2014

Examining the Effects of Interactive Dynamic Multimedia and Direct Touch Input on Performance of a Procedural Motor Task

Matthew Marraffino
University of Central Florida

EXAMINING THE EFFECTS OF INTERACTIVE DYNAMIC MULTIMEDIA AND DIRECT TOUCH INPUT ON PERFORMANCE OF A PROCEDURAL MOTOR TASK

by

MATTHEW D. MARRAFFINO

B.A. University of Central Florida, 2008
M.A. University of Central Florida, 2013

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

Summer Term
2014

Major Professor: Valerie K. Sims

ABSTRACT

Ownership of mobile devices, such as tablets and smartphones, has quickly risen in the last decade. Unsurprisingly, they are now being integrated into the training and classroom setting. Specifically, the U.S. Army has mapped out a plan in the Army Learning Model of 2015 to utilize mobile devices for training purposes. However, before these tools can be used effectively, it is important to identify how the tablets' unique properties can be leveraged. For this dissertation, the touch interface and the interactivity that tablets afford were investigated using a procedural-motor task. The procedural motor task was the disassembly procedures of a M4 carbine. This research was motivated by cognitive psychology theories, including Cognitive Load Theory and Embodied Cognition. In two experiments, novices learned rifle disassembly procedures in a narrated multimedia presentation presented on a tablet and then were tested on what they learned during the multimedia training involving a virtual rifle by performing a rifle disassembly on a physical rifle, reassembling the rifle, and taking a written recall test about the disassembly procedures. Spatial ability was also considered as a subject variable.

Experiment 1 examined two research questions. The primary research question was whether including multiple forms of interactivity in a multimedia presentation resulted in higher learning outcomes. The secondary research question in Experiment 1 was whether dynamic multimedia fostered better learning outcomes than equivalent static multimedia. To examine the effects of dynamism and interactivity on learning, four multimedia conditions of varying levels of interactivity and dynamism were used. One condition was a 2D phase diagram depicting the before and after of the step with no animation or interactivity. Another condition utilized a non-

interactive animation in which participants passively watched an animated presentation of the disassembly procedures. A third condition was the interactive animation in which participants could control the pace of the presentation by tapping a button. The last condition was a rifle disassembly simulation in which participants interacted with a virtual rifle to learn the disassembly procedures. A comparison of the conditions by spatial ability yielded the following results. Interactivity, overall, improved outcomes on the performance measures. However, high spatials outperformed low spatials in the simulation condition and the 2D phase diagram condition. High spatials seemed to be able to compensate for low interactivity and dynamism in the 2D phase diagram condition while enhancing their performance in the rifle disassembly simulation condition.

In Experiment 2, the touchscreen interface was examined by investigating how gestures and input modality affected learning the disassembly procedures. Experiment 2 had two primary research questions. The first was whether gestures facilitate learning a procedural-motor task through embodied learning. The second was whether direct touch input using resulted in higher learning outcomes than indirect mouse input. To examine the research questions, three different variations of the rifle disassembly simulation were used. One was identical to that of Experiment 1. Another incorporated gestures to initiate the animation whereby participants traced a gesture arrow representing the motion of the component to learn the procedures. The third condition utilized the same interface as the initial rifle disassembly simulation but included “dummy” gesture arrows that displayed only visual information but did not respond to gesture. This condition was included to see the effects (if any) of the gesture arrows in isolation of the gesture component. Furthermore, direct touch input was compared to indirect mouse input. Once again, spatial ability also was considered. Results from Experiment 2 were inconclusive as no significant effects were found. This may have been due to a ceiling effect of performance. However, spatial ability was a significant predictor of performance across all conditions.

Overall, the results of the two experiments support the use of multimedia on a tablet to train a procedural-motor task. In line with vision of ALM 2015, the research support incorporating tablets into U.S. Army training curriculum.

This work is dedicated to my family.

ACKNOWLEDGMENTS

This work was supported in part by the Army Research Institute. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army or U.S. Government.

I would like to thank the members of my committee: Dr. Valerie Sims, Dr. Cheryl Johnson, Dr. Matthew Chin, and Dr. Mustapha Mouloua. I have tremendous respect for the work you do and the guidance you provided me during, not only my dissertation, but also my graduate career as a whole. I would also like to thank IST, specifically Brian Plamondan and Dean Reed for developing the Rifle Disassembly Application. I would also like to thank Dr. Glenn Martin for allowing me to use his facilities and for continually encouraging me to work more I would like to thank John Killilea for helping get this research effort off the ground. I would like to thank all of the members of the ACAT lab for their help running participants, coding data, and providing feedback. I would like to thank Anne Sinatra for introducing me to Human Factors psychology and starting my career with the ACAT lab. I would like to thank Heather Lum for her tutelage and for guiding me through my first years in graduate school. To my wife, Andi, this dissertation is as much yours as it is mine. Thank you for always being there for support and for pushing me to succeed. To my daughter, your smile made every day brighter. To my Mom, Dad, and Sister for encouraging me and supporting me in everything I have ever done.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

The Army is interested in tablets for training. Because tablet technology has only recently taken off, little research exists to guide utilization practices for training. This dissertation represents a first step in identifying features that promote effective training by tablets, by examining how to capitalize on the touchscreen capability. This research was motivated by cognitive psychology theories, including Cognitive Load Theory and Embodied Cognition. In two experiments, novices learned rifle disassembly procedures in a narrated multimedia presentation presented on a tablet and then were tested on what they learned by performing rifle disassembly on a physical rifle, reassembling the rifle, and taking a written recall test about the disassembly procedures. Specifically, in Experiment 1, my primary research question concerned how the level of interactivity impacts learning and cognitive load; my secondary research question concerned whether animations were as effective as their static counterparts. In this experiment, Participants received training on a tablet from a Rifle Disassembly Simulation (RDS), Interactive Animation, Non-Interactive Animation or 2D Phase Diagram. In Experiment 2, I explored how to best interact with the RDS by incorporating gestures (tapping to interact with the simulation or dragging to simulate real rifle disassembly procedures). Additionally, I also examined the impact of using the touchscreen versus using a more indirect interface (a mouse). I had two primary research questions in Experiment 2 concerning the impact of gestures on learning and touchscreens on learning. Six conditions were used in Experiment 2. Three different variations of the RDS, including a tap interface where participants tapped on components to interact with them, a gesture interface, where participants traced gesture arrows on the screen to mimic the real-world action, and a tap interface that contained arrows to account for any information simply gained from the presence of additional visual information. These groups were crossed with input modality; either direct touch or mouse input.

The following sections will provide background research to support the hypotheses addressed by the experiments. First, I discussed current research involving incorporating tablets in training and education with a specific emphasis on how the Army would like to integrate mobile devices in their model for training. Next, Cognitive Load Theory and the Cognitive Theory of Multimedia Learning were discussed as theoretical bases for learning from multimedia presentations on tablet computers in the context of how these theories predict the effectiveness of interactive and non-interactive forms of multimedia. Next, I discussed the theory of Embodied Cognition in the context of learning from a touch interface. Following the background, I describe the importance of the task being investigated, disassembling a Colt M4 carbine, the standard issue weapon for the Army followed by how this task was incorporated into a rifle disassembly training application. Lastly, I described, in detail, two experiments addressing the posed research questions.

Tablets for Training and Education

Currently, mobile devices, specifically tablet computers, are being integrated into various fields. For example, tablets are used in cockpits as part of an electronic flight bag to eliminate the use of paper based texts (Bassanesi & Tindall, 2011). Another example comes from medicine, where Johnson et al. (2012) demonstrated tablets' efficacy as a display device for radiology outputs. However, one of the largest potentials for tablet computers is to offer effective training solutions available to anyone, at anytime due to their portability and ability to download applications anywhere a wireless network exists. Tablets offer several features that could be useful for training integration. For example, they offer a larger screen than other mobile devices such as smartphones. This means higher levels of interactive features that might prove otherwise challenging on smaller devices. Other features include GPS technology, and apps that can be downloaded, pushed, and accessed anywhere with a network connection. This means that trainers and educators can reach individuals in the field with just-in-time information and support.

Several research studies have focused on pushing tablets into an existing curriculum. However, during integration efforts, curriculums were not adjusted to incorporate the tablets' specific capabilities. Furthermore, plans regarding the tablets' use within the curriculum were not specific (Tucker, 2010). These research efforts have resulted in mixed success. For example, Bush and Cameron (2011) distributed iPads to college students in three Master's level courses and collected data regarding usability, faculty and student perceptions, personal and academic uses, impact on learning, and several other factors. Their study indicated that students rated the iPad as good or better than the traditional printed materials, and that the multimodal affordances (e.g., ability to browse the web, use GPS, and other applications in addition to being an e-reader) of the device aided in the adoption of the device. Notwithstanding, instructors did not rate any change in student involvement or engagement and a couple of instructors expressed concern about poorer student comprehension relative to traditional instructional materials. In another study, Marmarelli and Ringle (2011) distributed iPads out to students in a political science class at a university. Surveys were administered at the beginning and middle of the semester in order to provide feedback regarding the status and impact of tablet technology for curricular use. Although it was not found to be a perfect solution, students reported an overall positive perception of the devices. Specifically, the size, contrast, and resolution of the device were noted to be good for reading texts. Additionally, the battery life and ability to save paper were also mentioned as positive benefits of the tablet. Unfortunately, this particular study only utilized subjective ratings and did not have a specific uses for the tablet other than as a general tool for reading texts. Similarly, Handy, Suter, and Hooper (2011) handed iPads out to university students in another study examining students' perceptions of the technology. However, unlike the previously described studies, Professors encouraged the students to use the tablets inside and outside of class as well as made video lectures and electronic textbooks available in tablet-friendly formats. Pre- and post- surveys were administered to students to compare attitudes towards the tablets before and after tablets were distributed. The results of the study indicated that students had an overall positive perception of the tablets. Specifically, students reported that they spent more time on course related material after being given the tablet. However, only subjective ratings were used, and there were no data on how they performed in the class. Moreover, although video and texts were formatted to the tablet, no applications were developed specifically for the tablet. Several other studies have reported similar perceptions of tablet-centric curriculum based on subjective ratings (e.g., Rossing, Miller, Cecil, & Stamper, 2012; Tucker, 2010).

On the other hand, a recent study investigating the effects of integrating iPads into an Army training curriculum found contrary findings to the generally positive usage surveys found in the previously described studies. Killilea, Marraffino, and Singer (2013) distributed tablet computers to Soldiers in the Signal Captains Career Course at the Army Signal Center of Excellence. Soldiers completed an initial survey asking their predicted use, utility and attitudes regarding the tablets. Following the study, a post-test survey asked the soldiers to record their actual usage of the devices. Overall perception of the devices was positive with Soldiers tending to rate the devices as helping them complete tasks more effectively and being useful in the course. However, most participants reported using the tablet as little as on a weekly or monthly basis, contrary to their perceived usage ratings at the beginning of the study. Tablets, although perceived to be a potentially useful tool, were not effectively utilized. This experiment demonstrated that although tablets may appear as useful training tools, the authors concluded that courses must be designed around their functionality.

The existing literature exhibits two notable shortcomings. First, the measures were typically self-report ratings of perceived utility. It is unclear how integrating the tablets into the curriculum affected a measurable performance outcome (e.g., grades on tests of learning retention). Second, the tablets were used primarily as e-readers without any content specifically devoted to the available technology. In other words, instead of using some of the more interactive features, the tablets were used primarily as textbook replacements. In order to make tablet adoption successful in training and education curricula, research should focus on the tablets' unique properties in addition to their ability to present textual information. Personal computers have established themselves as successful pieces of technology for training, but tablets exhibit properties that set them apart from their PC counterparts. Integration efforts should design their curriculum around the specific capabilities of the tablets (Murphy, 2011). In order to help better integrate tablets into the training sector, the present studies described here investigated the ability of the touchscreen, ubiquitous to all tablet computers, to serve as an effective interface for learning a procedural motor task. In order to effectively evaluate the touchscreen, two different studies addressed its utility. The first study examined the touchscreen's ability to help foster interactivity by reducing the cognitive load associated with using an interface during a multimedia presentation. The second study examined whether physically touching a screen improves learning gains versus indirect forms of interaction, such as using a mouse.

Tablets for Army Training

In response to the potential of mobile computing for training, the U.S. Army is undergoing a paradigm shift in how Soldiers and Leaders are being trained. The Army Learning Model of 2015 (ALM 2015; Alley, 2010), a document released by the Army's Training and Doctrine Command (TRADOC), describes a new vision for training that is learner-centric. ALM 2015 envisions a classroom in which the instructor acts as a facilitator and learners are active participants, as opposed to an instructor-centric classroom where the instructor delivers a lecture and learners are passive participants. The previous learning model was met with several challenges. One issue was that course lengths were predetermined and could not adapt to meet individual learners' needs. Another challenge was the inability to provide training to commanders in timely fashion. Previous training efforts often lagged behind the learner's level of experience due to the time and resources required to send a mobile training team to a location. Although the Army began adopting a distributed learning model nearly 20 years ago, the goal of anytime, anywhere training was not met (ALM, 2015). To that end, a large part of this new vision calls for training to occur at the point of need and be available to Soldiers anytime, anywhere. Using mobile training technologies will play a key role in realizing the Army's vision for training for their ability to provide on-demand training requirements. However, despite the wide interest in using mobile devices in education, medicine and military training, there has been little research on how to make them effective as training tools. The current dissertation focuses on how to leverage the interactive touchscreen to train a procedural motor task.

Research in the area has focused on the mobility aspect of tablet computers in an attempt to provide meaningful learning outside the classroom, and to integrate the technology into existing curriculums. Unfortunately, little research has examined how to develop specific tablet-specific training applications leveraging the features inherent to tablets, specifically the touch interface (Tucker, 2010). To address the gap in the literature, this dissertation addressed two primary research questions in two experiments. First, can interactivity, afforded by tablet computers, be effective at training a procedural motor task? Second, does using the touchscreen offer a better interface than a traditional mouse and keyboard by which to encode new material? The procedural motor task used to address the research questions was disassembling a Colt M4 carbine.

Cognitive Load Theory

The theoretical framework for describing the potential advantages of utilizing interactive graphics to train procedural-motor tasks stems from Cognitive Load Theory (Cognitive Load Theory; Sweller, van Merriënboer, & Paas, 1998; Paas, Renkl & Sweller, 2003). Cognitive Load Theory was developed as a framework to describe how learners process information from working memory (Baddeley & Hitch, 1974; Baddeley, 1986) to long-term memory in the form of schemas. Schemas are categories of elements that are organized in the manner in which they are to be used. Schemas combine several elements of information into a single element that can be manipulated and expanded to account for new information. Schemas are efficient knowledge structures that allow the mind to deal with one element instead of several. Cognitive Load Theory posits that as individuals learn about a particular topic or skill, they combine lower level schemas with higher level ones to create increasingly complex schemas. The theory has been used to develop effective instructional designs that align with human cognitive architecture (Paas, Renkl, & Sweller, 2004). Cognitive Load Theory relies on a few primary assumptions. First, working memory is limited when dealing with new information (Miller, 1956). Second, working memory contains two sub components that independently handle visual and auditory information (Baddeley and Hitch, 1974). A third assumption is that, although working memory is limited, long-term memory is essentially unlimited and learning occurs by storing information from working memory into long-term memory in the form of schemas. When learning new information, schemas are brought into working memory to incorporate new information. Schemas that are more robust mean fewer pieces of information need to be handled by working memory. Therefore, learning new information in a familiar domain requires working memory to handle fewer pieces of information expanding the available space to handle information that is new and more complex. Instruction should be carefully designed to limit any unnecessary burden on working memory, and to maximize the potential to construct and automate schemas. Overloading the cognitive resources available to working memory will result in information that is not learned.

Cognitive Load Theory identifies three different sources of cognitive load: intrinsic, extraneous, and germane. Intrinsic load is the amount of load inherent to the content. Content that that requires multiple interacting pieces of information to be processed simultaneously in working memory to learn new material imposes high intrinsic load. For example, simple arithmetic like adding and subtracting numbers would impose minimal intrinsic load because of the relatively low number of interacting elements. However, calculus imposes high intrinsic load because it requires manipulating and executing simple arithmetic as well as additional, higher level mathematical concepts. Intrinsic load can be thought of as the experienced difficulty of the subject matter, and cannot be changed by adjusting instructional elements (Ayres, 2006; de Jong, 2010). Although intrinsic load cannot be changed without adjusting the content, experts experience intrinsic load differently than novices and must be considered when designing instructional elements. Experts in a domain experience less intrinsic load by having a set of complex schemas that can be processed implicitly. This in turn reduces the imposed intrinsic load within the domain by being able to handle more interacting elements in working memory. Being able to handle more information without using a significant amount of cognitive resources allows for more cognitive processing and additional schema creation to occur (Chi, Glazer, & Rees, 1982; Paas et al., 2004).

Extraneous load is the load added by the instructional design that does not contribute to learning and schema construction and may impede learning. For example, displaying text and pictures farther apart rather than close together imposes extraneous load because people need to integrate the two disparate pieces of information, a finding referred to as the spatial-contiguity effect (Chandler & Sweller, 1992; Johnson & Mayer, 2012). Ideally, a sound instructional design will minimize any extraneous load.

Lastly, germane load is the load imposed by constructing and automating new schemas in long-term memory. Whereas extraneous load interferes with learning, germane load directly contributes to it. Instructional designs should stimulate and engage germane processing and schema construction. For instance, prompting students to self-explain the process by which they solved problems promotes inference and germane processing by engaging learners in active learning (Atkinson, Renkl, & Merrill, 2003; Paas, & Van Gog, 2006). The three types of loads are considered to be additive. If the three loads exceed the capacity of working memory, learning becomes impeded (Sweller, van Merriënboer, & Paas, 1998; Paas et al., 2004; Ayres & Paas, 2007). Although Cognitive Load Theory asserts that intrinsic load cannot be changed without adjusting the content, a sound instructional design should manage intrinsic load, reduce the amount of extraneous load, and foster germane load (Mayer, 2005).

Cognitive Theory of Multimedia Learning

The Cognitive Theory of Multimedia Learning extends Cognitive Load Theory by incorporating a model of how learners process information from multimedia. Similar to Cognitive Load Theory, Cognitive Theory of Multimedia Learning also identifies three sources of cognitive processing: essential, extraneous, and generative, which map on to the three loads described by Cognitive Load Theory; intrinsic, extraneous and germane. Mirroring Cognitive Load Theory, the processes are considered additive and must work within the constraints of working memory for deep learning to occur.

Much like Cognitive Load Theory, Cognitive Theory of Multimedia Learning relies on three assumptions of the human cognitive architecture. First, based on dual-coding theory (Clark & Paivio, 1991; Paivio, 1990) and Baddeley's working memory, (Baddeley, 1992; Baddeley & Hitch, 1974), the dual channel assumption posits that the information processing system takes information from two channels, the visual/pictorial for information coming to the eyes, and the auditory/verbal for information entering the ears. The second assumption is that each channel is limited by the amount of information it can handle (e.g., Baddeley 1992, Miller, 1956). Overloading either channel results in information that is not actively processed. Third, learning requires a set of active cognitive processes: selecting relevant words, selecting relevant images, organizing the words into a coherent verbal representation, organizing images into a coherent representation and then integrating the verbal and pictorial information with prior knowledge (Mayer, 2005). Figure 1 provides an illustrated overview of Cognitive Theory of Multimedia Learning. A multimedia presentation presents words and pictures, which are picked up by the eyes and ears in the form of auditory and visual information. The learner then selects the relevant words and images and processes them in working memory by organizing and integrating the verbal and pictorial information.

Both Cognitive Load Theory and Cognitive Theory of Multimedia Learning emphasize reducing the amount of extraneous load in multimedia presentations to allow deeper learning to occur. Using these theories as a framework, interactivity and dynamic multimedia were investigated in the context of learning procedural motor tasks from a tablet computer.

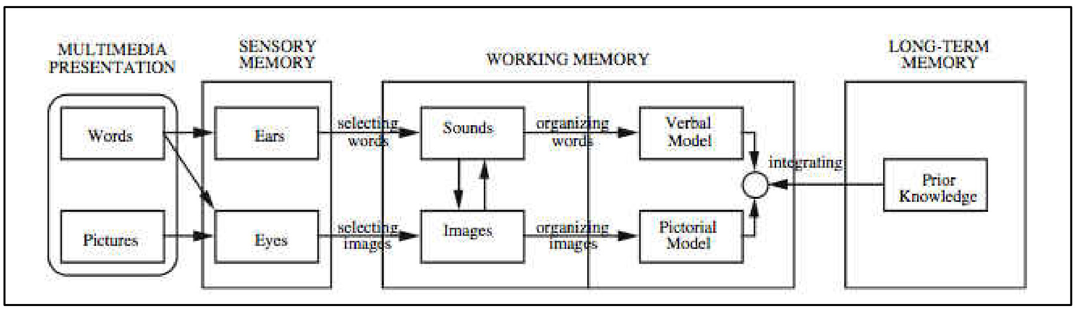


Figure 1. Cognitive Theory of Multimedia Learning (Mayer, 2005)

Dynamic Multimedia

Central to this dissertation is examining the effects of interactivity on dynamic multimedia. However, before discussing interactivity, dynamic multimedia must first be discussed in isolation to interactivity. There is an ongoing debate over whether people learn more deeply from dynamic multimedia or from static images (Betrancourt, 2005; Rieber, 1990; Tversky, Morrison, & Betrancourt, 2002). Multimedia is any presentation that utilizes words (spoken or written) and pictures (Mayer, 2005). Dynamic multimedia refers to any multimedia that depicts motion (or movement), such as animations and video. In contrast, static images are not dynamic and rely on sequences of images and symbols to represent motion. For example, phase diagrams depict transitions from a previous state to a new state in the form of a series of static images. It is not clear under what circumstances one is better than the other. Some argue that dynamic multimedia reduces extraneous load because it provides a more complete visualization, thus reducing the level of abstraction of temporal ideas (Lewalter, 2003). In addition, animations more clearly depict 3D spatial relationships and depth and may not require the use of mental transformations (Strobel, 2010). On the other hand, others argue that dynamic multimedia has the opposite effect and actually increases extraneous load by being potentially distracting (Ayres, Kalyuga, Marcus, & Sweller, 2005; Lowe, 1999; Ploetzner & Lowe, 2004) and by providing transient information. As an animation plays, working memory must constantly incorporate new incoming information while also trying to consolidate the previous information that is no longer present on the screen (Hegarty, 2004; Leahy & Sweller, 2011; Sweller, Ayres, & Kalyuga, 2011).

