remote environment, is performance primarily associated determining the location of an egocentric vehicle or point of reference?
  - If the measure is primarily associated the location of something in a remote environment, is performance primarily associated determining the location of an object that is exocentric to the vehicle or point of reference?
  - Is one required to perform the assessment from a camera or virtual image that moves through a remote environment?
3. Is the measure reported in units that require an operator to effectively navigate through a remote environment? Specific examples include obstacle avoidance, route selection, and proficiency in rerouting a vehicle through an environment.

4. Is the measure reported in units of time? A specific example includes the amount of time that is taken to complete a mission.
- If the measure is reported in units of time, does the primary task require an operator and/or autonomy to navigate a vehicle through a remote environment?
  - If the measure is reported in units of time, does the primary task require an operator to identify and/or localize objects in the remote environment (e.g. reconnaissance tasks)?
  - If the measure is reported in units of time, is it primarily a measure of reaction time (e.g., time taken report an outcome)?

## AUTONOMY CLASSIFICATION

**Identification Support:** Does the autonomy support the operator with the identification of objects in the environment and to what degree is this support provided?

- **Operator Identification [Rated as 0]:** The participant is not provided with any autonomous support for the identification of targets and objects.
- **Detection Support Identification [Rated as 1]:** The participant is informed of the presence of targets and objects to identify. This can be accomplished by providing an auditory queue or some other method of alerting the participant to presence of targets in the remote environment.
- **Fully Autonomous Identification [Rated as 2]:** The autonomy identifies targets and objects that are encountered by the participant.

**Localization Support:** Does the autonomy support the operator with the identification of objects in the environment and to what degree is this support provided?

- **Operator Localization [Rated as 0]:** The participant localizes targets and vehicles in the environment without autonomous support.
- **Autonomous Localization [Rated as 1]:** The autonomy provides a participant with information about location (e.g., coordinates) in a remote environment This can be accomplished through the provision of coordinates (e.g., via text) or display of a vehicle on a user interface (e.g., location of a vehicle on a virtual map). If one vehicle is capable of providing an exocentric view of another vehicle (e.g., viewing a UGV from the perspective of a UAV), this *is not* localization support.

**Navigation Support:** Does the autonomy support the operator with the identification of objects in the environment and to what degree is this support provided?

- Teleoperation Operator [Rated as 0]: The participant must manually maneuver a UV or avatar through a remote environment. Under this condition, the *operator has complete control*.
- Semi-Autonomous Navigation [Rated as 1]: A participant uses waypoints to plan a route through a remote environment that is executed by a UV or avatar. The autonomy *executes plans* that are created by the operator, but it *does not provide choices* to help with planning the task.
- Rerouting Support for UV Navigation [Rated as 2]: Navigation is largely autonomous, such that the autonomy will reroute, or provide suggestions to an operator who reroutes, a vehicle as it moves through a remote environment. In this condition, the *operator has final control* over how the vehicle navigates, but the *autonomy supports* this process by via *planning and execution of routes*.
- Fully Autonomous Navigation [Rated as 3]: The autonomy navigates through an environment without feedback from the participant. In this condition, the *operator has no control*. NOTE: If a participant is simply observing a recorded video, this would be coded fully autonomous navigation.

## **APPENDIX E: LIST OF FORMULAS**

## **STATISTICS FOR CODER RELIABILITY**

Cohen's Kappa:

$$\kappa = \frac{o - e}{1 - e}$$

Where  $K = \text{Kappa}$ ;  $o = \text{observed percentage of agreements}$ ; and  $e = \text{the expected percentage of agreements}$ .

## **STATISTICS FOR ESTIMATING A CORRELATION FROM REPORT DATA**

$$r = \frac{t}{\sqrt{t^2 + N - 2}}$$

with:  $r = \text{the correlation coefficient}$ ;  $t = \text{observed value for the } t\text{-statistic}$ ; and  $N = \text{the observed sample size}$

$$r = \frac{d}{\sqrt{d^2 + 4}}$$

with:  $r = \text{the correlation coefficient}$ ; and  $d = \text{the observed value for Cohen's } d$

$$r = \sqrt{F / (F + N - 2)}$$

with:  $r$  = the correlation coefficient;  $F$  = observed value for the  $F$ -statistic; and  $N$  = the observed sample size

$$r = \frac{m_E - m_C}{SD_p} (\sqrt{pq})$$

with:  $r$  = the correlation coefficient;  $m_E$  = the observed mean;  $m_C$  = the comparison mean;  $SD_p$  = the pools standard deviation,  $p$  = the percent of sample in  $m_E$ ; and  $q$  = the percent of sample in  $m_C$  ( $q = 1 - p$ )

### **STATISTICS FOR FISHER'S Z AND WEIGHT FOR REGRESSION**

Fisher's Z:

$$z = \ln \left( \frac{1+r}{1-r} \right)$$

with:  $z$  = Fisher's Z; and  $r$  = the correlation coefficient



Sampling variance for Fisher's Z:

$$v = \frac{1}{N_i - 3}$$

with:  $N_i$  is the sample size of study  $i$

Weight for Fisher's Z:

$$w = \frac{1}{v}$$

with:  $w$  = the weight for each estimate of Fisher's Z; and  $v$  = the sampling variance of Fisher's Z

## **STATISTICS FOR ESTIMATING EFFECTS FROM UNCORRECTED CORRELATION**

### **COEFFICIENTS**

Estimate of the mean correlation:

$$\bar{\rho}_0 = \frac{\sum_{i=1}^k N_i r_i}{\sum_{i=1}^k N_i}$$

With:  $r_i$  = the observed correlation of study  $i$ ;  $N_i$  = the sample size of study  $i$ ; and  $k$  = the total number of studies in the meta-analysis

Estimate of variance around the observed effect size:

$$\text{Var}(r) = \frac{\sum N_i (r_i - \bar{\rho}_0)^2}{\sum N_i}$$

With:  $\bar{\rho}_0$  = mean correlation;  $r_i$  = the observed correlation of study  $i$ ; and  $N_i$  is the sample size of study  $i$

Estimating the sampling variance of Study  $i$  [ $\text{Var}(e_i)$ ]:

$$\text{Var}(e_i) = \frac{(1 - \bar{\rho}_0^2)^2}{N_i - 1}$$

With:  $\bar{\rho}_0$  = squared mean correlation; and  $N_i$  is the sample size of study  $i$