Some studies have indicated an advantage of dynamic over static multimedia. For example, Kuhl, Scheiter, Gerjets and Edelmann (2011) compared static and dynamic multimedia to teach the physical principles underlying fish locomotion. The dynamic condition consisted of an animated fish undulating in a repetitive fashion. The static condition used nine sequential key frames taken directly from the animation. Their results from knowledge and pictorial recall tests indicated that the dynamic visualizations resulted in higher test performance. They concluded that the animations helped learners develop a better understanding of the dynamic aspect of the material. The dynamic aspect of the animation lessened the cognitive load associated with mentally animating the motion. Similar advantages for dynamic multimedia have been demonstrated in other educational domains including motion trajectory, rate of speed, and the circulatory system (Baek & Layne, 1988; Kaiser, Proffitt, & Anderson, 1985; Large, Behashti, Breuleux & Renaud, 1996). However, these studies used computers to show the effectiveness of animations. No research to date has examined the effects of dynamic multimedia on a tablet device.

Although animations appear to be superior to static images in some studies, other research has indicated no clear advantages to either form of multimedia. Byrne, Catrambone, and Stasko (1999) found that animations were more effective than static images on a post-test when teaching algorithms (i.e., depth-first searches and binomial heaps). Their results showed that for simple algorithms, animations were more effective than static images on post-test performance. However, this effect was attributed to the animations inherently prompting the participants to make predictions about the outcome. The effect disappeared in the second reported study when participants were prompted to make predictions in both the animated and static conditions for a second algorithm (binomial heaps). They concluded that it was prediction and not the animation per se that resulted in the additional learning outcomes in the initial animation condition. Hegarty, Kriz, and Cate (2003, Experiment 1) obtained similar benefits to prediction when comparing animations to static diagrams for learning a mechanical system. The experiment utilized four conditions. In the control condition, participants viewed a single static image diagram. In the prediction condition, participants viewed the static image diagram but then were shown a three-phase diagram and were subsequently asked prediction questions about the system. Participants in the animation condition watched an animation after examining the static image diagram. Lastly, the combination condition had participants view the static image diagram, view the three-phase diagram, answer the prediction questions and watch the animation. Participants who learned the system from an animation did not differ significantly from participants who learned the system using in the prediction condition. However, all three conditions were better than the control condition. They concluded that even though a single static image was inferior to animation, participants were able to infer motion from the phase diagram. Using a process referred to as mental animation, participants used the phase diagram to direct their own mental representation of the system and its movements. In a series of four studies examining various topics including a how a toilet cistern works, lighting formation, ocean waves, and car brakes, Mayer, Hegarty, Mayer and Campbell (2005) compared the transfer and retention effects of two conditions: static media and computer-based animation. Special care was taken to ensure the information contained in both conditions was similar. In each study, the static media condition either out performed or showed no significant difference from the animation condition on measures of recall and transfer. They concluded that static images reduced extraneous processing by not having participants attend to salient, but not relevant motion, associated with animations. They authors did note that animation may be more effective in teaching low spatial ability individuals who may not have the cognitive capacity to mentally animate from static images.

Notwithstanding varied results for using animations and other dynamic multimedia for learning processes and systems, Höffler and Leutner (2007) attempted to reconcile the mixed results in a meta-analysis of 27 primary studies involving dynamic multimedia. Their results found a medium effect size in favor of animations over static diagrams overall. In particular, effect sizes were largest when the animation was representational, realistic, and/or when the task to be learned was procedural-motor.

Later studies investigating animations in the context of learning procedural motor tasks have been in line with the Höffler and Leutner (2007) meta-analysis. For example, Ayres, Marcus, Chan and Qian (2009) demonstrated that animations support the acquisition of procedural tasks. In their study, animations were compared to static images in order to teach two procedural-motor tasks: tying knots and assembling and disassembling ring puzzles. The results indicated that participants in the animation condition performed better for both tying real knots and disassembling real ring puzzles. Wong et al. (2009) also found that animations were more effective than static diagrams when teaching individuals to fold paper in an origami task. These effects were hypothesized to be a result of the organization of sub-systems in working memory. Working memory theories (e.g., dual-coding theory) contain independent processors for the auditory and visual channel (c.f., Baddeley, 1992). By utilizing both channels, working memory capacity can be used more efficiently by not overloading any particular channel. Both Ayers et al. (2009) and Wong et al. (2009) assert that in addition to the visual and auditory sub-processors in working memory (Baddeley, 1992), humans also have a working memory store specifically for movement, which may increase the capacity of working memory specifically for procedural-motor tasks. This movement sub-system allows processing movement information without overloading the auditory and visual sub-system. This is consistent with modality effect of Cognitive Theory of Multimedia Learning and Cognitive Load Theory where deeper learning is optimized when textual elements are presented in an auditory format in conjunction with related visual information (Ginns, 2005). However, an additional sub-processor for movement would extend this theory to include movement as a modality in addition to the auditory and visual channels. Overall, the research suggests that pairing animations with a procedural-motor task, such as the one in this dissertation, benefits learning outcomes.

In the present set of experiments, one goal of Experiment 1 was to examine the effect of dynamic multimedia when learning a procedural motor task. It was expected that dynamic multimedia would result in higher learning outcomes on a transfer of training task than equivalent static image diagrams, consistent with previous research (Wong et al., 2009; Ayres et al., 2009). This experiment also extends the results of Wong et al., (2009) and Ayres et al., (2009) by investigating how interactivity affects procedural motor learning from animations.

Interactive Dynamic Multimedia

Due to the inconsistent findings related to the effectiveness of dynamic multimedia (Tversky, Morrison, & Betrancourt, 2002), recent research has investigated the circumstances under which animations can be effective. One technique found to improve the retention of dynamic multimedia is to add an element of interaction (Mayer & Moreno, 2003; Plass, Homer, & Hayward, 2009). According to Evans and Sabry (2002), interaction is described in a three-stage model of information exchange between the user and an interface. First, the interface invites input from the user; second, the user responds by providing input. Third, the interface provides information that is a direct result of the user input. The interactive element can be as simple segmenting information into bite size chunks (Mayer & Chandler, 2001) or adding stop and start functionality to the interface to allow the user control over the pace of the information

(Hasler, Kersten & Sweller, 2007). Table 1, adapted from Moreno and Mayer (2007) identifies several forms of interactivity that can be implemented into an interface. The current dissertation focused exclusively on controlling and manipulating.

Table 1

Types and Examples of Interactivity.

Type of Interactivity	Description	Example
Dialoguing	Learner receives questions and answers or feedback to his/her input	Seek help from an on-screen agent, click on a hyperlink to get additional information
Controlling	Learner determines pace and/or order of presentation	Use pause/play key or forward (continue) button while watching a narrated animation
Manipulating	Learner sets parameters for a simulation, or zooms in or out, or moves objects around the screen	Using gestures such as tracing one finger across the screen to zoom in on a digital model.
Searching	Learner finds new content material by entering a query, receiving options, and selecting an option	Seek information in an Internet search
Navigating	Learner moves to different content areas by selecting from various available information sources	Click on a menu to move from on Internet page to another

Implementing individual components of interactivity, specifically controlling the pace of a presentation, has been shown to be beneficial for learning from dynamic multimedia. For instance, Schwan and Riempp (2004) had participants learn how to tie nautical knots of various difficulties from either an interactive (i.e., the learner was able to start and stop the presentation) or non-interactive dynamic multimedia presentation (i.e., learner did not control the pace). Participants in the interactive condition were able to learn to tie real knots during a transfer test significantly more quickly than participants in the non-interactive condition. The researchers argued that participants in the interactive condition were able to distribute their cognitive resources in a manner that allowed them to spend more time on the difficult parts of the knot tying instructions, which led to a better performance on the transfer test. In another study, Hasler, Kersten and Sweller (2007) demonstrated the benefits of interaction by having individuals learn about the solar system with or without a stop button for presentation control over an animation. The group that was able to control the pace of the information performed better on a transfer test than those who did not have control over the pace. Similar findings have been reported in other studies testing the effects of learner control using other topics including lightning formation and Newton's laws of motion (Mayer & Chandler, 2001; Rieber, 1990).

In contrast to non-interactive forms of dynamic multimedia, interactivity allows users to adapt the media to work within the constraints of their working memory capacity (Plass, Homer & Hayward, 2009; Schwan & Riempp, 2004). This interaction "reduces the learner's cognitive load on working memory, thereby enabling the learner to progressively build a coherent mental model" (Mayer & Chandler, 2001, p. 390-391). Allowing individuals to control the pace of a presentation reduces the extraneous load associated with the transient information presented by animations. What is unclear, however, is whether there is an additive effect when combining other forms of interactivity with controlling the pace of a presentation that leads to reduced extraneous load and thus better learning outcomes; or if adding multiple types of interaction increases the amount of extraneous load placed on the student.

Utilizing interactivity poses its own challenges because it requires the use of an interface, which in and of itself can be a source of extraneous load (Hegarty, 2004). For example, manipulating the view of an object may add extraneous load, because the user must mentally compare previous states of the object with current states of the object. This effect is referred to as the temporal split-attention effect (see Ayres & Sweller, 2005; Kalyuga, 2007). Additionally, more interactivity requires the user to decide how to use the interactivity, as well as focus the individual's attention on relevant information, resulting in increased extraneous load (Lowe, 2004) and underutilization of the interactivity (Ainsworth, Bibby, & Wood, 2002; de Jong & van Joolingen, 1998; Reigeluth, & Schwartz, 1989). Further, the negative effects of high extraneous load may affect individuals with varying amounts of prior knowledge differently. Content in an area of experts' domain will pose lower intrinsic load on an expert than a novice and therefore experts may not suffer from the effects of high extraneous load posed by the interface (Boucheix & Guignard, 2005; Kriz & Hegarty, 2007; Park, Lee, & Kim, 2009; Sweller & Chandler, 1994;).

Unfortunately, the amount of research systematically investigating varying types of user interaction in dynamic multimedia is relatively small and contradictory, with some showing improvement in learning with the addition of interactive features, and others showing interactivity to interfere with learning. For instance, Kalet et al. (2012) had medical students learn the procedures of an abdominal exam on an online multimedia module, which incorporated three different types of interactivity. In one condition, participants were able to control the pace of the presentation as they watched using a start and stop button. Using a mouse, the other two conditions had participants either clicks on the relevant tools to start an animation sequence or click and drag the tool in manner simulating actual performance in the task. Learning was measured using a multiple-choice post-test addressing procedural knowledge, and a transfer test assessing their ability to perform an abdominal exam on a patient acted out by an experimenter. The results showed no differences on post-test scores between the conditions. However, performance on the transfer test indicated that individuals in the click condition outperformed the other two conditions. They concluded that the additional interactivity associated with the click and drag condition created higher levels of extraneous load by distracting the participants from processing the information, resulting in lower performance on the post-test. On the other hand, the click condition provided sufficient engagement, relative to the animation condition, without causing distraction. Essentially, the high levels of interactivity associated with the interface interfered with processing the actual information, but including slightly more interactivity than the animation condition resulted in higher engagement level that resulted in best transfer task performance. Similar issues with extraneous load and interactivity were found in an experiment by Schnotz, Böckheler, and Grzondziel (1999). Their experiment had participants learn about the simultaneous existence of different daytimes and dates on the earth using either an interactive animation or a static image presentation. Performance was measured with a transfer test and a deeper learning test. Learning from interactive animations resulted in higher learning outcomes in the transfer test, but inferior performance on the deeper learning questions as compared to learning from the static images.

In contrast, others have found presentations with multimedia interactive features to improve learning. For example, Plass et al. (2007; as cited by Plass, Homer, & Hayward, 2009) found that the ability to manipulate content within a simulation depicting the Ideal Gas Law resulted in higher comprehension scores relative to individuals who were only able to control the pace of the presentation. Evans and Gibbons (2007) compared a non-interactive form of multimedia to an interactive multimedia that utilized several types of interaction, including pace control, self-assessment questions, and parameter manipulation. The task was learning how a bicycle pump works. They found that the interactive multimedia condition was superior to the non-interactive multimedia condition on a transfer test. They concluded that the interactive multimedia increased their depth of learning on transfer problems, because the interactivity actively engaged the learner in the learning process. However, because so many forms of interactivity were included in the interactive condition, it is unclear which facet of interaction was beneficial. The difference between the two conditions included several forms of interactivity. The present dissertation attempted to more systematically compare multiple types of interactivity to better shed light on how multiple types of interactivity affect learning a task.

Using a more systematic approach, Wang, Vaughn, and Liu (2011) investigated how different levels of interaction affect learning statistics concepts. In this study, levels of interaction were systematically added across four experimental conditions. The conditions included the control that only received a static multimedia presentation, level 1 which received the static multimedia and a simple animation with pace control, level 2 which received the previous two forms of multimedia plus the ability to manipulate the content in a simulation, and level 3 which received the previous three forms of multimedia plus a practice mode in which they could test their ability in statistics. The authors hypothesized that the addition of increased levels of interactivity would result in progressively higher learning outcomes. However, the results indicated that there were no significant differences between the levels of interactivity, although those in the interactive multimedia conditions performed better on tests of understanding and lower-level applications relative to the control. The authors noted that the complex interface associated with the higher levels of interactivity may have resulted in the underutilization of the features (i.e., too much extraneous load associated with the interface to explore the interactive features).

Based on previous research, the inclusion of interactivity can either aid learning outcomes or inhibit them. The main goal of Experiment1 was to examine whether interactivity enables better learning outcomes in dynamic multimedia. On the one hand, interactivity may be beneficial because it manages the pace of presented information and can provide additional engagement through manipulation of the presented content. On the other hand, interaction may create extraneous load due to an overwhelming interface. In experiment 1, I addressed this question using the interactivity afforded by the touchscreen on tablet devices. The ability to not only control the pace of the information, but also manipulate (e.g., zoom, rotate, and translate) the objects on a screen was compared to less interactive forms of multimedia (i.e., no interaction/control of pace). It was hypothesized that the additional interactivity will foster better learning outcomes on a transfer task than multimedia presentations with less interactivity. In spite of research indicating that interactivity can interfere with learning, it was expected that the interface afforded by tablet computers would not subject participants to additional extraneous load or processing that could interfere with learning due to the intuitive nature of the touchscreen interface.

Embodied Cognition

Experiment 1 relied on the Cognitive Load Theory and Cognitive Theory of Multimedia Learning as the theoretical basis for describing how interactivity and dynamism may affect learning from multimedia. In a similar vein, Experiment 2 leveraged Embodied Cognition as the theoretical basis for how using the touchscreen on a tablet might facilitate learning through gestures afforded by the touchscreen. Although hotly debated (Wilson, 2002), the central tenet of Embodied Cognition posits that learning is grounded in action and that our motor processes and perception are deeply tied to our cognition[1]. The belief that knowledge is grounded in action is in sharp contrast with traditionally held theories that identify the mind as independent of our sensorimotor system and that cognition occurs by manipulating abstract symbols and representations (Zwaan, 1999; Barsalou, 1999). According to more traditional views, incoming sensations are first recoded, or transduced (Barsalou, 1999; 2008; Barsalou, Simmons, Barbey, & Wilson, 2003) into amodal symbols that omit the sensory input (Brachman & Levesque, 2003).

---

[1] It should be noted that the recent literature has identified a distinction between “embodied” cognition and “grounded” cognition mostly having to do with how humans encode abstract concepts (see Borghi, Scorolli, Caligiore, Baldassarre, & Tummolini, 2013; Barsalou, 2008). However, the difference between the two does not fall within the scope of this dissertation. Therefore, grounded cognition and embodied cognition from the literature were used interchangeably.

The symbols are amodal in the sense that their structures do not resemble the perceptual states that produced them and are only arbitrarily connected (Barsalou, 1999). For example, the perceptual states that produced the concept of a “mug” are transduced in the mind as an amodal semantic list that is stored as part of a semantic network of knowledge that include features like “handle” or “ceramic.” This “list” does not refer back to any physical states and exists purely as an abstract cognitive representation. However, this view of cognition has several problems including how transduction occurs, and why, neurologically, semantic knowledge structures are so deeply tied to the sensory-motor regions of the brain (Barsalou, 1999, 2008; Gallese, & Lakoff, 2005).

Embodied Cognition proponents on the other hand, argue that conceptual knowledge retains the sensory information and that concepts and mental representations are integrated (i.e., embodied) with the body’s sensory and perceptual systems (Barsalou, 2008, 1999; Garbarini & Adenzato, 2004; Pezzulo et al., 2011). Although the paradigm of Embodied Cognition is broad in scope (see Wilson, 2002), Experiment 2 focused exclusively on the simulation theory of Embodied Cognition to help explain how gestures could facilitate encoding information. Simulation theory argues that knowledge concepts retain the modality-specific sensory information from which they were formed (Barsalou, 2008). When these concepts are retrieved, multiple areas of activation occur in the brain, including sensory motor areas, creating a mental simulation of the instance in which they were encoded. Under simulation theory and other Embodied Cognition constructs, knowledge concepts are not transduced into amodal symbols, but rather refer directly back to sensorimotor states that encoded the information. To illustrate, the concept of a “mug” brings to mind not only a mental image and definition of a mug, but the actions associated with using the mug (e.g. picking the mug up and drinking from it; Barsalou, 2008). Additionally, abstract concepts are similarly encoded onto concrete concepts of physical experiences (Barsalou & Wiemer-Hastings, 2005; de Koning & Tabbers, 2011). For example, the concept of “scary” is connected to the physical feelings of trembling and breaking out into a sweat. However, abstract encoding is beyond the scope of this dissertation.

Several research studies have offered support of an embodied view of cognition. For instance, Hauk, Johnsrude, and Pulvermüller (2004) showed activation in motor areas in the brain corresponding to action words. That is, when participants read the word “kick,” activation was observed in the motor areas that activate when the leg is moved. Similar results occurred with the words “lick” and “pick” for the tongue and fingers, respectively. Motor processes also have been linked to low-level cognitive functions including spatial ability. For example, Wexler, Kosslyn, and Berthoz (1998) had participants rotate a joystick while performing a mental rotation task. Performance on the task increased (as measured by a decrease in reaction time) when the rotation direction of the joystick and the task matched compared to when the two tasks conflicted, indicating a link between mental rotation and motor processes. Similarly, James, Humphrey, and Goodale (2001) observed that individuals who actively explored novel, 3D objects using a trackball mouse had faster reaction times during a mental rotation task than individuals who passively viewed them. This connection between sensorimotor processes and cognitive encoding was explored in Experiment 2 in the context of learning from a touchscreen interface.

The clearest way to incorporate the sensorimotor system into an instructional design is to require the individual to move their body in the form of gestures. Recent research in Embodied Cognition has demonstrated the importance of gestures in a learning environment. Gestures, broadly defined as movements of the arms and hand (Edwards, 2009), are thought to be a result of sensory motor simulation (Hotstetter & Alibali, 2008). For instance, when individuals gesture when speaking, it is thought to be a result of simulating the concepts about which they are speaking. Working in the other direction, gestures have been shown to help encode information. The effects of using gestures to promote learning and problem solving have been demonstrated in numerous studies. For instance, Cook, Mitchell, and Goldin-Meadow (2009) had 3rd and 4th grade children learn to solve a math problem. Children were required to speak, gesture or both speak and gesture while learning. After learning a few sample problems, the children were tested with a novel problem. In a follow up test four weeks later, the children in the gesture and gesture + speech conditions solved more problems than the children in the speech only condition. Their results suggested that gesture had a causal role in learning.

In another set of experiments, Lozano and Tversky (2006) had participants learn to assemble a piece of furniture. In the first experiment, they examined the effects of gesture by having participants first learn how to assemble the piece of furniture and subsequently make an instructional video about the assembly procedures as if they were teaching someone else. In one condition, participants were required to use only gesture, in another condition participants were required to speak and use gesture. A third, control condition, had participants simply assemble the piece of furniture without gesture or speaking. In a subsequent retest, participants in all conditions reassembled the piece of furniture. Those who used only gestures committed the fewest number of errors during reassembly, followed by individuals in the speech + gesture condition. The control condition committed the greatest number of errors. In a second experiment, participants watched an instructional video describing the assembly procedures of the small piece of furniture in one of two conditions: either with only gestures or only speech. Participants who watched the gesture only video assembled the furniture faster and with fewer errors than the participants who watched the speech only video. Taken together, the two experiments demonstrated that gestures facilitated learning the assembly task above and beyond using speech alone for both learners and communicators. Similar results have been found using various tasks and domains including the Tower of Hanoi puzzle (Beilock & Goldin-Meadow, 2010), gear problems (Alibali, Spencer, Knox, & Kita. 2011), problem solving (Broaders, Cook, Mitchell, & Goldin-Meadow, 2007; Francaviglia, & Servidio, 2011; Cook, Yip, & Goldin-Meadow, 2012; Werner & Raab, 2013), and spatial ability including mental rotation and spatial orientation (Chu & Kita, 2011; Goksun, Goldin-Meadow, Newcombe, & Shipley, 2013; James et al., 2001; Wesp, Hesse, Keutmann, & Wheaton, 2001).