Estimating the average sampling variance across all studies [Var(e)]

$$\text{Var}(e) = \frac{\sum N_i \text{Var}(e_i)}{\sum N_i}$$

With:  $\text{Var}(e_i)$  = estimated sampling variance of study  $i$ ; and  $N_i$  is the sample size of study  $i$

Estimate of the variance  $\text{Var}(\bar{\rho}_0)$  from  $\text{Var}(r)$  and  $\text{Var}(e)$ :

$$\text{Var}(\bar{\rho}_0) = \text{Var}(r) - \text{Var}(e)$$

With:  $\text{Var}(e_i)$  = estimated sampling variance of study  $i$ ; and  $\text{Var}(r)$  = the observed variance around the effect size

## **STATISTICS FOR ESTIMATING EFFECTS FROM CORRECTED CORRELATIONS**

Attenuation Factor for Measurement Error:

$$a = \sqrt{R_{xx}}$$

with:  $a$  = the attenuation factor; and  $R_{xx}$  = the observed reliability coefficient

Attenuation Factor for Dichotomization:

$$a = \frac{\phi(c)}{\sqrt{pq}}$$

with:  $p$  &  $q$  represent the proportion of observations in each condition ( $q = 1 - p$ ); and

$$\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

Corrected Effect Size:

$$rc_i = r_i / A_i$$

With:  $A_i$  = the attenuation factor for study  $i$ ; and  $r_i$  is the observed correlation of study  $i$

Weight for Effect Size:

$$w_i = N_i * A_i$$

With:  $A_i$  = the attenuation factor for study  $i$ ; and  $N_i$  is the sample size of study  $i$

Mean Estimate of Effect Size:

$$\rho = \frac{\sum w_i rc_i}{w_i}$$

With:  $w_i$  = weight for study  $i$ ; and  $rc_i$  = correlation coefficient for study  $i$

Estimate of variance around the observed effect size for corrected correlations:

$$\text{Var}(rc) = \frac{\sum w_i (rc_i - \rho)^2}{\sum w_i}$$

With:  $\rho$  = mean of the corrected correlations;  $rc_i$  = the corrected correlation of the observed correlation for study  $i$ ; and  $w_i$  is the weight of study  $i$

Estimating the corrected sampling variance of Study I [ $\text{Var}(ec_i)$ ]:

$$\text{Var}(ec_i) = \frac{\text{Var}(e_i)}{A_i^2}$$

With:  $\text{Var}(e_i)$  = estimated sampling variance of study  $i$ ; and  $A_i$  = the attenuation factor for study  $i$

Estimating the average sampling variance for corrected correlations across all studies [ $\text{Var}(ec)$ ]:

$$\text{Var}(ec) = \frac{\sum w_i \text{Var}(ec_i)}{\sum w_i}$$

With:  $\text{Var}(ec_i)$  = estimated sampling variance of study  $i$  for the corrected correlation; and  $w_i$  = weight for study  $i$

Estimate of the true variance  $\text{Var}(\rho)$  from  $\text{Var}(r)$  and  $\text{Var}(e)$ :

$$\text{Var}(\rho) = \text{Var}(rc) - \text{Var}(ec)$$

With:  $\text{Var}(ec_i)$  = estimated sampling variance of study  $i$  for the corrected correlation coefficients; and  $\text{Var}(rc)$  = the observed variance around the effect size for the corrected correlation coefficients

### **STATISTICS FOR ESTIMATING THE CONFIDENCE INTERVAL**

Estimating Standard Error:  $SE_{\rho} = \frac{\sqrt{\text{Var}(\rho)}}{k}$

With:  $\text{Var}(\rho)$  = estimate of variance that is obtained by subtracting the sampling variance from the observed variance; and  $k$  = the number of unique datasets

Confidence Interval

$$\bar{\rho}_{\min} = \rho - 1.96 * SE_{\rho}$$

$$\bar{\rho}_{\max} = \rho + 1.96 * SE_{\rho}$$

With:  $p$  = the mean estimate of the effect size; and  $SE_p$  = the estimate of the standard error for  $p$

### **STATISTICS FOR ESTIMATING BIAS WITHIN THE META-ANALYSIS**

File Drawer Equation  $x = \frac{z_c^2}{2.706} - k$

With:  $x$  = the number of zero-effect size that could “nullify” the result of meta-analysis with uncorrected correlations;  $z_i$  =  $z$  transformation for the  $p$ -value of study  $i$ ;  $k$  is the observed number of studies; and

$$z_c = \sum_{i=1}^k z_i$$

### **STATISTICS FOR DISCUSSING MODERATION**

Q-statistic:

$$Q = \sum_{i=1}^k w_i (r_i - \bar{\rho}_0)^2$$

With:  $\bar{\rho}_0$  = mean correlation;  $r_i$  = the observed correlation of study  $i$ ; and  $w_i$  is the inverse of the sampling variance ( $Var(e_i)$ )

Credibility Interval:

$$CrI = \rho - 1.28 * SD_{\rho}$$

$$CrI = \bar{\rho} + 1.28 * SD_{\rho}$$

With:  $p$  = the mean estimate of the effect size; and  $SD_{\rho}$  = the square root of  $Var(p)$

Variance explained by the artifacts:

$$\%VA = \frac{Var(ec)}{Var(rc)}$$

With:  $Var(ec_i)$  = estimated sampling variance of study  $i$  for the corrected correlation coefficients; and  $Var(rc)$  = the observed variance around the effect size for the corrected correlation coefficients



**APPENDIX F: TABLES FOR STEM & LEAF PLOTS, NORMALITY  
ANLYSIS, AND POWER ANALYSES**

Table 23. Stem & leaf plot for the uncorrected association between identification and visuo-spatial ability

-0.0	1											
0.0	3	8	9									
0.1	2	4	8	9								
0.2	0	0	1	2	4	5	5	6	7	8	8	9
0.3	1	4										
0.4	8	9										
0.5	7											
0.6												
0.7	1											

Table 24. Stem & leaf plot for the uncorrected association between localization and visuo-spatial ability

0.0	5						
0.1	4						
0.2	5	9					
0.3	1	2	4	6	7	8	8
0.4	0	6					
0.5	2						
0.6	4						

Table 25. Stem & leaf plot for the uncorrected association between navigation and visuo-spatial ability

0.0	2	5	8				
0.1	2	3	7	8			
0.2							
0.3	2	2	4	5	7	9	
0.4	0	3	6	7			
0.5							
0.6							
0.7	4						

Table 26. Stem & leaf plot for the uncorrected association between mission time and visuo-spatial ability