Similar to gestures, physical interactions with the environment also have been shown to facilitate learning. De Koning and Tabbers (2011) further note that, "guided actions such as gestures or object manipulation related to movements can influence cognitive performance (p. 514)." For instance, Ferguson and Hegarty (1995) had participants learn the mechanics of a pulley system by either looking at line drawings or touching a real pulley system. Although all groups demonstrated improvement, those who learned on the real machine were superior in applied problem solving. Being able to touch the pulley system likely contributed to superior learning outcomes.

In addition to real world embodied interactions, the Embodied Cognition paradigm has been migrating into the field of human-computer interaction where "proponents of Embodied Cognition would say that what we are striving for is sensorimotor coupling with an environment (real or virtual) and that computational devices represent tools that mediate this coupling" (Gillespie & Modhrain, 2011, p. 482). The following research studies support the notion that embodied learning, or learning through embodied interaction, including direct manipulation and interaction with virtual objects, leads to better performance and learning outcomes. For example, Zacharia and Olympiou (2011) had participants learn how heat and temperature change work in a physics experiment. Participants learned the topics using a virtual lab and virtual materials, a real lab and real materials, or the control condition that had participants learn the concepts using written instructions and descriptions of the experiment. After taking part in the virtual or real experiment, participants were tested on their knowledge using a conceptual written test. Participants in the virtual group did not differ significantly from the real lab group. However, both groups were significantly better than the control condition. They concluded that touch and manipulation could be real or virtual so long as there was an element of physicality involved. In other words, touching and manipulating the real or virtual equipment contributed to learning. In another study, Akinlofa, Holt and Elyan (2012) had participants learn to disassemble a Lego™ truck using either static images, animations, or an interactive virtual workspace that allowed participants to interact via clicking and dragging components using a mouse. Individuals in the virtual workspace condition showed better accuracy and completion time on a physical Lego truck disassembly task. From an Embodied Cognition perspective, the additional embodiment associated with more active learning lead to the better performance outcomes. The present research combined gestures with simulated interaction in effort to examine how virtually manipulating an object on a screen via gestures using the touchscreen on a tablet. Although gestures have been traditionally viewed as arm waving and hand movements, the current dissertation will see if these types of motions extend to the context of gestures within a tablet. Instead of an individual improvising their own motions, tablet gestures prompt individuals to move their hands in a predetermined way. A primary goal of this dissertation is to determine if this type of gesture promotes learning outcomes in a similar fashion as other types of gestures.

Touchscreens on mobile devices represent great potential to generate sensorimotor coupling for a computer system through gesture interaction. Direct touch input offers the ability to embody gestures that directly correspond to the manipulation of content in multimedia. However, the question remains, does direct touch input and the associated gestures facilitate learning above and beyond the interactive component described in Cognitive Load Theory and Cognitive Theory of Multimedia Learning? Experiment 2 addressed the question by manipulating the level of gesture and type of input used to manipulate the content on the tablet.

Direct Touch vs. Indirect Touch

Another question addressed by Experiment 2 is whether direct touch input (e.g., making direct contact with ones finger on a screen) is more effective for training than indirect input methods such as a mouse or keyboard? Despite the fact that direct touch input (e.g., touching and dragging a finger across a screen) and indirect input (e.g., moving the mouse with your hand to affect the movement of a pointer on a screen) both contain an element of gesture, it seems reasonable to contend that direct touch input offers a more embodied experience due to the direct relationship between the action and reaction. This claim is supported by Jones, Minogue, Tretter, Negishi, and Taylor (2005) who had participants learn about viruses using a mouse, joystick or a 3D virtual probe capable of providing force feedback by providing simulated resistance when "touching" an object. Participants in the virtual probe condition performed significantly better than the other two groups on an assessment questionnaire. In this study, the 3D virtual probe most closely resembled direct touch input compared to the other two forms of input methods. However, very little research has examined direct touch input to indirect input methods, specifically in the context of mobile devices.

Research comparing direct touch to indirect mouse input has focused almost exclusively on speed and accuracy in selecting icons on a screen in order to address user-interface performance (Forlines, Wigdor, Shen & Balakrishnan, 2007; MacKenzie, & Buxton, 1992; Sears & Shneiderman, 1991). Although this line of research is important in the development of efficient user interfaces, it does not address whether or not direct touch input facilitates learning, specifically in the context of procedural tasks. Despite the few studies that exist on the topic, research comparing direct touch input to indirect mouse input indicates it improves spatial memory. In a study by Tan, Pausch, Stefanucci and Proffitt (2002), participants used either a mouse or direct touch input to drag target objects to a specific location on a tabletop screen. They were later tested and asked to recall the location of those objects. Direct touch input condition resulted in a significant 19% increase in accuracy during a spatial memory task. In a similar study, Jetter, Leifert, Gerken, Schubert and Reiterer (2012) investigated whether direct touch input or mouse input resulted in better performance in a spatial memory test when zooming and panning interfaces were included. Participants took part in a similar task to the previous study, but instead were required to pan and zoom the screen to place the objects in the required locations. The locations to which the objects needed to be moved were not visible without either panning or zooming. In the first experiment, participants were only able to translate the screen by panning. Results indicated that for panning, direct touch input performance was significantly better in both the spatial memory test and navigation task (i.e., shortest panning distance). However, this result was not replicated in the second experiment, when the ability to zoom was combined with the ability to pan. In this instance, there were no significant differences between the two groups for spatial memory. This was contrary to the mouse condition that had a wider spread of activity on the screen. This finding was attributed usability issues with the size and layout of the touchscreen. Because the screen surface was so large, participants in the touch condition would limit their interactions to the portion of the screen below the center of the tabletop because it was convenient to reach but also resulted in less movement. Furthermore, zooming out changed the scale factor of the layout which then needed to be integrated with the local view. Taken together, these studies support an embodied view of touch interaction relative to indirect mouse interaction. Although touch did not impact learning during the second experiment of Jetter et al. (2012) this was largely attributed to a sufficiently large tabletop interface. Smaller tablets, such as the one being used in this dissertation, may not likely see the same usability issues as larger interfaces.

The research described pertained directly to spatial memory. To date, no research has specifically looked at how input affects learning a procedural-motor task. As a result of these foundational studies, the proposed research aims to extend our understanding of direct touch input and its effects on learning procedural tasks using Embodied Cognition as a theoretical foundation. It was hypothesized that the additional hand and arm motion associated with tapping the screen support a stronger embodied interaction, which should create a stronger, more robust encoding of the information and therefore result in better learning outcomes than indirect input with few or no gestures.

Spatial Ability

Spatial ability is an individual difference variable shown to improve performance in a variety of fields including video games, aviation, and medicine (e.g. Sims & Mayer, 2001; Dror, Kosslyn, & Waag, 1993, Hedman et al., 2006). It has also been demonstrated that different levels of spatial ability also increase the knowledge gained during multimedia presentations (e.g., Mayer & Sims, 1994; Hegarty, Kriz, & Cate, 2003). In effect, high spatials have a higher working memory capacity for spatial information. Therefore, when handling spatial tasks, more cognitive resources may be devoted to germane and essential processing (Keehner, Hegarty, Cohen, Khooshabeh, & Montello, 2008; Stull, Hegarty, & Mayer, 2009). For example, Mayer and Sims (1994) compared high and low spatial ability students on their ability to learn from a multimedia presentation under two conditions. In one condition, students watched an animation while listening to the narration simultaneously. In the other condition, students were shown the animation and narration successively. Both high and low spatials performed better under the simultaneous condition in a subsequent transfer of training task. This effect became known as the temporal contiguity principle in which individuals learn better from a presentation that combines auditory and visual information simultaneously (Mayer, 2005). Simultaneous presentation of narration and animation requires less extraneous load because individuals do not need to hold the narration in their working memory while waiting for the animation to begin. However, the effect was particularly effective for high spatials. In other words, high spatial individuals benefitted more from the reduction of extraneous load in the simultaneous group than did the low spatials who required more resources to select, organize, and integrate the information. Because of this, spatial ability was considered as a subject variable for both experiments.

Rifle Disassembly

This dissertation approached tablets for training in two ways. Primarily, it addressed a theoretical question regarding how individuals learn procedural motor tasks from a tablet using several theories from cognitive psychology. But in doing so, it also addressed how and whether tablets are a suitable solution to contemporary training issues with special regard to the Army. In the present experiments, the procedural motor task chosen was the disassembly procedure of an M4 carbine, a standard issue rifle in the Army. The disassembly procedures of a Soldier's weapon are one of the most important and fundamental tasks an incoming recruit must learn. These skills are essential in order to correct malfunctions and maintain the rifle and are included in the list of Warrior Tasks and Battle Drill Critical Individual Supporting Tasks (Alley, 2010).

Rifle disassembly procedures are trained as part of Basic Combat Training (BCT)—Introduction to Basic Rifle Marksmanship (BRM). The class covers basic safety, maintenance procedures, function checks, loading and unloading and correcting malfunctions. Current point-of-instruction (POI) is an eight-hour, instructor-led PowerPoint with hands-on instruction. Approximately two hours are dedicated to disassembly and reassembly procedures while students follow along using their rifle.

The current dissertation identified two areas of concern regarding current the current POI. First, not all incoming trainees come in with a similar skill set, nor do all trainees learn at the same pace (Wisher, Sabol, & Ellis, 1999). For instance, some incoming recruits may have previous experience using military equipment and therefore may learn new equipment procedures more quickly than an individual with no previous experience. Second, procedural-based tasks are prone to degradation over time and the rate of skill decay increases as the complexity of the task increases (Wisher, et al., 1999). A task analysis of the disassembly procedures of a standard issue Colt M4 Carbine using the User's Manual for Predicting Military Task Retention (Johnson & Cosby, 1985) indicated that less than half of soldiers in a unit would correctly complete the rifle disassembly procedures after just 10 weeks. In an effort to combat these issues, Wisher et al. (1999) noted several different ways to improve skill retention including optimizing the schedule of refresher training and maximizing original learning by increasing the repetitions and length of training. Unfortunately, this additional training requires more instructors, time, and facilities. Effective mobile applications could provide these additional repetitions without the need for additional, more expensive resources by providing access anytime

Rifle Disassembly Application

The U.S. Army Research Institute (ARI) developed the Rifle Disassembly Application (RDA) with the intent to test the effectiveness of mobile devices for training in response to ALM 2015. The goal of the application was to provide a proof-of-principle for implementing ALM principles for a more learner-centric training environment. In lieu of relying solely on instructor-led classes, Soldiers could practice the necessary skills at a time and place convenient for them using a tablet.

The RDA was designed to train the basic aspects of the Colt M4 Carbine with a focus on the disassembly procedures required for field stripping the weapon. The application contains several training modules that cover topics including names of the components, how the weapon fires, how to clear the weapon, how to disassemble and reassemble, a nomenclature quiz to test knowledge, and a rifle disassembly simulation (RDS). A form of interactive dynamic multimedia, the RDS animates components being removed from the rifle when they are selected in the proper order. In contrast with non-interactive multimedia (e.g., animation), which only allows for controlling pace, the RDS requires more interactivity and participation from the user in order to go through the disassembly procedures of the rifle. The RDS used in the application is of particular interest to the current dissertation in order to address how effective the interactivity afforded by tablets is for training procedural-motor tasks compared to other forms of non-interactive, dynamic, and static multimedia (e.g. static diagrams and interactive animations). The

RDS draws on existing types of interactive multimedia to utilize specific interactions that should foster learning the procedures required for disassembly. Specifically, the simulation is a type of multimedia that allows users two types of interactivity identified by Moreno and Mayer (2007). First, users have the ability to manipulate the content (e.g. rotate and zoom). Second, users can control the pace of the content.

Current Studies

The challenge in designing interactive multimedia is promoting behavioral activity without creating excessive extraneous load will otherwise interfere with cognitive activity (Moreno & Mayer, 2007). Although the literature is rich with research addressing the role and effectiveness of interaction in multimedia, these studies generally address conceptual learning (e.g., how pressure affects heat in the Ideal Gas Law) and utilize only one form of interactivity. The RDA developed by ARI focuses exclusively on a procedural motor task, which has received less attention in the multimedia literature. Therefore, the current dissertation examined whether dynamic multimedia, with higher levels of interaction, will foster better learning outcomes than other forms of multimedia using Cognitive Load Theory and Cognitive Theory of Multimedia Learning as the theoretical bases. Furthermore, the touch interface unique to mobile devices also was examined within the context of Embodied Cognition, with the aim of addressing how direct touch input and gestures contribute to learning from interactive multimedia. Although gestures have been examined as instances of improvised movement of the arm and hands, the research described here will examine them in the context of prompted movement. Similarly, learning from a touchscreen has been examined in the context of spatial memory. The research described in the current dissertation will investigate learning from direct touch in a complex procedural-motor task.

For both experiments, learning outcomes were measured in two ways. First, a recall test measured retention of the information. Secondly, participants were tested in their ability to use the information learned from the multimedia presentation in order to disassemble and reassemble a real rifle. I refer to this as a transfer of training from multimedia to real-world application. This in a similar vein as Wong et al., (2009) and Ayers et al. (2009), who had participants learn a procedural-motor task (origami folding and knot tying respectively), and tested the ability of the multimedia training to transfer to a real world version of the same task.

The research questions were addressed in two Experiments. Experiment 1 examined the effectiveness of interactive dynamic multimedia to train a disassembly task by comparing it to other forms of interactive and non-interactive multimedia. In doing so, the Experiment 1 looked specifically at how dynamic multimedia compared to static media as well as investigated how introducing varying levels of interactivity afforded by tablets compared to less interactive media on measures of recall and transfer. In the Experiment 2, the touch UI afforded by mobile devices was isolated and compared to indirect mouse input to isolate a potential reason why (or why not) dynamic interactive multimedia is effective. To accomplish this, gesture interfaces were compared to non-gesture interfaces using either a mouse or touch to manipulate the content. Learning outcomes were once again measured in terms of recall and transfer task performance. Together, the two research studies will not only provide guidelines for using mobile devices for training, but also provide further insight into Cognitive Load Theory and Embodied Cognition.

CHAPTER TWO: EXPERIMENT 1

Experiment 1 addressed two primary research questions. First, is interactive dynamic multimedia more effective at training disassembly tasks than other forms of less interactive and dynamic multimedia such as 2D static images, non-interactive animations, or interactive animations? Secondly, is dynamic multimedia more effective at training a procedural-motor task than equivalent static multimedia? In this experiment, the impact of different levels of interactivity and dynamics on learning from a multimedia tablet application were investigated. The trained task was disassembling a Colt M4 Carbine. Participants took part in one of four different multimedia training conditions: a Rifle Disassembly Simulation (RDS), interactive animation, non-interactive animation or static phase diagrams. Based on evidence suggesting learning procedural motor tasks from dynamic multimedia is beneficial over static multimedia, it was hypothesized that the dynamic multimedia will be a superior training tool than static multimedia. Further, it is believed that the added interactivity in the RDS (controlling and manipulating) will lead to reduced extraneous load (controlling) and increased germane load (manipulating). In order to assess performance, two measures of learning outcomes were used, recall and transfer of training or a real M4. Recall is the ability to reproduce learned material and is a measure of how much information was remembered. However, reproducing information does not necessarily indicate a deep understanding of the information. To this end, participants were tested on two transfer of training tasks, disassembling and re-assembling a real M4 carbine. Transfer of training in the present research refers directly to the ability of the multimedia training to carry over to a real world task. The ability to disassemble the rifle from the training would indicate a sound understanding of the material in a novel situation (i.e., digital model to real world task). Re-assembling the rifle, would require a deeper understanding and more developed schema construction for two primary reasons. For one, the information was not explicitly taught. Second, in contrast with the disassembly test, during the re-assembly test it was possible to complete a step incorrectly adding to the difficulty of the test.

Experimental Hypotheses

Based on the existing literature and the specific research questions posed by this experiment, the following hypotheses were proposed:

Hypothesis 1

Participants in the interactive conditions will outperform groups in the non-interactive conditions in the recall test and the two transfer of training tasks.

The interactive animation condition will benefit from the segmentation principle which states that extraneous load due to transient information is reduced by allowing the user to reconcile the information in small chunks before moving on (Mayer & Chandler, 2001). The RDS condition will benefit from the segmentation principle as well as the additional benefits of being able to manipulate (e.g., zooming and moving the virtual rifle) by alleviating working memory associated with performing mental animations. Furthermore, extraneous load associated with deciding which portion of the animation to direct one's attention may be reduced because users must actively decide the components with which to interact based on the narrated instruction thus directing their attention to the relevant portion of the screen. This will result in more germane processing which will be exhibited in the recall and transfer of training task performance.

Based on Hypothesis 1, several specific predictions have been developed. Individuals in the RDS and interactive animation conditions will (1) remember more steps on the recall test, (2) perform more disassembly steps in a faster time, and (3) perform more reassembly steps in a faster time than the non-interactive animations and 2D phase diagram conditions. Furthermore, because of the additional interactivity, the RDS condition will (4) remember more steps on the recall test, (5) perform more disassembly steps, more quickly, and (6) perform more reassembly steps more.

Hypothesis 2

Participants in the dynamic conditions will outperform the 2D phase diagram condition in the recall test and the two transfer of training tasks.

It was hypothesized that the individuals in the animated conditions (RDS, interactive animation, and non-interactive animation) would learn more from the multimedia training than the individuals in the 2D phase diagram condition. Hypothesis 2 is based on theoretical research suggesting that the ability to see transformations in space will help free up cognitive load associated with mental transformations (Hegarty, 2004). These mental transformations are a central element to learning how the rifle components fit together. Furthermore, research also has indicated dynamic multimedia to be especially effective when teaching a procedural motor task (e.g., Höffler and Leutner, 2007). Because the trained task is procedural motor, participants should benefit from the dynamic multimedia.

Based on Hypothesis 2, three specific predictions have been developed. Individuals in the animated conditions will (1) remember more steps on the recall test, (2) perform more disassembly steps in a faster time, and (3) perform more reassembly steps in a faster time than individuals in the 2D phase diagram condition.

Hypothesis 3

Spatial ability will interact with the RDS condition and 2D phase diagram conditions, such that high spatials will demonstrate better learning outcomes relative to low spatials in the RDS condition, and low spatials will show decrements relative to high spatials in the 2D phase diagram condition.

Hypothesis 3 is based research indicating that high spatials have a higher capacity to handle spatial concepts when learning from multimedia (e.g., Stull et al., 2009). High spatials will experience less imposed intrinsic load associated with the spatial aspects of the task because of their ability to more efficiently process spatial information. Less imposed intrinsic load allows more cognitive resources to handle higher levels of extraneous load or the ability to direct resources to germane load. The 2D phase diagram condition provides the least amount of instructional design elements intended to help learn procedural motor tasks: the pace of the presentation cannot be controlled, and there are no animations to help alleviate load associated with mental animations. Therefore, in the 2D phase diagram, low spatials could experience cognitive overload while high spatials should be able to handle more of the imposed intrinsic spatial load, resulting in less of a performance decrement relative to the low spatials. Furthermore, because high spatials should not be as affected by the spatial intrinsic load, they should be able to handle that elevated levels of interactivity associated with the RDS condition. Low spatials on the other hand, although being helped by the instructional elements in the RDS (i.e., controlling and manipulating), may still suffer cognitive overload due to the extraneous processing required by using all of the interactive elements.

CHAPTER THREE: EXPERIMENT 1 METHOD

Participants

One hundred sixteen college students (62 males, 54 females) between the ages of 18-24 (*M*=19.3, *SD*=1.6) were recruited for the study using the UCF Psychology Department's online recruitment tool and received class credit for their participation. All participants were over the age of 18 at the time of the experiment. Participants included in this sample were inexperienced with weapon disassembly procedures as measured by a prior knowledge questionnaire. Because rifles frequently require similar disassembly procedures, naïve subjects were essential to ascertain what knowledge came directly from the training multimedia. The seven individuals who recorded a score two standard deviations above the mean on the rifle experience questionnaire were omitted from data analysis, resulting in 109 analyzed cases (55 Males, 54 Females). Participants were randomly assigned to one of the four training conditions: phase diagram, non-interactive animation, interactive animation, and interactive simulation.

Design

This experiment tested the effectiveness of interactive dynamic multimedia by comparing four different multimedia training conditions with varying amounts of interactivity and dynamism. To this end, a 4 (multimedia training condition) x 2 (spatial ability: high, low) between-subjects design was used. With the exception of the RDS, the three other multimedia conditions were created using screen captures and screen recordings from the interactive animation conditions to ensure information equivalence across the other conditions. Each multimedia training condition walked through the fifteen steps required to disassemble the M4 Carbine. Additionally, all conditions contained the identical verbal instructions for each step of the process. The narrations for each step are listed in Table 2. Redundant text instructions were omitted due to the high extraneous load they impose on users watching dynamic multimedia (see Mayer & Johnson, 2008; Marraffino & Johnson, 2014). Table 3 illustrates the varying amounts of interactivity and dynamism in each condition. The left column indicates whether the condition is dynamic or not, the right column indicates whether and what type of interactivity is utilized.

Participants went through their assigned training twice. In each condition, participants were instructed to not move back or repeat any of the steps during the training session to ensure everyone received the same amount of training. The experimenter watching the training session enforced this.

Multimedia Training Conditions

Each training condition contained varying amounts of dynamism and interactivity from high levels of interactivity and dynamism found in the Rifle Disassembly Simulation, to no interactivity or dynamism as found in the 2D phase diagram condition.

Rifle Disassembly Simulation

The RDS contained the highest level of interactivity (i.e., including controlling and manipulating) as well as incorporated animation into its presentation. After listening to a narrated instruction, participants tapped on the relevant component to highlight it. Tapping on the component again displayed an animation sequence of the component being removed from the rifle. Once the component was removed, the next step in the disassembly process was presented. In addition to selecting components, participants also had access to panning (sliding two fingers across the screen), rotating (sliding one finger across the screen) and zooming (pinching and separating the fingers) functions. Figure 2 shows a screenshot of the RDS.

Figure 2. Screenshot of RDS.