0.0	0	2	6						
0.1	4								
0.2	8								
0.3	1	1	4	5	8	9			
0.4	6	6	8						
0.5									
0.6	2	9	9						
0.7									
0.8									
0.9	9								

Table 27. Stem & leaf plot for the uncorrected association between identification and visualization (VZ)

-0.0	5												
0.0	8	9											
0.1	2												
0.2	1	2	4	5	5	5	6	7	8	8	9	9	
0.3	1	1	1										
0.4	1	7	9										
0.5	7												
0.6	1	8											

Table 28. Stem & leaf plot for the uncorrected association between localization and visualization (VZ)

0.0	5												
0.1	4												
0.2	1	2	3	9									
0.3	1	4	6	7									
0.4	0	2	6										
0.5	2												
0.6	4												

Table 29. Stem & leaf plot for the uncorrected association between navigation and *visualization* (VZ)

-0.0	8						
0.0							
0.1	1	3	4	7			
0.2	1	7					
0.3	2	2	4	5	7	9	
0.4	0	3	6	7			
0.5							
0.6							
0.7	4						

Table 30. Stem & leaf plot for the uncorrected association between mission time and *visualization* (VZ)

0.0	0	3					
0.1							
0.2	0	3	8				
0.3	1	2	4	5	8	9	
0.4	6	8					
0.5	0						
0.6	2	9	9				
0.7							
0.8							
0.9	9						

Table 31. Normality analyses for Tables 12, 24, 25, 26, and 27

<b>Outcome</b>	<b>Skewness Statistic (Std. Error)</b>	<b>Kurtosis Statistic (Std. Error)</b>
Identification	1.01 (.43)	1.81 (.85)
Localization	-.17 (.58)	1.25 (1.12)
Navigation	.46 (.54)	.43 (1.04)
Mission Time	.51 (.54)	.54 (1.04)

Table 32. Normality analyses for Tables 17, 28, 29, 30, and 31

<b>Outcome</b>	<b>Skewness Statistic (Std. Error)</b>	<b>Kurtosis Statistic (Std. Error)</b>
Identification	.46 (.46)	.59 (.90)
Localization	.13 (.58)	.26 (1.12)
Navigation	.16 (.54)	1.34 (1.04)
Mission Time	.59 (.54)	.92 (1.04)

Table 33. Power analyses for Table 12

<b>Outcome</b>	<b><i>k</i></b>	<b>Power for <i>r</i> = .10</b>	<b>Power for <i>r</i> = .20</b>
Identification	29	.99	1.00
Localization	15	.98	1.00
Navigation	18	.69	1.00
Mission Time	18	.52	.98

Table 34. Power analyses for Table 17

<b>Outcome</b>	<b><i>k</i></b>	<b>Power for <i>r</i> = .10</b>	<b>Power for <i>r</i> = .20</b>
Identification	25	.93	1.00
Localization	15	.95	1.00
Navigation	18	.74	1.00
Mission Time	18	.78	1.00

**APPENDIX G: MATRIX OF CORRELATIONS FOR MODERATOR  
VARIABLES**

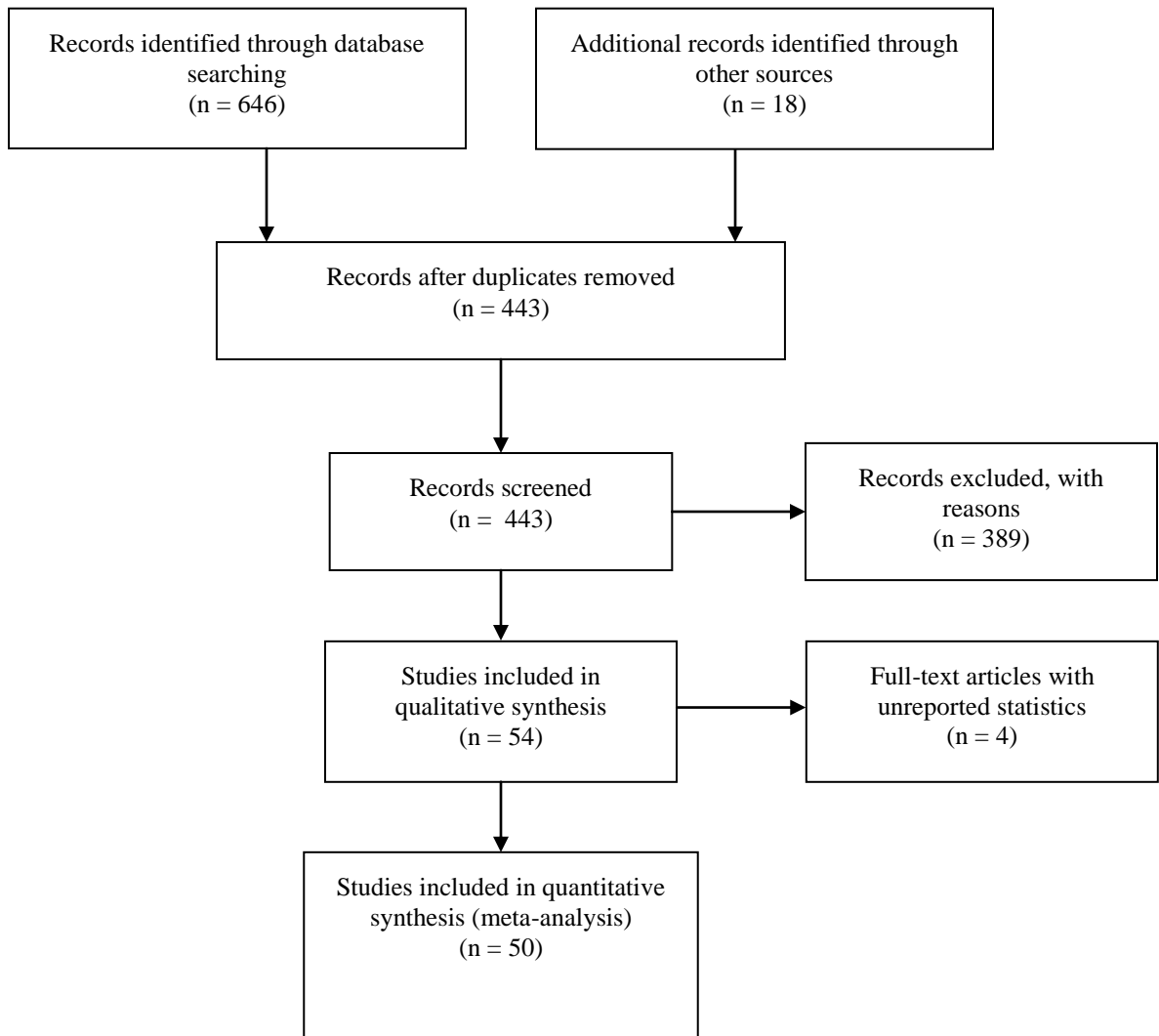
Table 35. Matrix for correlation between moderators

	1	2	3	4	5	6	7	8	9	10	11	12
1. IA	.84											
2. RIA	<b><i>-.87</i></b>	1.0										
3. LA	<b><i>.29</i></b>	<b><i>-.26</i></b>	.79									
4. RLA	<b><i>-.88</i></b>	<b><i>.97</i></b>	<b><i>-.25</i></b>	1.0								
5. NA	.03	.01	.10	.01	.92							
6. RNA	<b><i>-.37</i></b>	<b><i>.49</i></b>	<b><i>-.34</i></b>	<b><i>.51</i></b>	.00	1.0						
7. V#	<b><i>.58</i></b>	<b><i>-.69</i></b>	<b><i>.13</i></b>	<b><i>-.71</i></b>	-.03	<b><i>-.33</i></b>	1.0					
8. O#	<b><i>-.13</i></b>	.09	<b><i>.16</i></b>	.09	-.02	<b><i>.14</i></b>	<b><i>.17</i></b>	1.0				
9. ENV	<b><i>.37</i></b>	<b><i>-.35</i></b>	-.03	<b><i>-.37</i></b>	<b><i>.10</i></b>	<b><i>.19</i></b>	<b><i>.24</i></b>	<b><i>.14</i></b>	1.0			
10. HRI	<b><i>.24</i></b>	<b><i>-.20</i></b>	<b><i>.43</i></b>	<b><i>-.19</i></b>	<b><i>-.14</i></b>	<b><i>-.30</i></b>	<b><i>.25</i></b>	<b><i>.13</i></b>	<b><i>.31</i></b>	1.0		
11. SIM	-.06	<b><i>.11</i></b>	<b><i>-.28</i></b>	<b><i>.11</i></b>	<b><i>.29</i></b>	<b><i>.17</i></b>	<b><i>-.12</i></b>	-.07	<b><i>.20</i></b>	-.03	.77	
12. %W	<b><i>.16</i></b>	-.09	.06	-.07	-.02	<b><i>-.15</i></b>	<b><i>-.12</i></b>	<b><i>-.27</i></b>	<b><i>-.25</i></b>	.03	<b><i>-.12</i></b>	1.0

**Notes:** Diagonal contains reliability estimates (*Kappa*). Bold-italicized numbers indicate a significant correlation ( $p < .05$ ). N = 436. **Key:** IA- autonomy to support identification; RIA- reliability of autonomy to support identification; LA- autonomy to support localization; RLA- reliability of autonomy to support localization; NA- autonomy to support navigation; RNA- reliability of autonomy to support navigation, V#- vehicle number; O#- operator number, ENV- testbed environment (1 = physical; 0= virtual); HRI- intent to study HRI operation (1 = yes; 0 = no); SIM- use of simulation vehicle (1 = end user system; 0 = simulation of UV); %W- percent of women in the study

**APPENDIX F: FLOWCHART FOR OBTAINING AND SCREENING  
STUDIES FOR INCLUSION**





## REFERENCES

*References marked with an asterisk indicate studies included in the meta-analysis*

American Psychological Association (APA, 2010). *Publication manual of the American Psychological Association*. Washington, DC: American Psychological Association.

Biederman, I. (1987). Recognition-by-components: A theory of Human Image Understanding. *Psychological Review*, 94(2), 115-147.

\*Billings, D. & Durlach, P. (2008). The Effects of input device and latency on ability to effectively pilot a simulated micro-UAV. *Proceedings of the 52<sup>nd</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 2092-2096.

\*Billings, D. & Durlach, P. (2009). Mission completion time is sensitive to tele-operation performance during simulated reconnaissance missions with a micro-unmanned aerial vehicle. *Proceedings of the 53<sup>rd</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 1408-1412.

\*Blitch, J., Bauder, C., Gutzwiller, R., & Clegg, B. (2012). Correlations of Spatial Orientation with Simulations Based Robot Operator Training. *4<sup>th</sup> International Conference on Applied Human Factors and Ergonomics*, 1571-1580

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: John Wiley & Sons.

Bowers, C., Milham, L., & Price, C. (1998). Dual-task results and the laterization of spatial orientation: Artifact of test selection? *Journal of General Psychology*, 125 (1), 5-16.

Burke, J., & Murphy, R. R. (2004). Human-robot interaction in USAR technical search: two heads are better than one. *13th IEEE International Workshop on Robot and Human Interactive Communication*, 13, 307-312.

- Burke, J. L., Murphy, R. R., Coovert, M. D., & Riddle, D. L. (2004). Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction, 19*, 85-116.
- Casper, J., & Murphy, R. R. (2003). Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center. *IEEE Transactions on Systems, Man and Cybernetics, Part B, 33*, 367-385.
- Carroll, J. (1993). *Human cognitive abilities: A survey of factor analytical studies*. New York, NY: Cambridge University Press.
- Carroll, J. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports g and about ten broad factors. In H. Nyborg (Ed.). *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5–21). Amsterdam: Pergamon.
- \*Cassenti, D.N., Kelly, T.D., Swoboda, J.C., & Patton, D.J. (2009). The effects of communication style on robot navigation performance. *Proceedings of the 53<sup>rd</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 359-363.
- \*Chen, J. (2008). Effectiveness of concurrent performance of military and robotics tasks and effects of cueing in a simulated multi-tasking environment. *Proceedings of the 52<sup>nd</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 238-241.
- \*Chen, J. (2009). Concurrent performance of military and robotics tasks and effects of cueing in a simulated multi-tasking environment. *Presence, 18*, 1-15
- \*Chen, J. (2010a). UAV-guided navigation for ground robot tele-operation in a military reconnaissance environment. *Ergonomics, 53*(8), 940-950.