Interactive Animation

The interactive animation contained only a single element of interactivity, which was being able to control the pace of the information. Instead, verbal instructions and associated animations were segmented into the fifteen discrete steps. After watching and listening to a step in the disassembly process, participants were able to click the next button at their own pace to move forward. However, unlike the RDS, participants were unable to interact with the model rifle on the screen. Instead, pre-determined viewpoints were utilized. Figure 3 contains a screenshot taken from the interactive animation condition.

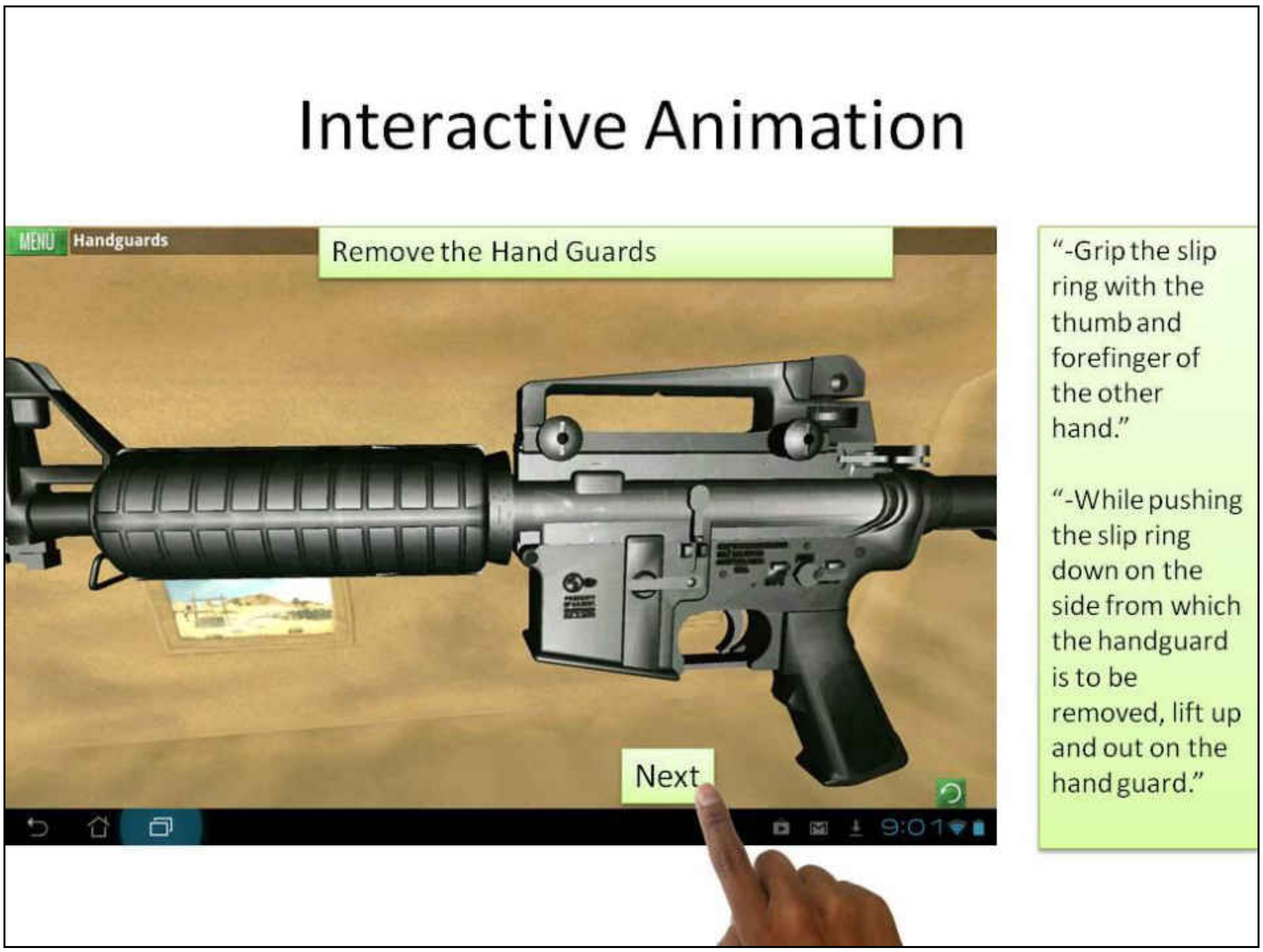


Figure 3. Screen shot of the Interactive Animation.

Non-Interactive Animation

The non-interactive animation was identical to the interactive animation except participants were not able to control the pace of the animation. The animation and verbal instructions played all the way through without stopping. Figure 4 is a screen shot of the non-interactive animation condition.

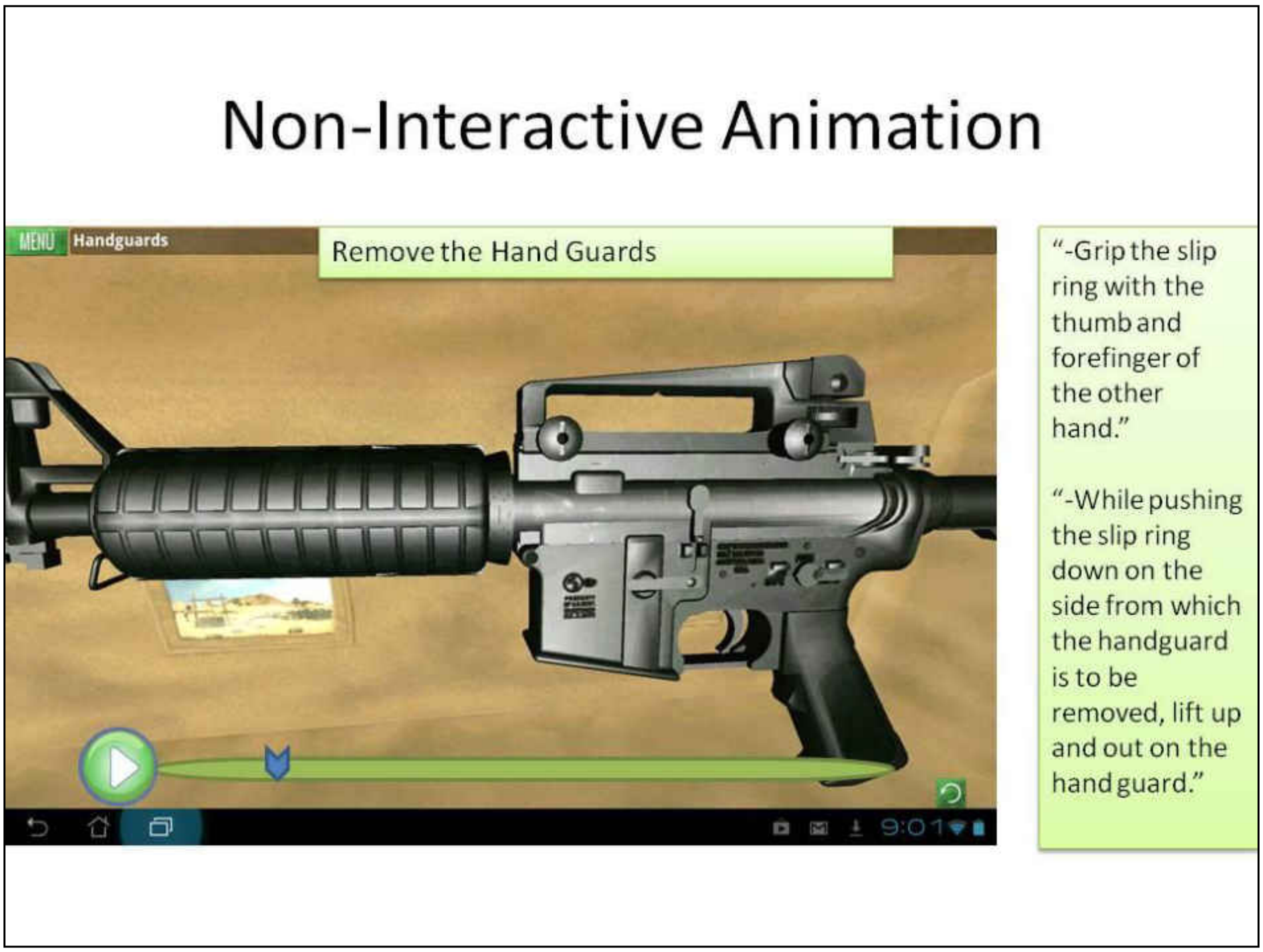


Figure 4. Screenshot of the Non-Interactive Animation.

2D phase diagrams

The 2D phase diagrams contained no elements of interactivity or dynamic animations. Instead, two images presented a *before* and *after* view of the current step. The *before* and *after* images that composed the phase diagram were screenshots taken directly from the non-interactive condition for each step. This condition was presented using Microsoft PowerPoint for Android. The presentation moved at a predefined pace in time with the narration. Figure 5 contains a screenshot of the 2D phase diagrams that were presented using PowerPoint.

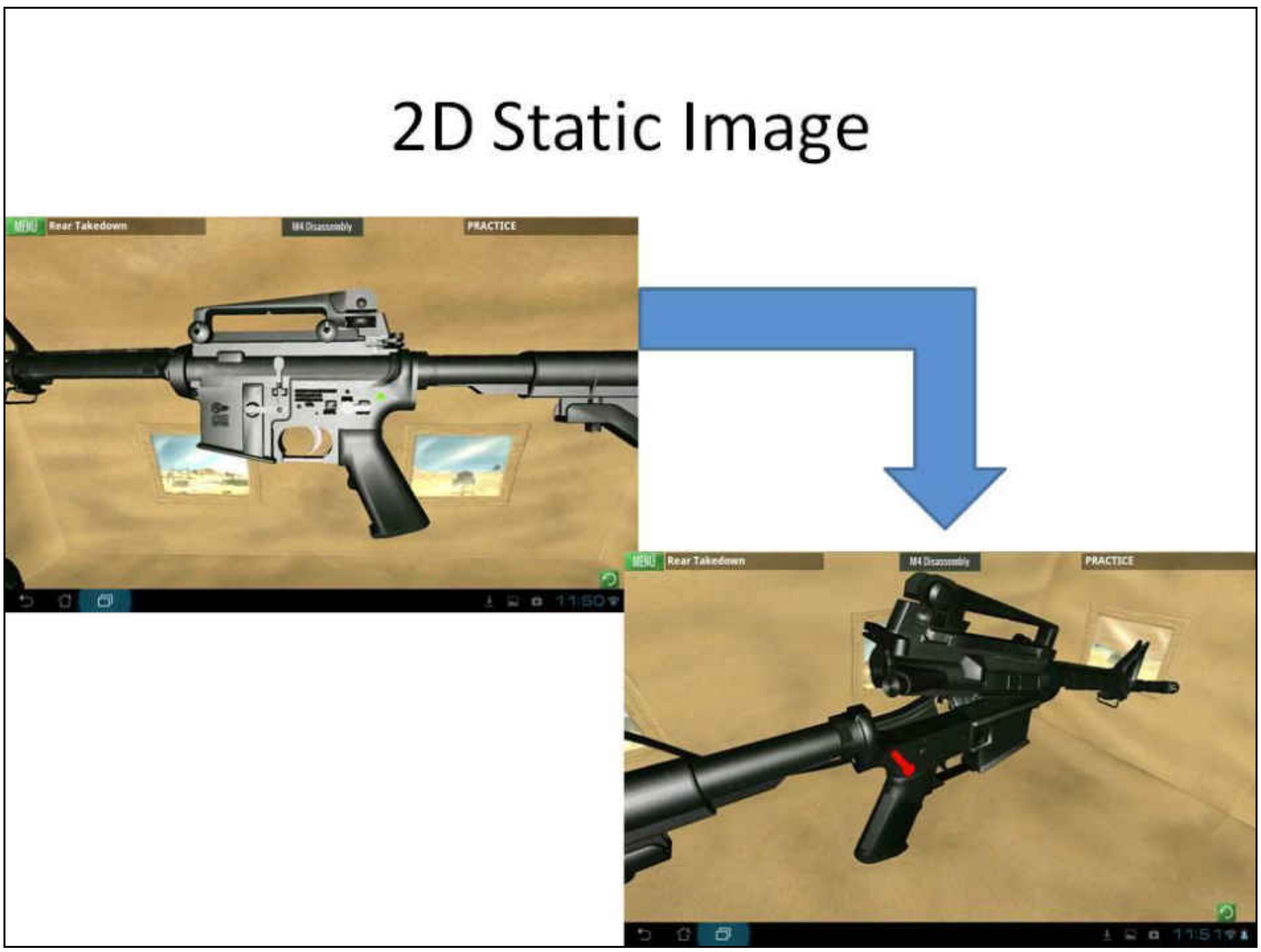


Figure 5. Screenshot of the 2D phase diagram.

Table 2

Fifteen steps with associated multimedia narration for the M4 disassembly.

Step	Instruction
1	"Clear the weapon"
2	"Remove sling. Disconnect the sling at the swivels on both ends."
3	"Remove hand guards. Grip the slip ring with the thumb and forefinger of the other hand. While pushing the slip ring down on the side from which the hand guard is to be removed, lift up and out on the hand guard. Caution, "do not use a screwdriver or any other tool when removing hand guards. Doing so may damage the hand guards, slip ring or both.
4	"Remove take down pins. Using a rifle round, push takedown pin as far as it will go."
5	"Push receiver pivot pin as far as it will go."
6	"Separate upper and lower receivers"
7	"Remove carrying handle. Loosen the round nuts on the left side of the carrying handle. Do not fully remove the round nuts from the threaded stud."
8	"Pull back charging handle and bolt carrier. Pull back on charging handle 5 to 7 cm while pressing the charging handle latch."
9	"Remove bolt carrier and bolt. Grasp the bolt carrier and pull it from the receiver."
10	"Remove charging handle. After the bolt carrier is removed, the charging handle will fall free of its groove in the receiver when pulled to the rear."
11	"Remove firing pin retaining pin. Press out the firing pin retaining pin by using the nose of a cartridge or similar pointed object."
12	"Push in bolt assembly to locked position. Rotate the bolt until the cam pin is clear of the bolt carrier."
13	Drop firing pin out of rear of bolt carrier. Elevate the front of the bolt carrier and allow the firing pin to drop from its well in the bolt."
14	"Remove bolt cam pin. Rotate cam pin 90 degrees (1/4 turn) and lift if out of the well in the bolt and bolt carrier."
15	"Remove bolt assembly from carrier. The bolt can be removed easily from its recess in the bolt carrier."

Table 3

Relative levels of dynamism and interactivity in each condition.

	Dynamic	Type of Interactivity
RDS	✔	Controlling Manipulating
Interactive Animation	✔	Controlling
Non-Interactive Animation	✔	
2D Phase Diagram		

Apparatus

Tablet

The experiment was conducted using an Asus Transformer Infinity tablet. The tablet had a 10.1-inch screen with 1920 x 1200 resolution using the Android operating system. For Experiment 1, the device only supported direct touch input.

Camera

A digital, hand-held video recorder situated on a tripod was used to capture the disassembly, re-assembly and tablet interaction.

Survey Administration

All surveys were administered on a desktop computer using the Qualtrics® survey creation website.

Replica Rifle

A replica Colt M4 carbine for training was used for the transfer of training task. The replica carbine was identical to a real carbine with two primary differences. First, for safety reasons, the model carbine's bullet chamber and bolt carrier assembly were machined so that the model could not fire any ammunition whatsoever. Second, the model's bolt carrier assembly was unable to be disassembled. To get around this, the model's bolt carrier assembly was switched out for a real bolt carrier assembly when the participant reached a point to disassemble the bolt carrier. Subsequently, the real bolt carrier assembly was swapped out for the model bolt carrier assembly once it had been reassembled and was ready to be placed back into the rifle.

Materials

Demographics

A demographics survey was administered to collect data regarding the participant's age, education, and previous experience using mobile devices and tablets. These data were primarily used to explain any oddities that may have occurred during data analysis and to ensure equivalence across conditions. The demographics survey can be found in Appendix A.

Nomenclature

In order to ensure participants started with the same basic knowledge, participants first learned the names of the major rifle components with which they were to interact (e.g., bolt carrier, firing pin, lower receiver). On the tablet, participants were shown an exploded view of the all the components. Clicking on each component would narrate the name of the component. Participants were instructed to use, as much time as they needed to familiarize themselves with the components and that a 100% was required on a follow-up quiz before they could move forward in the study. During the quiz, participants were shown a rifle diagram with blanks next to the components and they had to choose the correct component name from a drop-down box. A screenshot of the nomenclature quiz can be found in Appendix B.

Previous Firearm Experience

An important facet of Cognitive Load Theory is how expertise in a domain affects the amount of intrinsic load a task places on the individual. Experts in a domain have more robust schemas that can be accessed automatically without a detriment to workload. Similarly, novices will find the same content to have increased intrinsic load (Sweller & Chandler, 1994). Therefore, firearm experience was controlled for using a questionnaire ascertaining previous experience handling firearms with an emphasis on disassembly and assembly will be administered. The questionnaire asked a series of yes/no questions regarding firearm experience (e.g., Do you have a concealed weapons permit?). The total number of “yes” responses was summed and recorded. Additionally, the survey asked participants how comfortable they were handling the replica rifle on a 7-point likert scale anchored with "1-not comfortable at all," and "7-very comfortable." The previous firearm experience survey can be found in Appendix C.

Spatial ability

Two measures of spatial ability were used: the Card Rotations Test and the Paper-Folding Test (Ekstrom, French, Harman, & Dermen, 1976). The Card Rotations Test instructed participants to identify whether shapes were the same (rotated) or different (rotated and mirrored). Participants had three minutes to complete as many of the 80 problems as quickly as possible without sacrificing any accuracy. Appendix D contains a copy of the Card Rotations Test.

The Paper Folding Test contained 10 items in which participants were shown a series of paper folds in which at the end, a hole was punched through. Participants had three minutes to determine where the holes would be for each of the ten problems, after the paper was unfolded. These two tests were chosen because they load on to two separate spatial factors inherent to disassembly and assembly tasks (Carroll, 1993). Specifically, the Paper Folding Test required elements of spatial working memory while the Card Rotations Test measures the ability to perform mental rotations. Appendix E contains a copy of the Paper Folding Test.

Subjective Measures

Workload and Cognitive Load

In order to infer cognitive load, two measures of workload were used, the NASA Task Load Index (TLX; Hart & Staveland, 1986) and a single item measure of cognitive load. The

NASA TLX is a six-factor subjective workload scale that measures mental demand, physical demand, temporal demand, performance, effort, and frustration and is used extensively in the human factors literature as a subjective measure of overall workload (e.g., Brill, Mouloua, & Gilson, 2008). Users score their demand by sliding a slider bar to either side of two anchors ("Very Low" and "Very High") scored from 0-100 in five point increments. The NASA TLX can be found in Appendix F.

The measure of cognitive load asked participants how difficult a given task was at three points in the study: after the training phase, after the disassembly task and after the reassembly task. This type of subjective measure of cognitive workload is frequently used in the cognitive load literature (e.g. Kalyuga, Chandler & Sweller, 2000, 2001; Mayer & Chandler, 2001; Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Paas & van Merriënboer, 1994). The measure of cognitive load was administered after the training phase, disassembly task and reassembly task. All administrations of the measure used a 1-7 scale with the following anchors: very easy, somewhat easy, slightly easy, neutral, slightly difficult, somewhat difficult, or very difficult. For the training phase, the question asked, "How difficult was it for you to learn to about rifle disassembly from the presentation you just saw?" For the disassembly and reassembly tasks, participants were asked, "How difficult was it for you to perform this task?"

Usability

In an effort to infer any increases (or decreases) in extraneous load placed on the user by the multimedia interface, a measure of usability was administered. In conjunction with the measure of cognitive load, it could be inferred that if performance and usability are low in a condition, that the interface may to be blame for an increase in extraneous load. To measure usability, an adapted version of the IBM Computer Usability Satisfaction Questionnaire was utilized (Lewis, 1995). Agreement with statements is measured on a 7-point scale with "Strongly Agree" and "Strongly Disagree" as anchors. Example statements include "The information provided with the system is easy to understand," and "It is easy to find the information I need." The Usability Satisfaction Questionnaire can be found in Appendix G.

Motivation

One construct that has been shown to be related to germane load is motivation (Rieber, 1991; Paas, Renkl, & Sweller, 2003; Moreno & Mayer, 2007). In addition to affecting extraneous load, the multimedia interface, with the added interactivity, may increase motivation to learn and therefore promote germane processing. In order to examine differences in motivation, selected factors (Interest/Enjoyment, Perceived Competence, and Effort/Importance) from the Intrinsic Motivation Inventory (IMI) were administered to participants. The survey has been validated and used in several experiments related to intrinsic motivation and self-regulation (e.g. Ryan, 1982, Ryan, Mims, & Koestner, 1983; McAuley, Duncan, & Tammen, 1989) The IMI had participants rate how true various statements were reflective of them using a 1-7 scale with anchors at 1 (not at all true), 4 (somewhat true), and 7 (very true). Example statements include "I enjoyed doing this activity very much," and "I put a lot of effort into this." The IMI can be found in Appendix H.

Performance Measures

Performance was measured in two ways, declarative knowledge (measured in recall test performance) and transfer of training. Training transfer was assessed by having participants disassemble a physical rifle replica (trained task) as well as having participants reassemble the rifle (untrained task).

Recall Test

Participants were instructed to list, in as much detail as possible, as many of the fifteen-disassembly steps they could remember from the training session. They were given five minutes to type their answer. The total number of correctly identified steps was coded and recorded. Appendix I contains the specific coding instructions for the recall test.

Transfer of Training

Transfer of training was measured in two ways, disassembly and reassembly. Disassembly measured the knowledge directly obtained from the multimedia condition and represented a near transfer of training test. Reassembly on the other hand represented a far transfer of training because it was not explicitly taught and required a deeper understanding of the disassembly process in order to successfully apply the knowledge in a different sequence.