- \*Chen, J.Y.C. (2010). Effects of operator spatial ability on UAV-guided ground navigation. *Proceedings of the 5<sup>th</sup> International ACM/IEEE Conference on Human-Robot Interaction*, 139-140.
- \*Chen, J. (2011). Individual differences in human-robot interaction in a military multitasking environment. *Journal of Cognitive Engineering and Decision Making*, 5, 83-105.
- \*Chen, J.Y.C., Barnes, M.J. Supervisory control of robots using RoboLeader. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, 1483-1487.
- \*Chen, J.Y.C., & Barnes, M.J. (2012). Supervisory control of multiple robots: Effects of imperfect automation and individual differences. *Human Factors*, 54, 157-174.
- \*Chen, J.Y.C., & Barnes, M.J. (2012). Supervisory control of multiple robots in dynamic task environments. *Ergonomics*, 55(9), 1043-1058.
- Chen, J., Barnes, M., & Harper-Sciari, M. (2010) Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Application and Reviews*, 41(4), 435-454.
- \*Chen, J.Y.C., Barnes, M.J., & Kenny, C. (2011). Effects of unreliable automation and individual differences on supervisory control of multiple ground robots. *Proceedings of the International ACM/IEE Conference on Human-Robot Interaction*, 371-378.
- \*Chen, J.Y.C. Barnes, M.J. & Qu, Z. (2010). RoboLeader: An agent for supervisory control of multiple robots. *Proceedings of the International ACM/IEE Conference on Human-Robot Interaction*, 81-82.
- \*Chen, J.Y.C. Barnes, M.J. & Qu, Z. (2010). *RoboLeader: A surrogate for enhancing the human control of a team of robots*. Technical Report for the Army Research Laboratory, 1-14.

- \*Chen, J.Y.C., Barnes, M.J., Quinn, S.A., & Plew, W. (2011). Effectiveness of RoboLeader for dynamic Re-Tasking in an urban environment. *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting*, 1501-1505.
- \*Chen, J.Y.C., & Clark, B.R. (2008). UAV-guided navigation for ground robot operations. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, 1412-1416.
- \*Chen, J., Durlach, P., Sloan, J., & Bowens, L. (2008). Human robot interaction in the context of simulated route reconnaissance missions. *Military Psychology*, 20(3), 135-150.
- \*Chen, J., & Joyner, C. (2006). *Concurrent performance of gunner's and robotic operator's tasks in a simulated mounted combat system environment*. Technical Report for the Army Research Laboratory.
- \*Chen, J., & Joyner, C. (2006). Individual differences in concurrent performance of gunner's and robotic operator's tasks. *50<sup>th</sup> Proceedings of the Human Factors & Ergonomics Society*, 50, 1759-1763.
- \*Chen, J., & Joyner, C. (2009). Concurrent performance of gunner's and robotics in a multitasking environment. *Military Psychology*, 21(1), 98-113.
- \*Chen, J. & Terrence, P. (2007). Effects of tactile alerts on concurrent performance of the gunner's and robotic operator's tasks in a simulated mounted environment. *Technical Report for the Army Research Laboratory*.
- \*Chen, J., & Terrence, P. (2008). Effects of tactile cueing on concurrent performance of military and robotics tasks in a simulated multitasking environment. *Ergonomics*, 51(8), 1137-1152.

- \*Chen, J. & Terrence, P. (2008). Individual differences in concurrent performance of military and robotics tasks with tactile cueing. *Proceedings of the 52<sup>nd</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 52, 1408-1411.
- \*Chen, J., & Terrence, P. (2009). Effects of imperfect automation and individual differences on concurrent performance of military and robotics tasks in a simulated multitasking environment. *Ergonomics*, 52(8), 907-920.
- \*Chien, S., Wang, H., Lewis, M.(2011). Effects of spatial ability on multi-robot control tasks. *Proceedings of the Human Factors & Ergonomics Society*, 894-898.
- Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. (2004). Working memory is (almost) perfectly predicted by *g*. *Intelligence*, 32, 277-296.
- Conway, A., Kane, M., & Engle, R. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Science*, 7(12), 547-552.
- \*Cook, M.B., Smallman, H.S., Lacson, F.C., & Manes, D.I. (2010). Situation displays for dynamics UAV replanning: Intuitions and performance for display formats. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, 492-496.
- \*Cosenzo, K., Chen, J., Reinerman-Jones, L., Barnes, M., & Nicholson, D. (2010). Adaptive automation effects on operator performance during a reconnaissance mission with an unmanned ground vehicle. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, 2135-2139.
- Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Cucina, J., Gast, I., & Su, C. (2012). *g* 2.0: Factor analysis, filed findings, facts, fashionable topics, and future steps, *Journal of Industrial and Organizational Psychology: Perspectives on Science and Practice*, 5(2),166-171.

- \*Cutmore, T.R.H., Hine, T.J., Maberly, K.J., Langford, N.M., & Hawgood, G. (2000). Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human- Computer Studies*, 53, 223-249.
- \*Darken, R.P. & Cevik, H. (1999). Map usage in virtual environments: Orientation issues. Proceedings of IEEE Virtual Reality 99, 133-140.
- DARPA (2006). Report to Congress: DARPA Prize Authority (Fiscal Year 2005 Report). Retrieved September, 19, 2007, [http://www.darpa.mil/GRANDCHALLENGE/docs/Grand\\_Challenge\\_2005\\_Report\\_to\\_Congress.pdf](http://www.darpa.mil/GRANDCHALLENGE/docs/Grand_Challenge_2005_Report_to_Congress.pdf)
- Day, E.A., Arthur, W., Miyashiro, B., Edwards, B.D., Tubre, T.C., & Tubre, A.H. (2004). Criterion-related validity of statistical operationalizations of group general cognitive ability as a function of task type: Comparing the mean, maximum, and minimum. *Journal of Applied Social Psychology*, 34(7), 1521-1549.
- \*Diaz, D.D., & Sims, V.K. (2001). The use of a route and survey composite display for navigational training in virtual environments. *45th proceedings of the Human Factors and Ergonomics Society*, 1857-1861.
- \*Diaz, D.D., & Sims, V.K. (2003). Augmenting virtual environments: the influence of spatial ability on learning from integrated displays. *High Ability Studies*, 14, 191-212.
- \*Driscoll, I., Hamilton, D.A., Yeo, R.A., Brooks, W.M., & Sutherland, R.J. (2005). Virtual navigation in humans: The impact of age, sex, and hormones on place learning. *Hormones and Behavior*, 47, 326-335.
- Ekstrom, R., French, J., Harman, H, Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, New Jersey: Educational Testing Service.
- Eliot, J. & Smith, I. (1983). An international directory of spatial tests. Windsor: NFER-Nelson.

- Endsley, M. & Kiris, E. (1995). The out-of-the-loop performance problem and level of control. *Human Factors*, 37(2), 381-394.
- Fatolitis, P. (2008). Initial validation of novel performance-based measures: Mental rotation and psychomotor ability. Unpublished Dissertation, University of Central Florida, Orlando, FL.
- \*Fincannon, T., Curtis, M., & Jentsch, F. (2006). Familiarity and expertise in the recognition of vehicles from and an unmanned ground vehicle. *Proceeding of the 50<sup>th</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 1218-1222
- \*Fincannon, T, Evans, III, A.W., Jentsch, F, & Keebler, J. (2008). Interactive effects of backup behavior and spatial abilities in the prediction of teammate workload using multiple unmanned vehicles. *Proceeding of the 52nd Annual Meeting of the Human Factors and Ergonomics Society*. 995-999.
- \*Fincannon, T, Evans, III, A.W., Jentsch, F, & Keebler, J. (2010a). Dimensions of spatial ability and their influence on performance with unmanned systems. In D. H. Andrews, R. P. Herz, & M. B. Wolf (eds.), *Human factors in defense: Human factors in combat identification* (pp. 67-81). Burlington, VT: Ashgate Publishing.
- \*Fincannon, T., Evans, A. W., Jentsch, F., Phillips, E., & Keebler, J. R. (2009). The effects of sharing control of unmanned vehicles on backup behavior and workload in distributed teams. *Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society*, 1300-1303.
- \*Fincannon, T., Evans, A. W., Phillips, E., Jentsch, F., & Keebler, J. (2009). The influence of team size and communication modality on team effectiveness with unmanned systems. *Proceedings of the 53rd Annual Meeting of the human Factors and Ergonomics Society*, 419-423.



- \*Fincannon, T., Jentsch, F., Sellers, B., & Keebler, J. R. (2012). Beyond “spatial ability”:  
Examining the impact of multiple individual differences in a perception by proxy  
framework. *Proceedings of the 7th ACM/IEEE International Conference on Human-  
Robot Interaction*, 127-128.
- Fincannon, T., Keebler, J. R., & Jentsch, F. (in press-a). Examining external validity issues in  
research with human operation of unmanned vehicles. *Theoretical Issues in Ergonomic  
Science*.
- \*Fincannon, T., Keebler, J. R., Jentsch, F., & Curtis, M. (in press-b). The influence of  
camouflage, obstruction, familiarity, and spatial ability on target identification from an  
unmanned ground vehicle. *Ergonomics*.
- \*Fincannon, T., Keebler, J. R., Jentsch, F., & Evans, III, A.W. (2008). Target identification  
support and location support among teams of unmanned system operators [Electronic  
version]. *Proceedings of the 26<sup>th</sup> Annual Army Science Conference*.  
[http://www.dtic.mil/cgi-  
bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA504210](http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA504210)
- Fincannon, T, Keebler, J. R., Jentsch, F., Phillips, E., & Evans, III, A.W. (2011). Team size, team  
role, communication modality, and team coordination in the distributed operation of  
multiple heterogeneous unmanned vehicles. *Journal of Cognitive Engineering and  
Decision Making*, 5(1): 106-131.
- \*Fincannon, T., Ososky, S., Keebler, J., Evans, III, A.W., & Jentsch, F (2010d). *The influence  
UGV operator spatial ability on a UAV teammate’s reconnaissance report*. Poster  
presented at the 118<sup>th</sup> Annual Meeting of the American Psychological Association, San  
Diego, CA.

- \*Fincannon, T., Ososky, S., Keebler, J., & Jentsch, F. (2010b). Acquisition of skill sets and mental models over time. In Kaber, D. & Boy, G. (Eds.). *Advances in Cognitive Ergonomics* (pp. 597-606). New York: CRC Press.
- \*Fincannon, T., Ososky, S., Jentsch, F., Phillips, E., & Keebler, J. (2010c). Some good and bad with spatial ability in three person teams that operate multiple unmanned systems. *Proceedings of the 54th Annual Meeting of the human Factors and Ergonomics Society*, 1615-1619.
- \*Gomer, J.A. (2010). Spatial perception and robot operation: The relationship between visual spatial ability and performance under direct line of sight and teleoperation. Unpublished Dissertation. Clemson University, Clemson, SC.
- \*Gomer, J.A., & Pagano, C.C. (2011). NASA Task Load Index for human-robot interaction workload measurement. *ITEA Journal*, 32, 210-214.
- \*Gomer, J.A., & Pagano, C.C. (2012). Spatial perception and robot operation: Selection, training, gaming, & design. *Naval Engineers Journal*, 124(1), 131-138.
- \*Gomer, J.A., & Pagano, C.C. (2011). Spatial Perception and robot operation: Should spatial abilities be considered when selecting robot operators. *Proceedings of the 55th Human Factors and Ergonomics Society Annual Meeting*, 55, 1260-1264.
- Goodrich, M., Morse, B., Gerhardt, D., Cooper, J., Quigley, M., Adams, J., et al. (2008). Supporting wilderness search and rescue using a camera-equipped minim UAV. *Journal of Field Robotics*, 25(1-2), 89-110.
- \*Gugerty, L., & Brooks, J. (2004). Reference-frame misalignment and cardinal direction judgments: Group differences and strategies. *Journal of Experimental Psychology: Applied*, 10(2), 75-88.

- \*Gugerty, L., Brooks, J., & Treadaway, C. (2004). Individual differences in situation awareness for transportation tasks. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory, measures and application* (pp. 193–212). London: Ashgate Publishers.
- \*Gugerty, L. & Rodes, W. (2007). A cognitive model of strategies for cardinal direction judgments. *Spatial Cognition and Computation*, 7(2), 179-212.
- Guilford, J. P. (1985). The structure-of-intellect model. In B. B. Wolman ( Ed.), *Handbook of Intelligence* (pp. 225-266). New York: Wiley.
- Guilford, J. P. & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1980). Fluid and crystallized intelligences: Two fanciful concepts. *Psychological Bulletin*, 88, 406-412.
- Haris, C. (1967). On factors and factor scores. *Psychometrika*, 13, 363-379.
- \*Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., Lovelace, K. (2006) Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34, 151-176.
- Hegarty, M. & Waller, D. (2005). Individual differences in spatial ability. In P. Shah & A. Miyake (Eds.), *The Cambridge Handbook of Visuospatial Thinking* (pp. 121-169). Cambridge University Press.
- Horn, J. (1998). A basis for research on age differences in cognitive abilities. In J.J. McArdle & R.W. Woodcock (Eds.), *Human Cognitive Abilities in Theory and Practice* (pp. 57-92). Mahwah, NJ: Erlbaum.
- Horn, J. & Cattell, R. (1966). Refinement and test of the theory of fluid and crystallized intelligences. *Journal of Educational Psychology*, 57(5): 283-270