For the disassembly task, the replica rifle was placed in front of participants where they were instructed to disassemble the rifle in two stages. In the first stage, they were to disassemble the exterior of the rifle up to and including the step where the bolt carrier was removed. In the second stage, they disassembled the interior bolt carrier assembly. This separation of segments was done in lieu of providing feedback to the participant if they became “stuck” at any one step in the sequence. Separating the disassembly procedures into two stages allowed participants the opportunity to attempt the full range of disassembly procedures even if they were unable to complete certain steps. For example, if a participant were unable to remove the take down pin and separate the upper receiver from the lower receiver, they would not have the opportunity to demonstrate their ability to disassemble the bolt carrier. In this case, not knowing one step would preclude a participant from demonstrating their knowledge of several other steps. Participants were instructed to work as quickly as possible. Five minutes were given to complete each stage, and the clock was stopped when either time ran out or the participant indicated they were finished. The experimenter recorded the total number of steps correctly completed and the total time for completion for both stages of disassembly. Similar to the disassembly test, re-assembly occurred in two stages. Once again, participants had five minutes and were instructed to work as quickly as possible.

The step “clear the rifle” was omitted during the transfer of training procedures because the procedure to remove the magazine on the replica rifle was not the same as the procedure listed in the tablet training. This resulted in fourteen total scored steps for disassembly and fourteen steps for reassembly. The experimenter recorded the time to complete and number of steps correctly completed. Appendix J contains the steps for disassembly and reassembly as coded by the experimenter. The total number of completed steps and total number of correct steps in sequence was summed and recorded for both disassembly and reassembly.

Procedure

Participants were run in individual sessions lasting approximately 1.5 hours. After obtaining consent, participants first learned the essential component names before being tested to 100% proficiency on the nomenclature quiz. This was done to ensure all participants began on an equal footing and could better understand the multimedia presentations.

Following nomenclature quiz, participants were randomly assigned to one of the four types of training media (RDS, non-interactive animation, interactive animation or 2D static image). All training conditions took place on the Asus tablet to control for screen size and resolution. Participants went through the training condition twice. With the exception of the non-interactive conditions (i.e., 2D phase diagram and non-interactive animation), participants moved through the training at their own pace. The non-interactive conditions took a similar amount of time as the interactive conditions. The total training time for completing two training conditions was approximately 12 minutes for all participants. After the training phase, participants filled out the NASA TLX, measure of cognitive workload, usability questionnaire and the IMI. Filling out the questionnaires also served to clear working memory before moving on to the testing phase. Afterwards, participants took the recall test followed by the Paper Folding Test.

Following the recall test, participants took part in the disassembly task. They were handed the fully assembled and cleared replica to disassemble. Participants were instructed to disassemble the rifle in two stages (exterior followed by interior) in the same sequence as described in the training conditions. To ensure that participants were not cued to any steps during the disassembly, the experimenter completed any steps not performed outside the view of the participant. Afterwards, the participant filled out the NASA TLX and measure of cognitive workload before taking the Card Rotations Test.

Following the Card Rotations Test, the participant was handed the fully disassembled rifle and instructed to reassemble it. Similar to the disassembly, this process was also divided into two stages; the exterior and interior bolt assembly. Finally, participants filled out the demographic and rifle experience questionnaires.

CHAPTER FOUR: EXPERIMENT 1 RESULTS

Analyses were performed using SPSS 20 for Windows. Unless otherwise stated, an alpha level of .05 was used for all analyses.

Preliminary Analysis

Several preliminary analyses and variable coding were conducted prior examining the research question and hypotheses. First, an outlier analysis was conducted to remove any individuals with sufficiently high previous experience with firearms. To accomplish this, the rifle experience questionnaire was summed and recorded. A higher score indicated a higher level of experience with firearms. The mean score was 1.13 with a standard deviation of 1.47. Individuals scoring higher than two standard deviations from the mean were removed from analysis (9 total) leaving 109 analyzed cases.

An one-way ANOVA was conducted to ensure equality across each condition for a variety of demographic measures including spatial ability, video-game experience and tablet ownership. No significant differences were found.

A correlation analysis was conducted between the two measures of spatial ability (Card Rotations Test and Paper Folding Test). The two measures were found to be significantly correlated, $r(116) = .278$, $p=.003$. Therefore these variables were standardized and combined into a single measure of spatial ability. A median split was conducted separating spatial ability into high and low spatial ability groups. A follow up t-test was conducted between the high and low groups to ensure differences between the two. The t-test yielded a significant difference between the two groups, $t(107) = 14.69$, $p < .001$, $d = 2.84$. The high group ($M = 1.29$, $SD = .746$) had a higher measured spatial ability and the low group ($M = -1.23$, $SD = 1.01$). Table 4 shows the breakdown of high and low spatial ability by condition.

Table 4

Number of participants in each condition by spatial ability.

	Spatial Ability		
Condition	High	Low	Total
2D Phase Diagram	17	10	27
Non-Interactive Animation	15	13	28
Interactive Animation	10	16	26
RDS	14	14	28
Total	56	53	109

For the recall test, two raters independently coded the participant responses and awarded a point for every step that could be clearly identified. The total number of steps was summed and recorded. To verify the reliability of the scoring, the two raters' scores were correlated resulting in a Pearson's correlation coefficient of .880 indicating a very high overall agreement between the raters.

For the disassembly and reassembly tasks, the number of completed steps for each stage was combined and summed. Similarly, the time to complete each stage for reassembly was summed and recorded in seconds.

Primary Analysis

In order to assess whether spatial ability and the multimedia training conditions affected performance on the recall test and two transfer of training tasks, separate 4 (multimedia training condition) x 2 (spatial ability) ANOVAs were conducted for each. Similar ANOVAs were also conducted for each of the subjective measures.

Performance Measures

Recall Test

In order to assess the effects of the multimedia training conditions and spatial ability on recall, a 4 (multimedia training condition) x 2 (spatial ability) ANOVA was conducted. Effect sizes, means, and standard deviations were reported.

Using the number of steps correctly listed during the recall test, a statistically significant main effect was found for spatial ability, $F(1,101) = 11.66$, $p = .001$, $\eta_p^2 = .103$, such that high spatials were able to recall more steps ($M = 10.31$, $SD = 2.63$) than low spatials ($M = 8.38$, $SD = 2.95$). There was no significant main effect for training condition in spite of a moderate effect size, $F(1,101) = 1.674$, $p = .177$ $\eta_p^2 = .047$, nor a significant condition by spatial ability interaction, $F(1,101) = .815$, $p = .489$, $\eta_p^2 = .024$.

In terms of recall test performance, the lack of a main effect for condition did not lend support to either Hypothesis 1 or Hypothesis 2 in spite of a moderate effect size for condition. Hypothesis 1 predicted that the interactive conditions (interactive animation and RDS) would recall more steps than the non-interactive conditions (phase diagram, and non-interactive animation). Although the interactive conditions had higher mean scores than the non-interactive conditions, they were unable to reach statistical significance. Hypothesis 2 predicted that individuals in the animated conditions (non-interactive animation, interactive animation, and

RDS) would outperform those in the phase diagram condition. Even though the mean number of steps recalled for each of the animated conditions, as seen in Table 5, was higher than the phase diagram condition, statistical significance was not achieved. Hypothesis 3 predicted an interaction between spatial ability and condition. In spite of an apparent interaction seen in Figure 6, statistical significant was not achieved. However, because of the specific a priori interaction hypothesis, an exploratory t-test was conducted using a Bonferonni correction resulting in an adjusted p-value of .013. A statistically significant difference was found for spatial ability in the RDS condition, $t(26) = 3.32$, $p = .003$, $d = 1.30$. High spatials ($M = 11.7$, $SD = 2.59$) in the RDS condition recalled more steps than low spatials ($M = 8.64$, $SD = 2.59$). This result was in support of Hypothesis 3, which predicted that high spatials in the RDS condition would outperform low spatials.

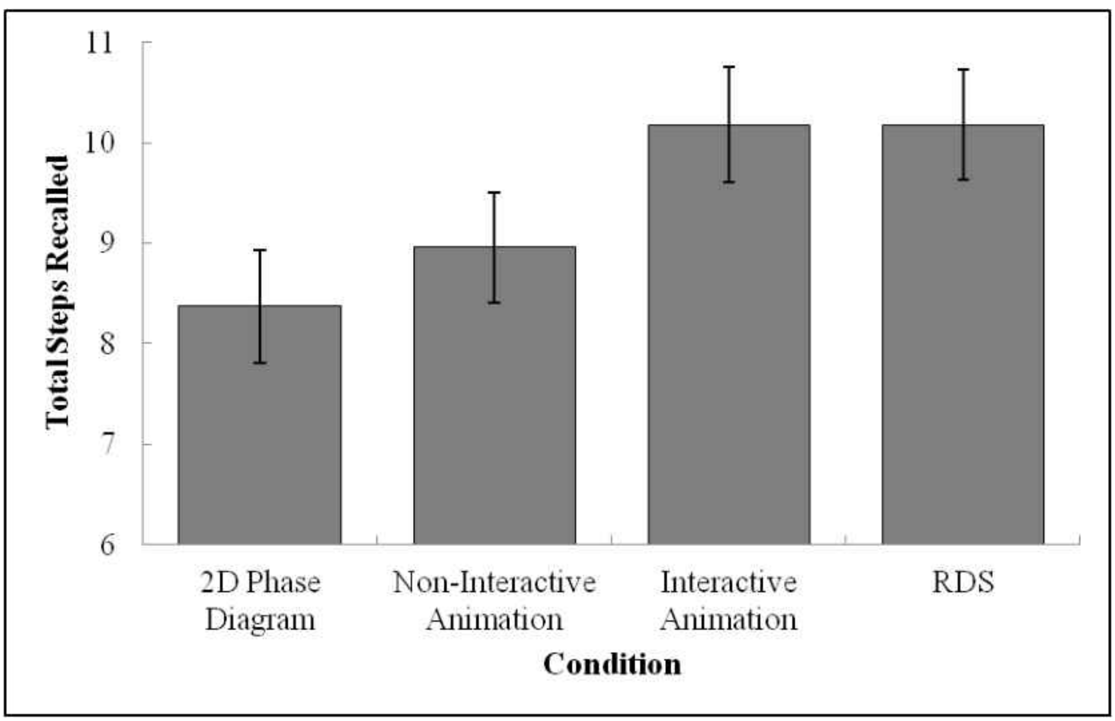


Figure 6. Number of correctly listed steps during the recall test by condition. Error bars represent standard errors.

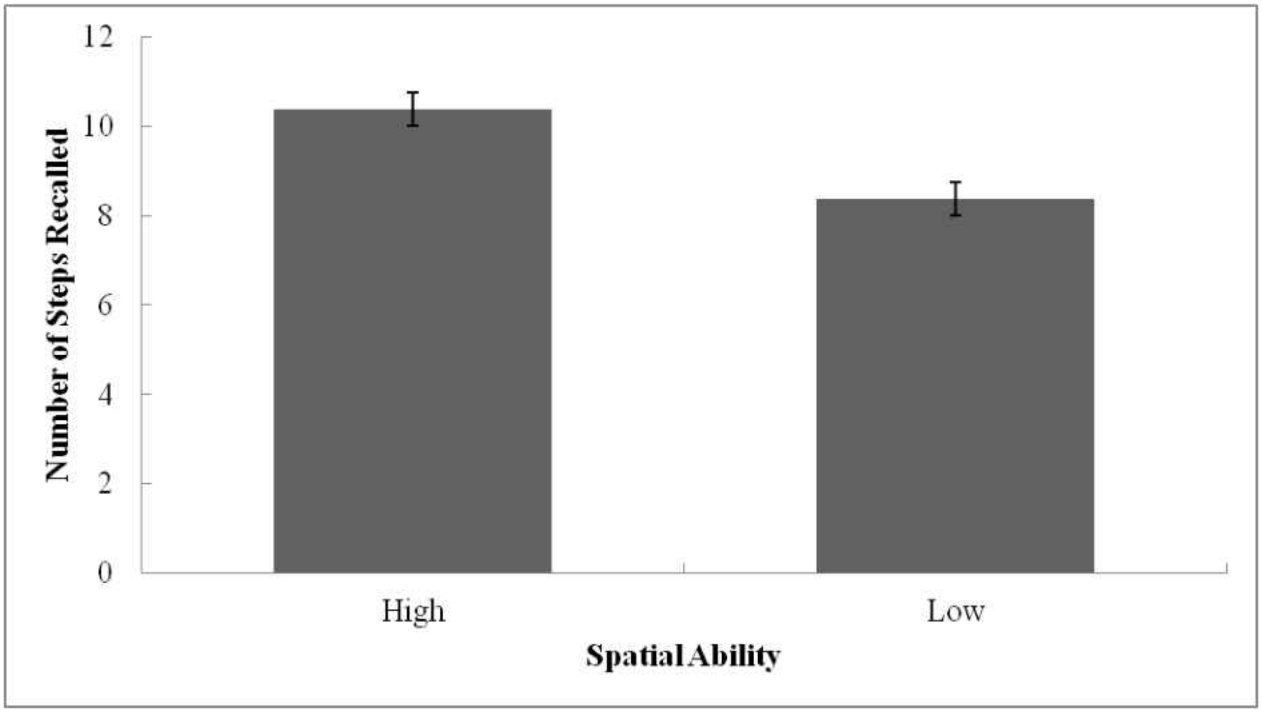


Figure 7. Number of correctly listed steps during the recall test by spatial ability. Error bars represent standard errors.

Table 5

Means and Standard Deviations for number of steps recalled for high and low spatials by condition.

	Spatial Ability					
	High		Low		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
2D Phase Diagram	9.70	2.67	7.59	3.14	8.37	3.10
Non-Interactive Animation	9.46	2.50	8.53	3.50	8.96	3.06
Interactive Animation	10.37	2.47	9.10	2.61	9.88	2.55
RDS	11.71	2.59	8.64	2.31	10.18	2.87

Total	10.38	2.63	8.38	2.95

Note. Means were out of 15 possible steps.

Disassembly

In order to assess the effects of multimedia training condition and spatial ability on disassembling the rifle, a 4 (multimedia training condition) x 2 (spatial ability) ANOVA was conducted for both number of steps completed and completion time. Effect sizes, means, and standard deviations were also reported.

Steps Completed

For number of steps completed, a statistically significant main effect was found for condition, $F(3,101) = 3.13$, $p = .029$, $\eta_p^2 = .085$. Planned comparisons using least significant differences (LSD) revealed the RDS condition significantly differed from both the non-interactive animation ($p = .003$) and the phase diagram conditions ($p = .014$). Participants in the RDS condition performed more disassembly steps ($M = 12.82$, $SD = 3.54$), than participants in the non-interactive animation condition ($M = 10.75$, $SD = 3.78$) and the phase diagram condition ($M = 10.3$, $SD = 3.54$). This lends partial support to Hypothesis 1, which predicted the interactive conditions would outperform the non-interactive conditions. However, only the RDS, showed a significant performance increase relative to the other two non-interactive conditions as seen in Figure 8. There was only partial support for Hypothesis 2, which predicted the three animated conditions would outperform the phase diagram condition. Only participants in the RDS condition outperformed participants in the phase diagram condition. Hypothesis 3 was not supported because there was no statistically significant condition by spatial ability interaction, $F(3,101) = .715$, $p = .546$, $\eta_p^2 = .021$, nor was there a statistically significant main effect for spatial ability, $F(1,101) = 2.09$, $p = .151$, $\eta_p^2 = .020$. Table 6 contains a complete list of means and standard deviations for disassembly steps completed.

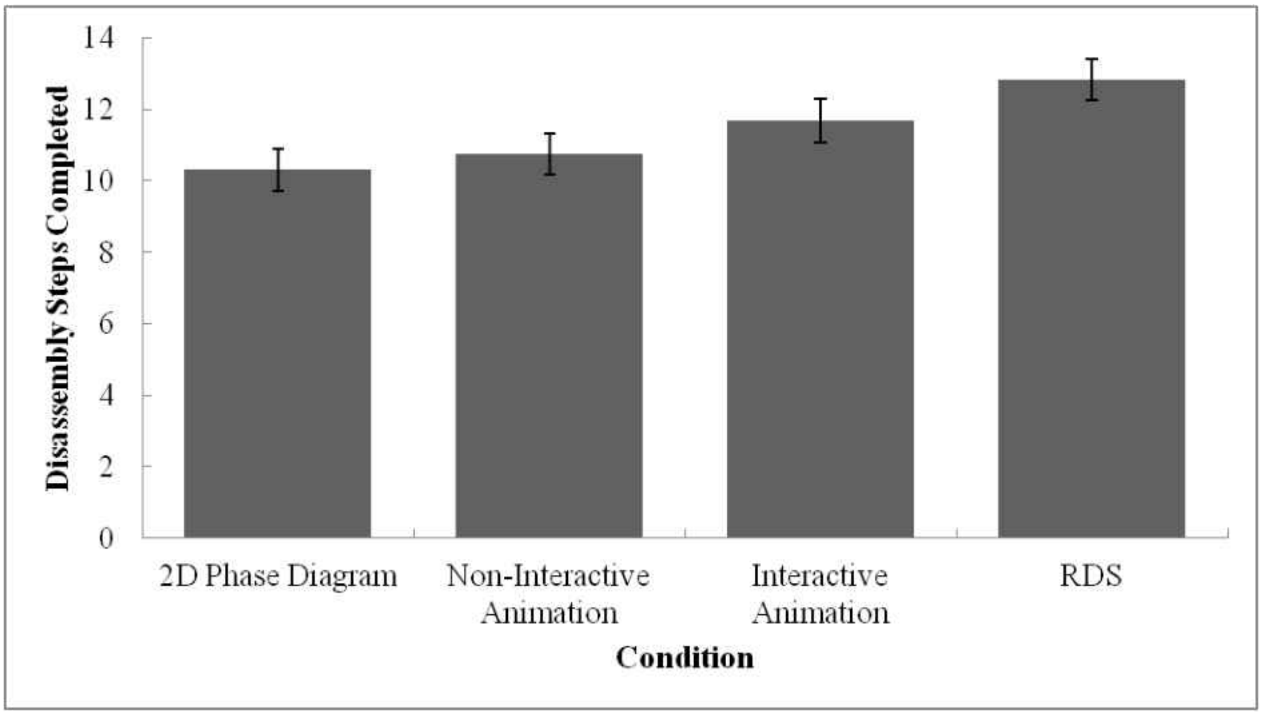


Figure 8. Number of steps completed during the disassembly task by condition. Error bars represent standard errors.

Table 6

Number of disassembly steps completed means and standard deviations for high and low spatials by condition.

	Spatial Ability					
	High		Low		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
2D Phase Diagram	11.30	3.06	9.71	3.75	10.30	3.54
Non-Interactive Animation	10.54	3.48	10.93	4.13	10.75	3.78
Interactive Animation	12.38	2.22	10.60	2.95	11.69	2.62
RDS	13.07	2.06	12.57	2.14	12.82	2.07
Total	11.91	2.80	10.91	3.48		

Note. Means were out of 14 possible steps

Completion Time

For disassembly completion time, a statistically significant main effect for spatial ability was found, $F(3,101)= 9.03$, $p = .003$, $\eta_p^2 = .082$. High spatials (M = 333sec, SD = 136) completed the disassembly task faster than low spatials (M = 419sec, SD = 146). No statistically significant main effect for condition or condition by spatial ability interaction was found. In spite of a moderate effect size for condition ($\eta_p^2 = .041$) no support for Hypothesis 1 or 2 was found. Figure 9 shows a non-significant trend whereby time to complete decreases with additional levels of dynamism and interactivity. Figure 10 depicts the main effect for spatial ability. Table 7 contains a complete list of means and standard deviations for completion time.

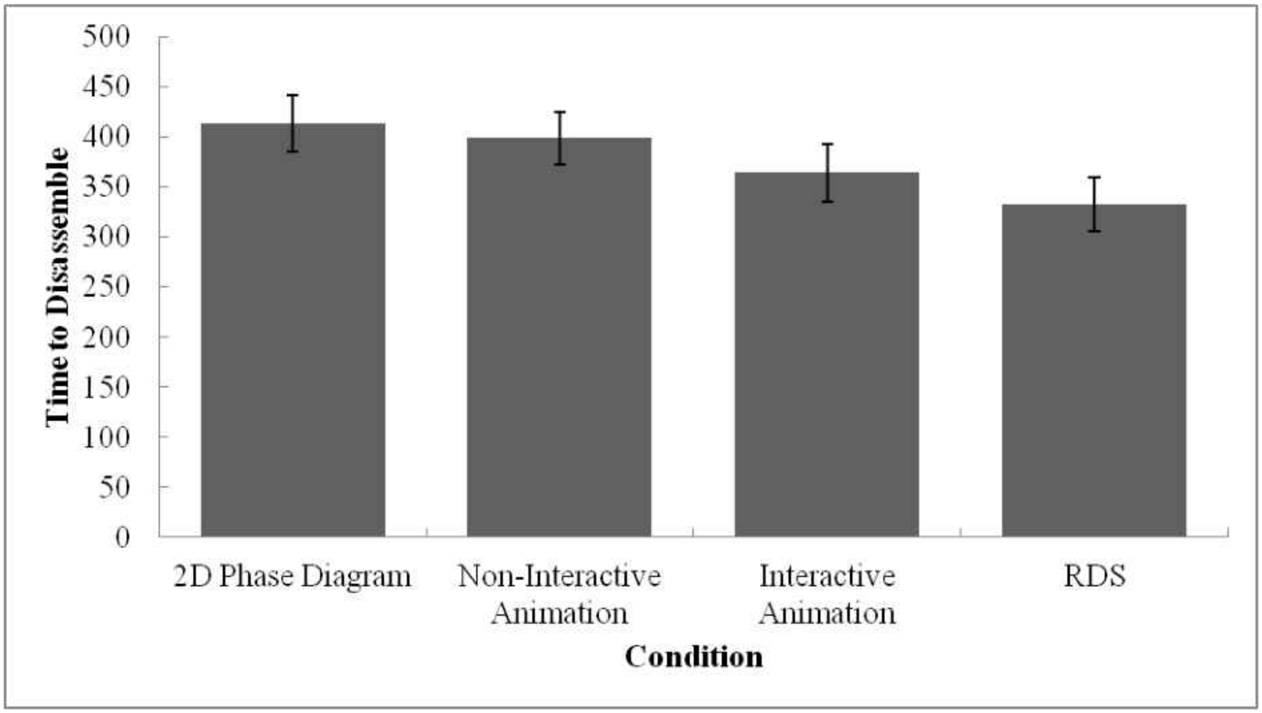


Figure 9. Time to complete disassembly task in sec. by condition. Error bars represent standard errors.