- Horn, J. & Cattell, R. (1982). Whimsy and misunderstandings of Gf-Gc theory: A comment on Guilford. *Psychological Bulletin*, 91(3): 623-633.
- Horn, J. & Knapp, J. (1973). On the subjective character of the empirical base of Guilford's Structure-of-Intellect model. *Psychological Bulletin*, 80, 33-43.
- Horn, J. & Knapp, J. (1974). Thirty wrongs do not make a right: Reply to Guilford. *Psychological Bulletin*, 80, 33-43.
- Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.). *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-81). New York: Guilford
- Hunter, J. & Schmidt, F. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage Publications, Inc.
- Huthmann, A. (2009). *Presence-dependent performance differences between virtual simulations and miniature worlds*. Unpublished Dissertation, University of Central Florida, Orlando, FL.
- \*Jentsch, F., Barnes, M. J., Fincannon, T., Keebler, J., Ososky, S., Phillips, E., & Schuster, D. (2011). *Safe operation of autonomous robot teams: Impact of autonomy, teaming, and workload on operator performance*. Poster presented at the 2011 Army Research Laboratory Robotics Technical Assessment Board (TAB), Aberdeen Proving Ground, MD, 08/09 August 2011.
- \*Jentsch, F., Fincannon, T., Ososky, S., Keebler, J., Leotaud, P., & Phillips, E. (2010). *Effects of robot teaming and reliability on operator supervisory control*. Orlando, FL: University of Central Florida
- \*Jentsch, F., Fincannon, T., Sellers, B., Keebler, J., Ososky, S., Phillips, E., & Schuster, D. (2011). *Supervisory control of robot teams by human teams. [Task Order 132] Safe*

*Operation of Autonomous Robot Teams: Impact of Autonomy, Teaming, and Workload on Operator Trust and Performance*. Orlando, University of Central Florida.

\*Jentsch, F., Fincannon, T., Sellers, B., Keebler, J., Ososky, S., Phillips, E., & Schuster, D.

(2012). Year 2 final report: Supervisory control of robots in urban operations. [Task Order 132] *Safe Operation of Autonomous Robot Teams: Impact of Autonomy, Teaming, and Workload on Operator Trust and Performance*. Orlando, University of Central Florida.

Kalaian, H. & Raudenbush, S. (1996). A multi linear model for meta-analysis. *Psychological Methods*, 1(3), 227-235.

Kaufman, S. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence*, 35, 211-223.

\*Keebler, J.R. (2011). Effects of 3D stereoscopy, visuo-spatial working memory and enjoyment on memorization of confusable objects. Unpublished Dissertation, University of Central Florida, Orlando, FL.

\*Keebler, J. R., Sciarini, L., Jentsch, F., Fincannon, T., & Nicholson, D. (2008). Effects of training modality on military vehicle identification in a virtual environment. *Proceeding of the 52nd Annual Meeting of the Human Factors and Ergonomics Society*. 2052-2056

\* Keebler, J. R., Sciarini, L.W., Jentsch, F., & Fincannon, T. (In Press). Using physical 3D objects as training media for military vehicle identification. *Journal of Ergonomics*.

\*Keebler, J. R., Sciarini, L., Jentsch, F., Nicholson, D., & Fincannon, T. (2010). A cognitive basis for vehicle misidentification. In D. H. Andrews, R. P. Herz, & M. B. Wolf (Eds.). *Human factors in defense: Human factors in combat identification* (pp. 113-129). Burlington, VT: Ashgate Publishing.

- Keehner, M., Tendick, F., Meng, M., Anwar, H., Hegarty, M., Stoller, M., & Duh, Q. (2004). Spatial ability, experience, and skill in laparoscopic surgery. *The American Journal of Surgery, 188*, 71-75.
- Landis, R. & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.
- \*Lathan, C., & Tracey, M. (2002). The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence: Teleoperators & Virtual Environments, 11*(4), 368-377.
- Lohman, D. (1989). Human intelligence: An introduction to advances in theory and research. *Review of Educational Research, 59*(4): 333-373.
- Lohman, D. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.) *Human abilities: Their nature and measurement* (pp. 97-116). Mahwah: Erlbaum.
- Lohman, D., Pellegrino, J., Alderton, D., & Regian, J. (1987). Dimensions and components of individual differences in spatial abilities. In S. Irvine & S. Newstead (Eds.) *Intelligence and cognition: Contemporary frames of reference* (pp. 253-312). Dordrecht: Martinus Nijhoff.
- \*Long, L.O., Gomer, J.A., Moore, K.S., Pagano, C.C. (2009). Investigating the relationship between visual spatial abilities and robot operation during direct line of sight and teleoperation. Proceedings of the 53rd Human Factors and Ergonomics Society Annual Meeting, 53, 1437-1441.
- \*Long, L.O., Gomer, J.A., Moore, K.S., Pagano, C.C. (2011). Visual spatial abilities in uninhabited ground vehicle task performance during teleoperation and direct line of sight. *Presence: Teleoperators & Virtual Environments, 20*, 466-479.

- McDermott, P., Luck, J., Allender, L., & Fisher, A. (2005). Effective human to human communication of information provided by an unmanned vehicle. *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, 402-406.
- McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889-918.
- McGrew, K.S. (2005) The Cattell-Horn-Carroll theory of cognitive abilities: Past, present and future. In D. Flanagan & Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues- Second Edition* (p. 136-202). New York: Guilford Press.
- McGrew, K.S. (2009). Editorial: CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1-10.
- Mendenhall, W. & Sincich, T. (2007). *Statistics for engineering and the sciences* (5<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Prentice Hall, Inc.
- Michael, W., Zimmerman, W. & Guilford, J. (1951). An investigation of the nature of the spatial-relations and visualization factors in two high school samples. *Educational and Psychological Measurement*, 11(4): 561-577.
- Michael, W., Guilford, J., Fruchter, B., & Zimmerman, W. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement*, 17: 185-199.
- Miyake, A., Friedman, N., Rettinger, D., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, 130(4), 621-640.
- \*Moffat, S., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a “virtual” maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution & Human Behavior*, 19, 73-87.