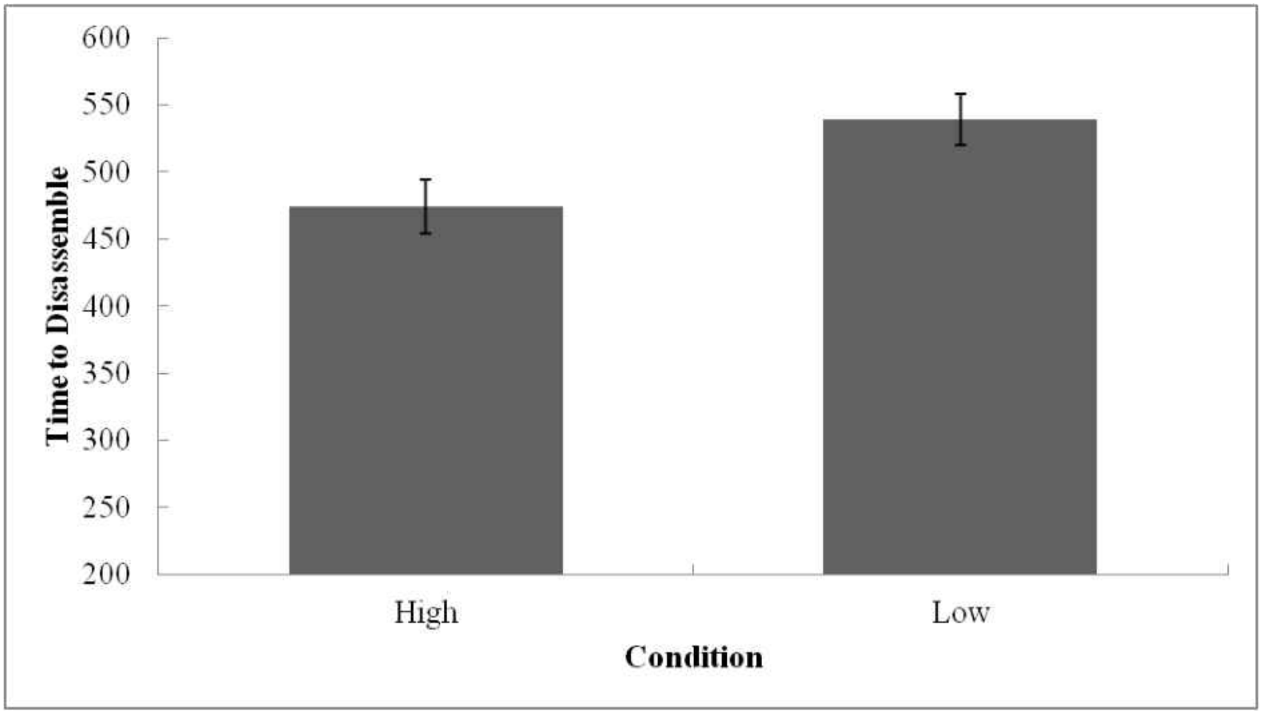


Figure 10. Time to complete disassembly task in sec. by spatial ability. Error bars represent standard errors.

Table 7

Disassembly completion time means (in sec) and standard deviations for high and low spatials by condition.

	Spatial Ability					
	High		Low		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
2D Phase Diagram	370	72.5	440	120.0	414	108.8
Non-Interactive Animation	381	170.9	414	168.3	399	167.2
Interactive Animation	314	128.9	445	151.2	364	149.8
RDS	284	135.1	419	145.8	333	149.9
Total	333	136.5	419	145.8		

Reassembly

In order to assess the effects of multimedia training condition and spatial ability on reassembling the rifle a 4 (multimedia training condition) x 2 (spatial ability) ANOVA was conducted for both number of steps completed and completion time. Effect sizes, means, and standard deviations were also reported.

Steps Completed

For total number of reassembly steps completed, a statistically significant main effect was found for spatial ability, $F(3,101) = 11.94$, $p = .001$, $\eta_p^2 = .106$. High spatials ($M = 9.87$, $SD = 2.38$) completed more reassembly steps than low spatials ($M = 8.14$, $SD = 2.37$). However, this main effect is better explained by a statistically significant condition by spatial ability interaction, $F(3,101) = 3.29$, $p = .024$, $\eta_p^2 = .089$. Planned comparisons revealed significant differences between high and low spatials in the phase diagram condition, $t(25) = 2.49$, $p = .020$, $d = .997$, and the interactive simulation, $t(25) = 4.016$, $p < .001$, $d = 1.61$. As seen in Figure 9, high spatials performed more reassembly steps than low spatials in those two conditions. This is consistent with Hypothesis 3 in that low spatials showed a drop in performance relative to the high spatials in the lowest and highest interactivity and dynamic conditions. Once again, no statistically significant main effect for condition was found in spite of a moderate effect size, $F(3,101) = 1.64$, $p = .184$, $\eta_p^2 = .047$, rendering no support for Hypotheses 1 or 2. Table 8 contains a complete list of means and standard deviations for reassembly steps completed.

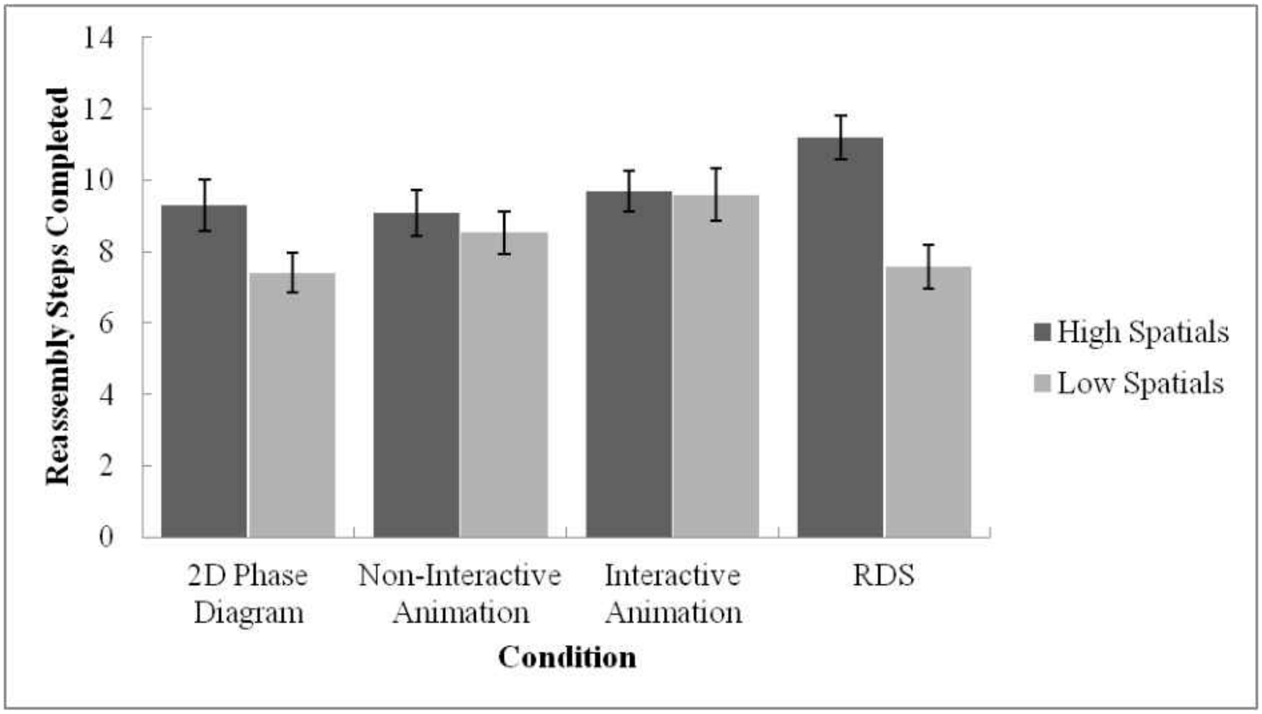


Figure 11. Number of steps completed during the reassembly task. Error bars represent standard errors.

Table 8

Number of reassembly steps completed means and standard deviations for high and low spatials by condition.

	Spatial Ability					
	High		Low		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
2D Phase Diagram	9.30	1.77	7.41	1.97	8.11	2.08
Non-Interactive Animation	9.08	2.29	8.53	2.39	8.79	2.32
Interactive Animation	9.69	2.65	9.60	2.12	9.65	2.42
RDS	11.21	2.16	7.57	2.62	9.39	3.00
Total	9.87	2.38	8.14	2.37		

Note. Means were out of 14 possible steps.

Completion Time

For time to complete reassembly, a statistically significant main effect for condition was found, $F(3,101) = 5.46$, $p = .002$, $\eta_p^2 = .140$. Planned comparisons using LSD revealed that the 2D phase diagram condition completed the reassembly phase slower than the animated and interactive conditions (non-interactive animation, $p = .032$; interactive animation, $p = .001$; RDS, $p < .001$). Figure 10 shows the main effect, which is consistent with Hypothesis 2. Hypothesis 1 received partial support such that both of the interactive conditions (RDS, $p < .001$, and interactive animation, $p = .001$) completed the task faster than the phase diagram condition. However, neither of the interactive conditions differed significantly from the non-interactive animation condition.

Once again, a significant main effect for spatial ability was found, $F(1,101) = 7.82$, $p = .006$, $\eta_p^2 = .072$, such that high spatials completed the reassembly task faster than low spatials. No statistically significant spatial ability by condition interaction was found in spite of a moderate effect size, $F(3,101) = 1.61$, $p = .191$, $\eta_p^2 = .046$. Table 9 contains a complete list of means and standard deviations for reassembly completion time.

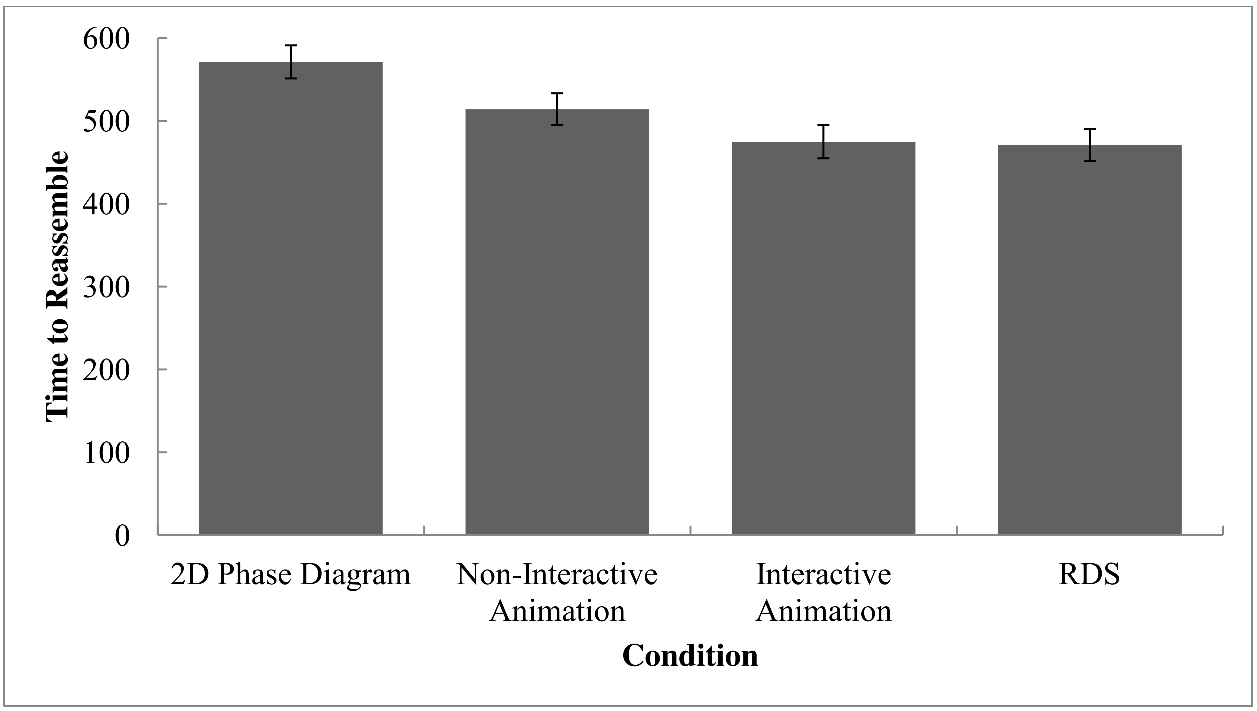


Figure 12. Time to complete reassembly task in sec. by condition. Error bars represent standard errors.

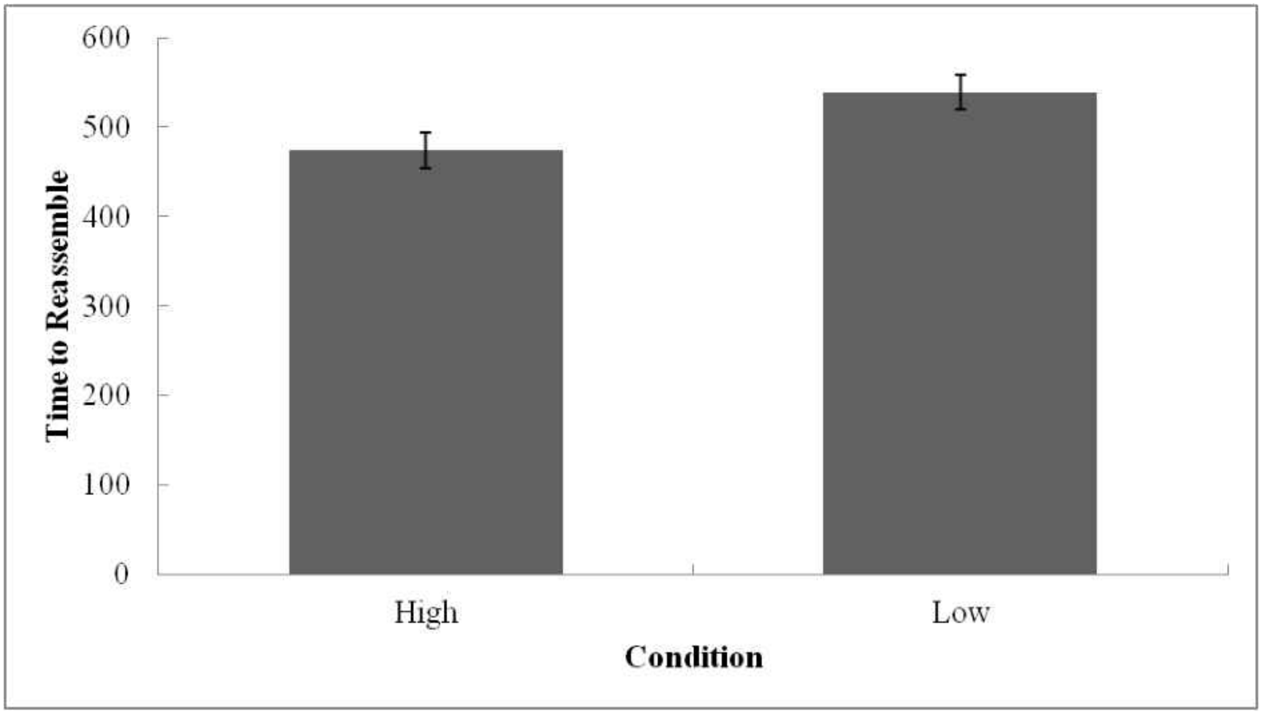


Figure 13. Time to complete reassembly task in sec. by spatial ability. Error bars represent standard errors.

Table 9

Reassembly completion time means (in sec) and standard deviations for high and low spatials by condition.

	Spatial Ability					
	High		Low		Total	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*
2D Phase Diagram	565	67.3	574	69.1	571	67.3
Non-Interactive Animation	492	135.3	533	100.1	514	117.3
Interactive Animation	459	111.3	501	100.4	475	107.3
RDS	410	108.9	531	65.0	471	107.4
Total	474	120.0	539	85.3		

Subjective Measures

For each subjective measure, a 4 (Multimedia Training Condition) by 2 (Spatial Ability) ANOVA was conducted to investigate the effects of condition and spatial ability on measures of perceived workload, cognitive load, usability and motivation. Appendix N contains means and standard deviations for each measure.

NASA TLX

To score the NASA TLX, the performance factor was reverse coded and summed with the other five factors to produce a single score of workload. A higher score indicated more perceived workload. In order to assess the effects of multimedia training condition and spatial ability on the NASA TLX, a 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted for each administration of the NASA TLX which followed the multimedia training phase, disassembly task, and reassembly task.

Multimedia Training Phase

For the training phase, there were no statically significant main effects for condition, $F(3,101) = .651$, $p = .584$, $\eta_p^2 = .019$, or spatial ability, $F(1,101) = .252$, $p = .617$, $\eta_p^2 = .002$, nor was there a significant condition by spatial ability interaction, $F(3,101) = .038$, $p = .990$, $\eta_p^2 = .001$.

Disassembly

For the disassembly task, there were no statically significant main effects for condition, $F(3,101) = .110$, $p = .954$, $\eta_p^2 = .003$, or spatial ability, $F(1,101) = 1.350$, $p = .248$, $\eta_p^2 = .013$, nor was there a significant condition by spatial ability interaction, $F(3,101) = .681$, $p = .566$, $\eta_p^2 = .020$.

Reassembly

For the reassembly task, once again, there were no statically significant main effects for condition, $F(3,101) = .454$, $p = .715$, $\eta_p^2 = .013$, or spatial ability, $F(1,101) = 1.07$, $p = .304$, $\eta_p^2 = .010$, nor was there a significant condition by spatial ability interaction, $F(3,101) = .730$, $p = .537$, $\eta_p^2 = .021$.

Measure of Cognitive Load

In order to assess the effects of multimedia training condition and spatial ability on the subjective measure of cognitive load, a 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted for each administration of the measure of cognitive load which followed the multimedia training phase, disassembly task, and reassembly task.

Multimedia Training Phase

For the multimedia training condition, a statistically significant main effect was found for condition, $F\ (3,101) = 3.96$, $p = .010$, $\eta_p^2 = .105$. Post-hoc tests using Tukey's Honestly Significant Difference (HSD) revealed that the phase diagram condition rated the training as significantly more difficult than both the interactive animation (p = .048) and RDS (p = .036) conditions. Means and standard deviations for the measure of cognitive load can be found in Table 10. There was no statistically significant main effect for spatial ability, $F(1,101) = .554$, $p = .459$, $\eta_p^2 = .005$, nor was there a significant condition by spatial ability interaction, $F(3,101) = 1.39$, $p = .250$, $\eta_p^2 = .040$. This finding is consistent with both Hypotheses 1 and 2, which stated that a lack of interactivity coupled with a lack of animation would result in higher levels of cognitive load.

Table 10

Means and standard deviations for the measure of cognitive load (out of 7).

Condition	*M*	*SD*
2D Phase Diagram	3.74	1.48
Non-Interactive Animation	3.89	1.13
Interactive Animation	4.69	1.41
RDS	4.71	1.24

*Lower scores indicate higher reported cognitive load

Disassembly

For the disassembly task, no statistically significant main effect for condition, $F(3,101)$ = ..989, $p = .401$, $\eta_p^2 = .029$, or spatial ability was found, $F(1,101) = 1.55$, $p = .217$, $\eta_p^2 = .015$, nor was there a statistically significant condition by spatial ability interaction, $F(3,101) = 1.48$, $p = .226$, $\eta_p^2 = .042$. The moderate effect size found for the interaction was further investigated with a follow up ANOVA. However, so statistically significant effects were found between high and low spatials in any of the conditions indicating a high degree of variability in the data.

Reassembly

For the reassembly task, a statistically significant spatial ability by multimedia training condition interaction was found, $F(3,101) = 3.07$, $p = .031$, $\eta_p^2 = .083$. Follow up ANOVAs revealed that, consistent with the interaction found during the reassembly task, low spatials (M = 2.36, SD = 1.08) rated the reassembly task as more difficult than high spatials (M = 3.64, SD = 1.55) in the RDS condition. This result is consistent with low spatials dealing with higher levels of cognitive processing to deal with the reassembly task. The reassembly task for high spatials on the other hand, although demanding, did not require as much cognitive resources as low spatials.

For condition, no statistically significant main effect was found in spite of a moderate effect size, $F(3,101) = 2.08$, $p = .108$, $\eta_p^2 = .058$. Furthermore, no statistically significant main effect for spatial ability was found, $F(1,101) = 1.30$, $p = .258$, $\eta_p^2 = .013$.

Usability

The usability measure was administered once right after the multimedia training phase. The mean score (out of 7) was recorded. A higher score indicated higher rated usability. In order to assess the effects of multimedia training condition and spatial ability on perceived usability of the training, a 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted. No statistically significant main effect for condition, $F(3,101) = ..880$, $p = .454$, $\eta_p^2 = .025$, or spatial ability was found, $F(1,101) = 3.751$, $p = .055$, $\eta_p^2 = .036$, nor was there a significant spatial ability by condition interaction, $F(3,101) = 1.81$, $p = .149$, $\eta_p^2 = .051$. Spatial ability was close to significance however, the effect size was no large enough. On the other hand, the effect size for the interaction was moderate. A follow up ANOVA investigating the interaction revealed a similar pattern to previous findings, $F(1,27) = 4.52$, $p = .044$, $\eta_p^2 = .153$. For the 2D phase diagram condition, low spatials ($M = 4.95$, $SD = 1.57$) rated the multimedia training as less usable than high spatials ($M = 6.15$, $SD = 1.10$). However, the difference in rating by spatial ability did not reach significance for the RDS condition in spite of a moderate effect size, $F(1,26) = 1.21$, $p = .282$, $\eta_p^2 = .044$.

Motivation

The average response for the IMI (out of 7) was recorded. A higher score indicated a higher perceived motivation. In order to assess the effects of multimedia training condition and spatial ability on motivation, a 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted. No statistically significant main effect for condition, $F(3,101) = .1.03$, $p = .381$, $\eta_p^2 = .030$, or spatial ability, $F(1,101) = .015$, $p = .904$, $\eta_p^2 < .000$ was found. For spatial ability by condition interaction, a moderate effect size was found in spite of a lack of statistical significance, $F(3,101) = 2.32$, $p = .079$, $\eta_p^2 = .065$. The interaction was further investigated with a follow up ANOVA that did not reach statistical significance indicating a high degree of variability within the measurement.

CHAPTER FIVE: EXPERIMENT 1 DISCUSSION

Experiment 1 examined four conditions of varying amounts of interactivity and dynamism to answer two research questions. The primary research question was whether including multiple forms of interactive features in a dynamic multimedia increase learning outcomes of a procedural-motor task. The secondary research question was whether dynamic multimedia was superior to equivalent static multimedia in training a procedural-motor task. Spatial ability was also considered as a subject variable when addressing the research questions. To address these questions, the four conditions were compared in terms of retention, as measured by a recall test, and as measured by disassembling (trained) and reassembling (untrained) a replica M4 carbine. In general, high spatials outperformed low spatials in most performance measures. Additionally, the higher levels of interactivity with animation outperformed the 2D phase diagram condition. However, this result, in terms of reassembly performance, was dependent on spatial ability. High spatials demonstrated enhanced performance in the RDS condition while also compensating for the lack of interactivity and dynamism in the 2D phase diagram condition relative to low spatials. Low spatials appeared unable to effectively utilize the high levels of interactivity or compensate for a lack of interactivity and dynamism in the 2D phase diagram. Overall, with regard to the primary research question, it appeared that multiple forms of interactivity did foster learning outcomes, however this result was contingent on spatial ability. For the secondary research question, dynamic multimedia was superior to static multimedia. The following sections discuss the results in terms of the specific hypotheses.

Effects of Interactive Multimedia on Learning

Hypothesis 1 predicted that the interactive conditions (i.e., interactive animation and RDS) would outperform the non-interactive conditions (i.e., non-interactive animation and 2D phase diagram) in the recall test and transfer of training tests. The results partially supported this hypothesis. The RDS was more effective than both of the non-interactive conditions for the disassembly task as measured by number of steps completed. The RDS and interactive animation was also superior to the 2D phase diagram in the reassembly task as measured by completion time. This pattern of results is consistent with the idea that interactivity facilitates learning above and beyond non-interactive multimedia by reducing cognitive overload and facilitating germane load (Mayer & Moreno, 2007). The results of Experiment 1 are in line with Kalet et al., (2012) who found that including just the right amount of interactivity increased learning outcomes and that too much interactivity resulted in inferior performance on a transfer of training test. Although Experiment 1 did not find inferior performance with increased interactivity, the RDS was never significantly better than interactive animation on any measures of performance. The present study results indicate that adding additional forms of interactivity, did not directly result in performance decrements as in Kalet et al., (2012), but did result in diminishing returns. Each additional level of interaction added to a multimedia presentation may not increase the learning outcomes to the same extent as the previous level. Future research should systematically investigate the rate of return for each layer of interactivity that is added to a multimedia presentation.

The subjective measures of cognitive load were met with few significant results. The measure of cognitive load, taken directly after the training phase, supported the assertion that multimedia presentation with static images and no interactivity would impose the highest amount of cognitive load. The 2D phase diagram condition was rated as being significantly more difficult to learn from than the RDS and interactive animation conditions, which was mirrored in the transfer of training task performance. Unfortunately, it is unclear based on the other subjective measures whether the interactivity reduced extraneous load by being perceived as more usable, or if it increased germane load, as would have been partially indicated by measuring motivation with the IMI. However, a non-significant trend existed whereby subjective ratings of motivation increased as interactivity and dynamism increased. Motivation has been linked to germane load such that higher levels of motivation are thought to increase the overall working memory capacity of the individual thereby allowing more resources to be dedicated to germane load (Rieber, 1991; Paas, Renkl, & Sweller, 2003; Moreno & Mayer, 2007). Although not significant, the trend is in line with other results suggesting the benefits of interactivity in multimedia.

The results of the NASA TLX were inconclusive. This may have been because the construct of workload is not the same as cognitive load. The NASA TLX includes measures such as temporal demand, and physical demand that were not directly related to the task. Participants were under no time pressure to complete the training, nor were they required to perform any physical actions. Although these constructs are important in other domains, they may not have been relevant for this particular set of tasks.

Effects of Dynamic Multimedia on Learning

Hypothesis two suggested that dynamic multimedia would be superior to static multimedia in terms of recall test and transfer of training task performance. Results indicated that in the absence of interactivity, the non-interactive animation condition did not show any performance difference compared to the 2D phase diagram condition in the recall test and disassembly test. However, participants in the non-interactive animation condition were able to reassemble the rifle faster than those in the 2D phase diagram condition. This is consistent with Wong et al. (2009) and Ayers et al. who argue that extraneous load due to animations' transient information is lessened when the task being learned involves human motion. They argue that in addition to the auditory and visual channel associated with working memory (see Baddely, & Hitch, 1974; Paivio, 1990), there also exists a subsystem specifically for movement. Because the disassembly inherently requires movement, the working memory allocation for the task becomes more efficient due to the use of an additional subsystem. Within the context of Cognitive Theory of Multimedia Learning, a third channel could exist to aid in selecting, organizing, and integrating information in the form of a sensorimotor component. Utilizing the additional channel would help lessen the load placed on the other two existing channels resulting in less instances of cognitive overload.

However, with the addition of interactivity, the non-interactive animation did not significantly differ with the interactive animation in any of the measures of performance. This could be a result of how these two conditions were structured. In both conditions, the animation sequences occurred one at a time (i.e., only one component moved at any given time). Even though the entire presentation did not come to a complete halt after each step in the non-interactive animation condition, the presentation may have moved slowly enough that the addition of a “next” button was unnecessary. The information being directed to the learner was coming across slowly enough that it did not suffer from the transient effect described by Hegarty (2004). Overall, the results of this study support the conjecture that dynamic multimedia is effective when training a procedural motor task.

Effects of Spatial Ability on Learning

Hypothesis 3 suggested that high and low spatials would perform differently depending on the amount of interactivity and dynamism was present in each multimedia condition. Specifically that low spatials would suffer in the 2D phase diagram and RDS conditions relative to high spatials. Overall, spatial ability was a key predictor of recall and re-assembly performance, such that high spatial ability individuals outperformed those with lower spatial ability in the 2D phase diagram and RDS conditions. The ability for high spatials to process spatial concepts more efficiently aided in their performance. This was particularly evident in the reassembly task, which required the deepest understanding of the procedures and how components interacted without the cues provided during the disassembly test. In the 2D phase diagram condition, there was a high amount of extraneous load due to requiring the individual to conduct their own mental transformations as opposed to an animated sequence, which would have displayed the transformation. In this condition, low spatials may have suffered more cognitive overload than high spatials who had more working memory capacity available for the spatial processing required. This is supported by the main effect for condition reported for perceived usability of the system indicating that low spatials rated the training as less usable than high spatials, specifically in the 2D phase diagram condition. The RDS condition, which provided animated sequences of the steps, required interacting with a more complex interface than the other conditions. This was also to the detriment of low spatials who showed a drop in performance relative to the high spatials. The addition of a more complex interface created high extraneous load demands that precluded low spatials from sufficient germane processing. Alternatively, high spatials were able to excel with the additional forms of interactivity that was able to foster superior schema creation. This finding can also be attributed to high spatials ability to use the interactivity more effectively by finding an optimal viewpoint in order to see the animation sequence take place. This is consistent with Keehner et al. (2010) who noted that it may be an individual's ability to identify the optimal viewing angle and not the interactivity per se that predicts subsequent task performance. High spatials may have been better able to identify proper viewing angles than low spatials. Because high spatials excelled in this condition, their ability to find optimal viewing angles may have been due to the additional working memory resources available for them to spend effectively manipulating the interface. This could also explain why the interactive animation condition did not differ significantly in terms of spatial ability during the reassembly test. In the interactive animation condition, the optimal viewpoints were given to the participants and therefore they did not have to expend any cognitive resources determining the ideal way to manipulate the interface.

It is interesting that spatial ability was not a significant predictor of performance during the disassembly task. This is probably because less spatial working memory was involved when figuring out how to take apart the rifle. In essence, one could figure out some disassembly steps simply by looking at rifle and deducing the steps necessary to take it apart. This would not require spatial ability per se in the sense that no mental transformations were required to remember a step. For example, knowing to take out the receiver pivot pin would not require spatial knowledge of how the pivot pin fits into the rifle. Simply knowing that the pivot pin needed to be removed was enough.

Theoretical and Practical Implications

From a theoretical perspective, the results suggested that multiple forms of interactivity placed in a single multimedia presentation, as in the RDS, reduced the overall amount of cognitive load for high spatials as measured by their performance on the transfer of training tasks. The opposite was true for low spatials who did not perform as well after learning from the RDS condition. Based on these results, it appears as though the extraneous load placed on low spatials by the interactive interface left few resources that were able to be committed to germane load for encoding the information. Instead of actively processing the information, low spatials may have engaged in processes devoted to determining how to best use the interactivity, which may have interfered with the learning process. Although high spatials may have experienced similar amounts of extraneous load due to the interface, their ability to handle spatial information more efficiently allowed more resources to be devoted to germane load processes.

However, despite the effort to capture the individual facets of cognitive load (i.e., extraneous and germane) to support the inferences made from the performance measures, it is still unclear exactly how working memory was affected. The inability to measure the individual components of cognitive load has been a noted challenge within the Cognitive Load Theory literature (Paas, Tuovinen, Schnotz & Kurschner, 2007; Tabbers, & Van Gerven, 2003; Paas et al., 2003). Experiment 1 took the approach of measuring subjective measures that could contribute to extraneous (i.e., usability) and germane load (i.e., motivation), however this proved to be ineffective in the current study. However, this should not preclude Cognitive Load Theory as a viable framework to investigate instructional design. As Schnotz and Kurschner (2007, pg. 500) mention when referring to Cognitive Load Theory, "A framework does not require that each theoretical construct needs its own measurement procedure. Other theoretical frameworks – such as schema theory or production systems – have also been very fruitful without offering an empirical measurement procedure for each specific construct." The challenge in measuring intrinsic, extraneous, and germane load should not preclude designing instructional multimedia so that extraneous load is reduced and germane load is fostered.

From a practical perspective, the results of Experiment 1 suggest that a one-size-fits-all approach to designing a multimedia training presentation does not apply. Individual differences may affect the ability of an individual to learn effectively from the presentation. In the case of Experiment 1, spatial ability affected learning outcomes differently, specifically in the RDS and 2D phase diagram conditions. Based on these results, two different approaches might be taken when developing training multimedia. In order to reach the most amount of people, a middle ground of interactivity (e.g., the interactive animation condition) may be incorporated the multimedia, which would provide adequate training to both high and low spatials. The other approach may be to pre-test individuals for spatial ability and then determine which training multimedia they should interact with. This method would presumably provide the highest training outcomes for both types of individuals at the expense of needing to develop two different types of multimedia, which might result in higher development costs. Overall, Experiment 1 indicated that at least some form of interactivity coupled with dynamic multimedia was an effective way to train the disassembly procedures of a Colt M4.

CHAPTER SIX: EXPERIMENT 2

The results of Experiment 1 provided insight as to how added interactivity using a touchscreen interface on a tablet computer affects performance on a transfer of training task. Experiment 2 further investigated tablets by addressing whether there is something special about the touchscreen. Two research questions were posed. First, are gestures, as a means to interact with a tablet, effective at teaching a procedural-motor task due to the sensorimotor component they incorporate? Two, does touching a touchscreen result in a more embodied interaction that facilitates learning more than indirect touch inputs, such as mouse?

Tablet computers afford using interactive graphics partly due to their touch interface and ability to manipulate content. In order to investigate the effect of interaction in dynamic multimedia on learning from mobile devices more fully, the touch UI must be considered. Mobile devices offer users the ability to manipulate digital content via direct touch input without mediation from an external piece of equipment (e.g., mouse, keyboard, joystick etc.). That is, a user incorporates gestures (e.g., pinching to zoom and dragging a finger to rotate) to manipulate content on the screen. Furthermore, the gestures themselves may also contribute to learning by adding an additional process by which to encode new information. Since there is little research examining the use of mobile devices in training and education, it is unclear whether the added ability to use a touch interface will affect learning from a mobile platform. In addition, this research may shed some light on how the mind uses multiple modalities to encode information. Embodied Cognition was used as an explanatory foundation of how a touch UI may provide learning benefits.

In Experiment 2, participants interacted with the multimedia presentation interacting by way of direct touch input, or indirect mouse input. The training conditions in Experiment 2 were identical to the RDS used in Experiment 1 with the addition of gesture arrows that simulated the motion required to complete the disassembly step. The conditions using gesture elicited arm and hand motions associated with the selecting and "removing" the component aimed to create stronger encoding when learning the information via embodied interaction. The conditions not utilizing gesture were not thought to be as effective for encoding information. Furthermore, participants' interaction with the tablet was coded and summed to investigate how the conditions encouraged interactivity and whether the interactivity was related to performance on the recall test and transfer of training tasks. It was hypothesized that directly touching the screen would facilitate embodied interaction more than indirectly interacting with the mobile device with a mouse.

Hypotheses

Based on the existing literature and the specific research questions posed by this study, several hypotheses were proposed.

Hypothesis 1

It was hypothesized that participants in the RDS + Gestures condition, regardless of input (direct or indirect) would learn more from the multimedia presentation than those in the RDS condition.

Hypothesis 1 is based on the Embodied Cognition paradigm, which emphasizes the role of the sensorimotor system during encoding and recall of information (Barsalou, 2008). Similar to the way Cognitive Theory of Multimedia Learning and CTL leverage the cognitive architecture of the mind to develop effective principles for multimedia instruction, further incorporating the potential benefits of sensorimotor encoding within a multimedia interface should increase learning outcomes. Including an interactive gesture will facilitate a sensorimotor coupling between the user and the interface, which should provide motor cues associated with the material that may help encoding and therefore improve learning outcomes.

Based on Hypothesis 1, the following predictions have been made. If the sensorimotor component associated with tracing the gestures leads to stronger encoding and therefore better learning, I predict that the RDS + Gestures condition will have a (1) higher recall test performance, (2) higher transfer of training task performance for both disassembly and reassembly based on time and number of steps completed. However, if the addition of gestures to the interface imposes high levels of cognitive load, I would expect that (1) the RDS + Arrows condition would show higher performance on the recall test and transfer of training tasks than the RDS + Gestures condition. In this situation, it would not be expected that the RDS + Gesture condition would significantly differ from the RDS condition in terms of recall and transfer of training task performance. The RDS + Arrows condition contains additional visual information relative to the RDS condition but lacks the gesture component, which could overload the user.

Hypothesis 1(a)

It was hypothesized that the RDS + Arrows condition would show higher learning outcomes than the RDS condition, but lower learning outcomes than the RDS + Gestures condition.

Hypothesis 1(a) is based on the additional information that the RDS + Arrows condition displayed relative to the RDS condition. It follows that the additional embodied interaction present in the RDS + Gestures condition will further facilitate learning above and beyond what is facilitated by the arrows. On the other hand, if the RDS + Arrows condition outperforms the RDS + Gestures condition on the measures of recall and transfer of training, it could be theorized that it was the arrows, but not the gestures that were facilitating learning.

Hypothesis 1(b)

It was hypothesized that spatial *ability would interact with the RDS such that high spatials will demonstrate better learning outcomes relative to low spatials.*

Hypothesis 1(b) mirrors Hypothesis 1(a) from Experiment 1 using the assumption that high spatials have more capacity handle spatial information in working memory than low spatials. Based on this assumption, it is hypothesized that the embodied interaction in the RDS + Gestures condition and the informational arrows present in the RDS + Arrows condition will facilitate learning even for low spatials. However, because the RDS condition does not contain those instructional elements, low spatials will be subject to the cognitive overload demonstrated in Experiment 1.

Hypothesis 2

It was hypothesized that direct touch input (finger) would lead to better learning outcomes than indirect mouse input.

This hypothesis was based on the idea that directly touching the screen will lead to more embodied interaction due to the one-to-one nature of the interaction. Although the mouse requires some movements of the arm and hand, the fact that this motion is mediated by a mouse could lead to less embodied interaction than that of the direct touch input. This hypothesis is strongly supported by Tan et al. (2002) and Jetter et al. (2012) who found that spatial learning was better when the interface interaction utilized direct touch input as compared with indirect mouse input.

Based on Hypothesis 2, the following predictions have been made. The touch conditions, compared to the mouse conditions, will exhibit (1) better performance on the recall test and (2) better performance, as measured by time and number of steps completed, on the two transfer of training tasks.

Hypothesis 2(a)

It was hypothesized that input would interact with the multimedia training condition such that participants in the touch condition will have higher learning outcomes in the RDS + Gestures condition than the other multimedia training conditions.

This hypothesis was based on combining the embodied interaction of gesture with the one-to-one interaction that touch provides. The sensorimotor experience during the gestures should have a stronger encoding when the associated motions occur directly with the screen instead of indirectly with the mouse.

CHAPTER SEVEN: EXPERIMENT 2 METHOD

Participants

One hundred fifty three college students (80 males, 72 females) between the ages of 18-44 (*M*=20.6, *SD*=3.8) were recruited for the study using the UCF Psychology Department's online recruitment tool and received class credit for their participation. All participants were over the age of 18 at the time of the experiment. Participants included in this sample were inexperienced with weapon disassembly procedures as measured by a prior knowledge questionnaire. Because rifles frequently require similar disassembly procedures, naïve subjects were essential to ascertain what knowledge came directly from the training multimedia. The nine individuals who recorded a score two standard deviations above the mean on the rifle experience questionnaire were omitted from data analysis resulting in 144 analyzed cases (72 Males, 72 Females). Participants were randomly assigned to one of the four training conditions: phase diagram, non-interactive animation, interactive animation, and interactive simulation.

Design

A 2 (input: mouse or touch) x 3 (multimedia training condition) design was used resulting in six between-participants experimental conditions.

Multimedia Training Conditions

Experiment 2 examined the effects of including interactive gestures as part of the multimedia training. To that end, three training conditions were included for experiment 2: RDS, RDS + Gestures, and RDS + Arrows.

RDS

The RDS condition was identical to that of Experiment 1 in which participants tapped a component once to select it and then tapped the component again to engage the animation. Participants listened to a narrated instruction and subsequently tapped on the relevant component to highlight it. After the correct component was selected, tapping on it again completed the step. The other conditions were derived from the RDS.

RDS + Gestures

This condition was identical to the RDS condition with the exception of the following. After hearing the narrated instruction and selecting the relevant component, instead of tapping on the component again to complete the step, a gesture arrow appeared representing the way in which the component moved. Tracing the gesture (with either the mouse or a finger) completed the step.

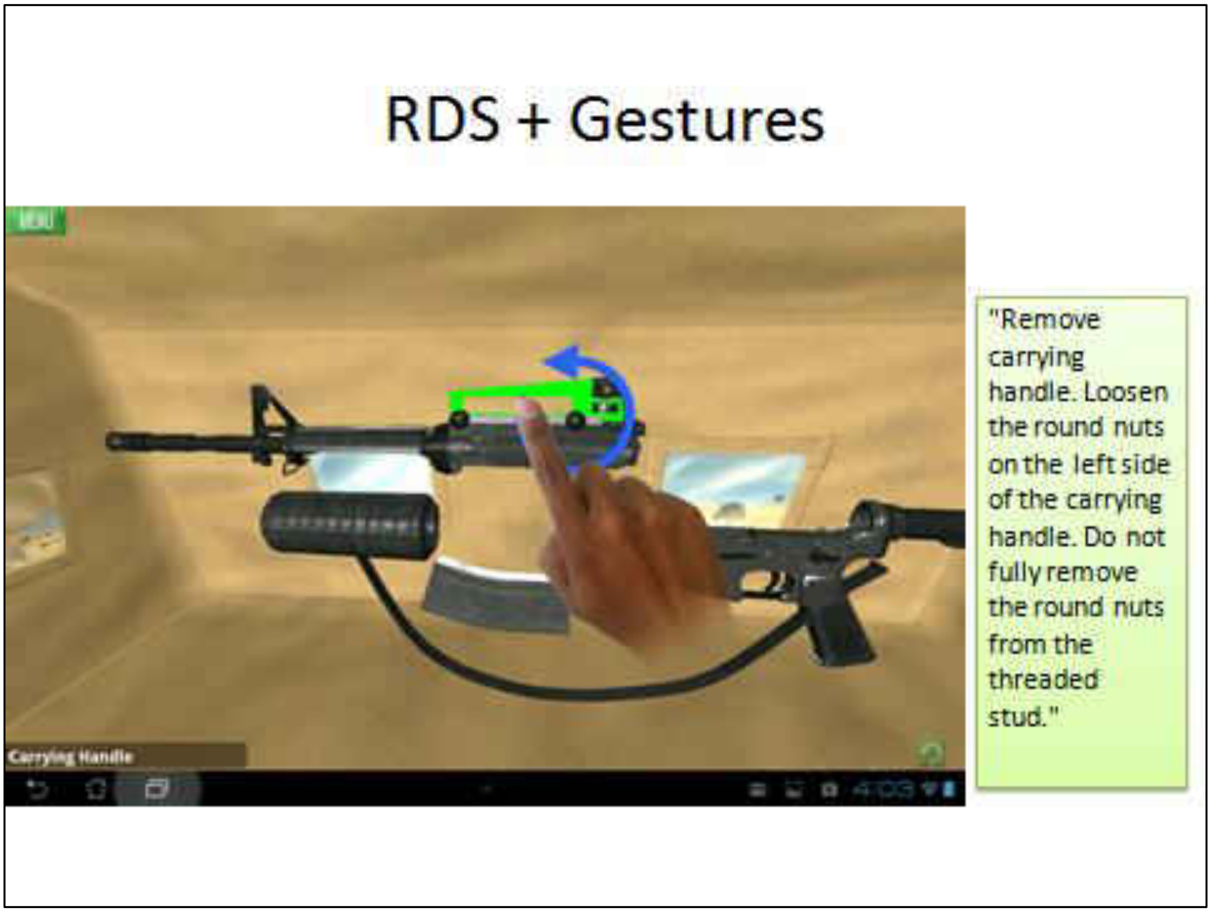


Figure 14. Screen shot of the RDS + Gestures condition.