- \*Ophir-Arbelle, R., Oron-Gilad, T., Borowsky, A., & Parmet, Y. (2012). Is More Information Better? How Dismounted Soldiers Use Video Feed from Unmanned Vehicles: Attention Allocation and Information Extraction Concerns. *Journal of Cognitive Engineering & Decision Making*, 7(1), 26-48.
- Parasuraman, R., Molloy, R., & Indramani, S.L. (1993). Performance consequences of automation-induced "Complacency". *The International Journal of Aviation Psychology*, 3(1), 1-23.
- Parasuraman, R. & Riley, V. (1997) Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.
- Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 30(3), 286-297.
- \*Phillips, E., Fincannon, T., Leotaud, P., & Jentsch, J. (2010a). Effects of robot teaming and operator memory on the rate of human interaction with multiple unmanned vehicles. *Proceedings of the 27th Army Science Conference*, Orlando, Florida. <http://www.armyscienceconference.com/manuscripts/E/EP-005.pdf>
- Phillips, E., Ososky, S. & Grove, J. (2011) From tools to teammates: Toward the development of appropriate mental models for intelligent robots. *Proceedings of the 55<sup>th</sup> Annual Meeting of the Human Factors and Ergonomics Society*. Las Vegas, NV.
- \*Phillips, E., Rivera, J., Fincannon, T., Dietz, A., & Afek, A. (2010b). The relationship of closure speed and training on target identification for unmanned vehicle operators. *Proceedings of the 54th Annual Meeting of the Human Factors and Ergonomics Society*, 1478-1482.
- \*Rehfeld, S. (2006). Mental rotation training across levels of automation in human-robot interaction. Unpublished Dissertation, University of Central Florida, Orlando, FL.



- \*Rehfeld, S., Jentsch, F., Curtis, M., & Fincannon, T. (2005). Collaborative teamwork with unmanned vehicles in military missions. *Proceedings of the 11th International Conference on Human-Computer Interaction* [Electronic Version]. Las Vegas, NV.
- \*Rodes, W., Gugerty, L., Brooks, J., & Cantalup. (2009). The effects of electronic map displays and spatial ability on performance of navigational tasks. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*, 53, 369-373.
- \*Schuster, D., Fincannon, T., Jentsch, F., Keebler, J., & Evans, A. W., III. (2008). *The role of spatial ability in the relationship between video game experience and route effectiveness among unmanned vehicle operators* [Online, Accessed 29 March, 2011]. Proceedings of the 26<sup>th</sup> Annual Army Science Conference. Available from: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA505705&Location=U2&doc=GetTRDoc.pdf>
- \*Schuster, D., Fincannon, T., Keebler, J., & Jentsch, F. (2010). *Spatial ability and video game experience in vehicle identification*. Poster session presented at the meeting of the American Psychological Association, San Diego, CA.
- \*Schuster, D., Keebler, J.R., Zuniga, J., & Jentsch, F. (2012). Individual differences in SA measurement and performance in human-robot teaming. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 187-190.
- Schuster, D., Ososky, S., Jentsch, F., Phillips, E., & Evans, A. W. (2011). A research approach to shared mental models and situation assessment in future robot teams. *Proceedings of the 55<sup>th</sup> Annual Meeting of the Human Factors and Ergonomics Society*. Las Vegas, NV.
- \*Sekmen, A.S., Wilkes, M., Goldman, S.R., & Zein-Sabatto, S. (2003). Exploring importance of location and prior knowledge of environment on mobile robot control. *International Journal of Human-Computer Studies*, 58, 5-20.

- \*Sellers, B., Fincannon, T., & Jentsch, F. (2011). Indexing spatial ability across team size: The influence of the weakest link, strongest link, and aggregate ability on performance with multiple unmanned systems. *Proceedings of the 55th Annual Meeting of the Human Factors and Ergonomics Society*, Las Vegas, NV.
- Shepard, R. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
- \*Sloan, J. (2005). Human-robot interaction in the context of simulated route reconnaissance missions. Unpublished Dissertation. University of Central Florida: Orlando FL.
- Steiner, I.D. (1972). *Group process and productivity*. New York: Academic Press.
- Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Schultz, J., Schultz, A., & Goodrich, M., (2006). Common metrics for human-robot interaction. *Proceedings of the 1<sup>st</sup> International ACM/IEEE Conference on Human-Robot Interaction*, 33-40.
- Tabachnick, B. & Fidell, L. (2007). *Using multivariate statistics (4<sup>th</sup> ed.)*. New York, NY: Pearson Education, Inc.
- \*Takayama, L., Marder-Eppstein, E., & Harris, H. (2011). Assisted driving of a mobile remote presence system: System design and controlled user evaluation. *IEEE International Conference on Robotics and Automation*, 1883-1889.
- \*Taylor, G. (2012). Comparing types of adaptive automation within a multi-tasking environment. Unpublished Dissertation. University of Central Florida: Orlando, FL.
- \*Taylor, G., Reinerman-Jones, L., Cosenzo, K., & Nicholson, D. (2010). Comparison of multiple physiological sensors to classify operator state in adaptive automation systems. *Proceedings of the 54<sup>th</sup> Annual Meeting of the Human Factors and Ergonomics Society*, 195-199.
- Thurstone, L. (1938) *Primary mental abilities*. Chicago: University of Chicago Press.

- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-604.
- \*Velez, M., Silver, D., & Tremaine, M. (2005). Understanding visualization through spatial ability differences. *IEEE Conference on Visualization*, 511-518
- Vogel, J., Bowers, C., & Vogel, D. (2003). Cerebral laterization of spatial abilities: A meta-analysis. *Brain and Cognition*, 52(2), 197-204.
- Voyer, D., Voyer, S., & Brydent, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117(2): 250-270.
- \*Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, 4, 307-321.
- \*Waller, D. (2000). Individual differences in virtual environments. *Unpublished manuscript*.  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.46.4687&rep=rep1&type=pdf>
- \*Waller, D., Napp, D., & Hunt, E. (2001). Spatial representations of virtual mazes: The role of visual fidelity and individual differences. *Human Factors*, 43, 147-158.
- Wang, J., & Lewis, M. (2007). Human control for cooperating robot teams. *Proceedings of the 2nd ACM Conference on Human Robot Interaction*, 9-16.
- Wong, J.T. (2009). Spatial abilities and performance in robot navigation. *The College of New Jersey Journal of Student Scholarship*, 11, 1-10.
- Woods, D. D., Tittle, J., Feil, M., & Roesler, A. (2004). Envisioning human-robot coordination in future operations. *Systems, Man and Cybernetics, Part C, IEEE Transactions on*, 34(2), 210-218.
- Zacks, J. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience* 20(1): 1-19.