RDS + Arrows

In an effort to control for any information obtained by the gesture arrows presented in the RDS + Gestures condition, the RDS + Arrows contained the gesture arrows but did not include the tracing component. After a participant selected the relevant component, the gesture arrow appeared in the same form as it did in the RDS + Gesture condition, however, instead of tracing the gesture arrow, participants only tapped on the component again, similar to the RDS condition.

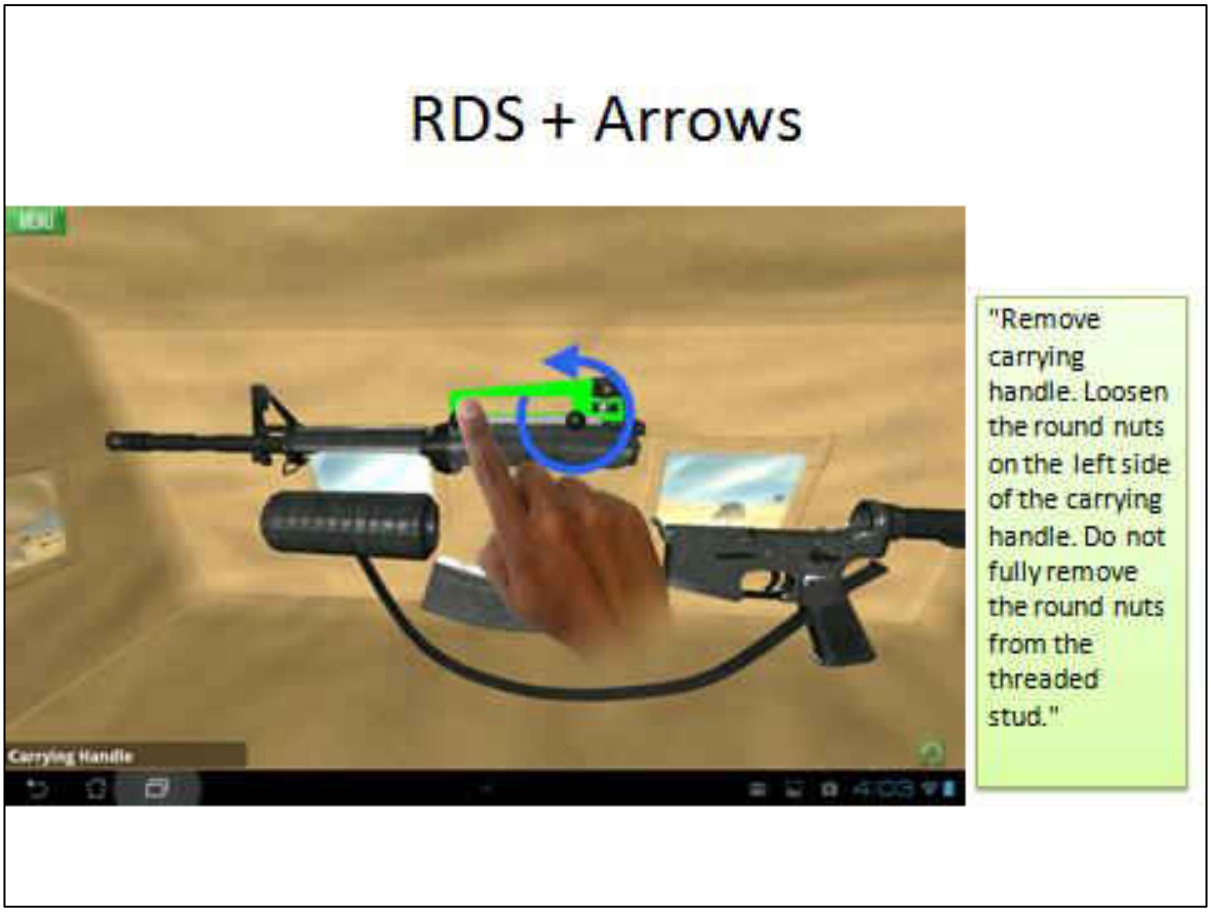


Figure 15. Screenshot of the RDS + Arrows condition.

Input

Two forms of input were compared in addition to the multimedia training conditions: direct touch and indirect mouse. The direct touch input conditions utilized the touchscreen interface present on the tablet device. To interact with the application, participants used their fingers to tap and drag essential areas on the screen. For the indirect mouse condition, an attachment was used to connect a standard 3-function mouse (left-click, right-click, and scroll wheel) to the tablet in lieu of the touchscreen. Instead of touching the screen, participants controlled aspects of the application using gestures that were mapped to the mouse functionality. Table 12 contains the complete list of interaction gestures for the Mouse and Touch conditions.

Table 11

Interaction gestures for Touch and Mouse conditions.

Interaction	Touch	Mouse
Rotate	1 finger + swipe	Left mouse button + swipe
Translate	2 finger + swipe	Left & Right button + swipe
Select	1 finger tap	Left button click
Zoom in/out	Pinch in/out	Scroll wheel

Apparatus

Tablet

Experiment 2 was conducted using the same Asus Transformer Infinity tablet as Experiment 1. The tablet had a 10.1-inch screen with 1920 x 1200 resolution using the Android operating system. For Experiment 2, a mouse attachment was used for the indirect touch conditions.

Camera

A digital, hand-held video recorder situated on a tripod was used to capture the disassembly, re-assembly and tablet interaction.

Survey Administration

All surveys were administered on a desktop computer using the Qualtrics survey creation website.

Replica Rifle

The same replica rifle utilized in Experiment 1 was also used for the transfer of training tasks in Experiment 2.

Materials

Experiment 2 used the same materials listed in Experiment 1.

Performance Measures

The recall and transfer of training tasks used in Experiment 1 were also used in Experiment 2.

Procedure

Experiment 2 used the identical procedures as Experiment 1 with the following change. Because performance in the RDS condition during Experiment 1 was approaching a ceiling effect, Experiment 2 only had participants go through the training portion once instead of twice.

CHAPTER EIGHT: EXPERIMENT 2 RESULTS

Preliminary Analysis

Several preliminary analyses and variable coding, mirroring Experiment 1, were conducted prior to examining the research question and hypotheses. First, an outlier analysis was conducted to remove any individuals with sufficiently high previous experience with firearms. To accomplish this, the rifle experience questionnaire was summed and recorded. A higher score indicated a higher level of experience with firearms. The mean score was 1.05 with a standard deviation of 1.87. Individuals scoring higher than two standard deviations from the mean were removed from analysis (9 total) leaving 144 analyzed cases.

An one-way ANOVA was conducted to ensure equality across each condition for a variety of demographic measures including spatial ability, video-game experience and tablet ownership. No significant differences were found.

A correlation analysis was conducted between the two measures of spatial ability (Card Rotations Test and Paper Folding Test). Once again, the two measures were found to be significantly correlated, $r(143) = .343$, $p < .001$. Therefore these variables were standardized and combined into a single measure of spatial ability. A median split was conducted separating spatial ability into high and low spatial ability groups. A follow up t-test was conducted to ensure that both groups significantly differed in their spatial ability. Both groups were significantly different, $t(141) = 16.84$, $p < .001$, $d = 2.84$. Individuals in the high group ($M = .905$, $SD = .648$) had a higher measured spatial ability than individuals in the low group ($M = -1.04$, $SD = .732$) Table 12 shows the breakdown of high and low spatial ability by condition.

Table 12

Number of participants in each condition by spatial ability.

		Mouse		Touch	
Condition	Spatial Ability	High	Low	High	Low
RDS		13	10	15	9
RDS + Gestures		12	13	10	13
RDS + Arrows		8	16	14	10

For the recall test, two raters independently coded all of the participant responses and awarded a point for every step that could be clearly identified. The total number of steps was summed and recorded. To verify the reliability of the scoring, the two raters' scores were correlated resulting in a Pearson's correlation coefficient of .880 indicating a very high overall agreement between the raters.

For the transfer of training tasks (disassembly and reassembly), the number of completed steps for each stage was combined and summed. Similarly, the time to complete each stage for reassembly was summed and recorded in seconds.

Performance Measures

Recall Test

For the recall test, a 3 (multimedia training condition) x 2 (input modality) by 2 (spatial ability) ANOVA was conducted using the number of correctly identified steps recalled. Effect sizes, means, and standard deviations were reported.

For multimedia training condition, there was a statistically significant main effect for spatial ability, $F(1,131) = 22.7$, $p < .001$, $\eta_p^2 = .146$. High spatials ($M = 8.96$, $SD = 2.90$) listed more correct steps than low spatials ($M = 6.62$, $SD = 2.71$). No other statistically significant effects were found. Table 13 contains the ANOVA table for the recall test. Table 14 contains a complete list of means and standard deviations for all conditions by spatial ability.

Table 13

ANOVA Table for the recall test.

	df	*F*	*p*	η_p^2
Training Condition	2	1.81	.167	.027
Input Modality	1	.214	.644	.002
Spatial Ability	1	22.6	.000	.147
Training Condition x Input Modality	2	1.69	.189	.025
Training Condition x Spatial Ability	2	.441	.644	.007
Training Condition x Spatial Ability	1	.447	.505	.003
Error	131			
Total	143			

<u>Disassembly</u>

For the disassembly test, a 3 (multimedia training condition) x 2 (input modality) by 2 (spatial ability) ANOVA was conducted for the number of steps completed as well as completions time. Effect sizes, means, and standard deviations were reported.

Steps Completed

For multimedia training condition, a statistically significant main effect for spatial ability was found, $F(1,131) = 19.7$, $p < .001$, $\eta_p^2 = .129$. High spatials ($M = 11.6$, $SD = 2.02$) completed more steps than low spatials ($M = 9.63$, $SD = 3.44$). No other statistically significant effects were found rendering no support for any of the hypotheses. Table 15 contains a complete list of means and standard deviations for all conditions by spatial ability.

Table 14

ANOVA table for disassembly steps completed.

	df	*F*	*p*	η_p^2
Training Condition	2	2.00	.140	.030
Input Modality	1	1.50	.223	.011
Spatial Ability	1	19.8	.000	.131
Training Condition * Input Modality	2	.202	.817	.003
Training Condition* Spatial Ability	2	.646	.526	.010
Training Condition * Spatial Ability	1	.010	.920	.000
Error	131			
Total	143			

Completion Time

For multimedia training condition, a statistically significant main effect for spatial ability was found, F $(1,131) = 24.7$, $p < .001$, $\eta_p^2 = .157$. High spatials (M = 334sec, SD = 150) completed the disassembly faster than low spatials (M = 449sec, SD = 130). No other statistically significant effects were found. Table 16 contains a complete list of means and standard deviations for all conditions by spatial ability.

Table 15

ANOVA table for disassembly completion time.

	df	*F*	*p*	η_p^2
Training Condition	2	.612	.544	.009
Input Modality	1	.001	.972	.000
Spatial Ability	1	24.3	.000	.157
Training Condition x Input Modality	2	.809	.447	.012
Training Condition x Spatial Ability	2	.692	.502	.010
Training Condition x Spatial Ability	1	.066	.798	.001
Error	131			
Total	143			

Table 16

Means and Standard Deviations for all Conditions by Spatial Ability for Number of Steps Recalled.

Input	Mouse				Touch				Total			
Spatial Ability	High		Low		High		Low		High		Low	
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RDS	9.62	3.01	6.40	2.76	8.40	3.16	6.78	3.39	8.96	3.10	6.58	2.99
RDS + Arrows	8.25	3.32	7.25	2.74	10.86	2.11	6.50	2.32	9.91	2.84	6.96	2.57
RDS + Gestures	8.50	2.35	6.92	2.63	7.40	2.59	5.69	2.75	8.00	2.47	6.31	2.71
Total	8.88	2.85	6.92	2.66	9.03	2.98	6.25	2.77	8.96	2.90	6.62	2.71

Table 17

Means and Standard Deviations for all Conditions by Spatial Ability for Number of Disassembly Steps Completed.

Input	Mouse				Touch				Total			
Spatial Ability	High		Low		High		Low		High		Low	
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RDS	11.54	1.56	7.50	4.27	11.07	2.43	9.44	3.32	11.29	2.05	8.42	3.88
RDS + Arrows	10.38	3.20	10.81	2.56	12.50	0.94	9.00	2.83	11.73	2.25	10.12	2.76
RDS + Gestures	11.88	1.96	9.31	4.05	12.10	1.60	10.77	3.19	11.98	1.76	10.04	3.65
Total	11.38	2.2	9.46	3.73	11.85	1.87	9.84	3.12	11.63	2.03	9.63	3.45

Note. Out of 14

Reassembly

For the reassembly test, a 3 (multimedia training condition) x 2 (input modality) by 2 (spatial ability) ANOVA was conducted for the number of steps completed as well as completions time. Effect sizes, means, and standard deviations were reported.

Steps Completed

For multimedia training condition, a statistically significant main effect for spatial ability was found, F (1,131) = 21.5, $p < .001$, $\eta_p^2 = .139$. High spatials (M = 11.4, SD = 2.50) completed more reassembly steps than low spatials (M = 9.37sec, SD = 2.84). No other statistically significant effects were found. Table 17 contains a complete list of means and standard deviations for all conditions by spatial ability.

Table 18

ANOVA table for reassembly steps completed.

	df	*F*	*p*	η_p^2
Training Condition	2	1.88	.157	.028
Input Modality	1	.884	.349	.007
Spatial Ability	1	21.6	.000	.142
Training Condition x Input Modality	2	1.05	.352	.016
Training Condition x Spatial Ability	2	.160	.852	.002
Training Condition x Spatial Ability	1	2.57	.111	.019
Error	131			
Total	143			

Completion Time

For multimedia training condition, a statistically significant main effect for spatial ability was found, *F* (1,131) = 13.2, $p < .001$, $\eta_p^2 = .091$. High spatials (*M* = 454sec, *SD* = 125) completed the reassembly faster than low spatials (*M* = 521sec, *SD* = 99.0). No other statistically significant effects were found. Table 18 contains a complete list of means and standard deviations for all conditions by spatial ability.

Table 19

ANOVA table for reassembly completion time.

	df	*F*	*p*	η_p^2
Training Condition	2	1.44	.240	.022
Input Modality	1	1.04	.310	.008
Spatial Ability	1	13.7	.000	.095
Training Condition x Input Modality	2	.487	.615	.007
Training Condition x Spatial Ability	2	2.19	.116	.032
Training Condition x Spatial Ability	1	3.12	.080	.023
Error	131			
Total	143			

Table 20

Means and Standard Deviations for all Conditions by Spatial Ability for Disassembly Completion Time (in sec).

Input	Mouse				Touch				Total			
Spatial Ability	High		Low		High		Low		High		Low	
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RDS	327	149	458	132	389	159	470	166	360	155	463	145
RDS + Arrows	346	180	434	120	279	139	507	78.6	303	154	462	110
RDS + Gestures	338	150	456	119	327	135	392	152	333	140	424	137
Total	335	152	448	120	333	150	450	143	334	150	449	130

Table 21

Means and Standard Deviations for all Conditions by Spatial Ability for Number of Reassembly Steps Completed.

Input	Mouse				Touch				Total			
Spatial Ability	High		Low		High		Low		High		Low	
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RDS	10.65	2.85	7.65	3.47	10.90	2.74	10.11	2.30	10.79	2.74	8.82	3.16
RDS + Arrows	12.75	1.83	9.25	2.43	11.54	3.02	10.05	2.10	11.98	2.67	9.56	2.30
RDS + Gestures	11.54	1.83	9.46	3.26	11.60	1.96	9.69	3.13	11.57	1.84	9.58	3.13
Total	11.49	2.37	8.91	3.02	11.31	2.62	9.92	2.54	11.39	2.50	9.37	2.84

Note. Out of 14

Table 22

Means and Standard Deviations for all Conditions by Spatial Ability for Reassembly Completion Time (in sec).

Input	Mouse				Touch				Total			
Spatial Ability	High		Low		High		Low		High		Low	
Condition	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RDS	472	144	541	117	485	109	491	139	479	124	517	127
RDS + Arrows	360	117	557	54.1	433	157	490	89.5	406	145	531	75.7
RDS + Gestures	490	74.1	536	88.0	447	109	494	110	470	91.9	515	100
Total	451	124	546	83.1	456	127	492	110	454	125	521	99.0

Subjective Measures

The analyses for the subjective measures mirrored that of the performance measures whereby a 3 (multimedia training condition) x 2 (input modality) x 2 (spatial ability) ANOVA was conducted for each measure. Means and standard deviations for each measure are listed in Appendix O.

NASA TLX

To score the NASA TLX, the performance factor was reverse coded and summed with the other five factors to produce a single score of workload. A higher score indicated more perceived workload. A 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted for each administration of the NASA TLX which followed the multimedia training phase, disassembly task, and reassembly task.

Multimedia Training Phase

For the NASA TLX administered after the multimedia training, no statistically significant effects were found.

Table 23

ANOVA table for the NASA TLX following the multimedia training phase.

	df	*F*	*p*	η_p^2
Training Condition	2	.271	.763	.004
Input Modality	1	.046	.831	.000
Spatial Ability	1	.989	.322	.007
Training Condition x Input Modality	2	.053	.948	.001
Training Condition x Spatial Ability	2	.101	.904	.002
Training Condition x Spatial Ability	1	.013	.909	.000
Error	131			
Total	143			

Disassembly

For the NASA TLX administered after the disassembly task, a main effect was found for spatial ability, $F(1,133) = 9.33$, $p = .003$, $\eta_p^2 = .066$. High spatials ($M = 288$, $SD = 108$) reported less workload during disassembly than low spatials ($M = 347$, $SD = 113$). This result mirrors performance during the disassembly task whereby high spatials had better performance relative to low spatials. No other statistically significant effects were found.

Table 24

ANOVA table for the NASA TLX following the disassembly task.

	df	*F*	*p*	η_p^2
Training Condition	2	1.79	.171	.027
Input Modality	1	.033	.856	.000
Spatial Ability	1	9.22	.003	.066
Training Condition x Input Modality	2	.576	.564	.009
Training Condition x Spatial Ability	2	1.36	.259	.020
Training Condition x Spatial Ability	1	.889	.348	.007
Error	131			
Total	143			

Reassembly

For the NASA TLX administered after the reassembly task, no statistically significant effects were found.

Table 25

ANOVA table for the NASA TLX following the reassembly task.

	df	*F*	*p*	η_p^2
Training Condition	2	1.59	.207	.024
Input Modality	1	1.34	.250	.010
Spatial Ability	1	3.83	.053	.028
Training Condition x Input Modality	2	.709	.494	.011
Training Condition x Spatial Ability	2	2.49	.087	.037
Training Condition x Spatial Ability	1	.002	.961	.000
Error	131			
Total	143			

Measure of Cognitive Load

A 4 (Multimedia Training Condition) x 2 (Spatial Ability) ANOVA was conducted for each administration of the measure of cognitive load, which followed the multimedia training phase, disassembly task, and reassembly task.

Table 26

ANOVA table for the measure of cognitive load following the multimedia training phase.

	df	*F*	*p*	η_p^2
Training Condition	2	.272	.762	.004
Input Modality	1	.043	.837	.000
Spatial Ability	1	1.39	.241	.010
Training Condition x Input Modality	2	.342	.711	.005
Training Condition x Spatial Ability	2	.189	.828	.003
Training Condition x Spatial Ability	1	.312	.578	.002
Error	131			
Total	143			

Multimedia Training Phase

For the measure of cognitive load following the multimedia training phase, no statistically significant effects were found.

Disassembly

For the measure of cognitive load following the disassembly task, a main effect was found for spatial ability, $F(1,131) = 15.2$, $p < .001$, $\eta_p^2 = .102$. High spatials ($M = 3.79$, $SD =$

1.69) reported less cognitive load during disassembly than low spatials (*M* = 2.65, *SD* = 1.60). This result mirrors performance during the disassembly task whereby high spatials had better performance relative to low spatials. This result is in contrast to Experiment 1, which did not find significant differences between high and low spatials for their subjective rating of cognitive load. This may have been a result of only going through the multimedia training phase once, which added additional cognitive load particularly to low spatials. No other statistically significant effects were found.

Table 27

ANOVA table for the measure of cognitive load following the disassembly task.

	df	*F*	*p*	η_p^2
Training Condition	2	.833	.437	.013
Input Modality	1	.079	.779	.001
Spatial Ability	1	15.0	.000	.103
Training Condition x Input Modality	2	.128	.880	.002
Training Condition x Spatial Ability	2	.572	.566	.009
Training Condition x Spatial Ability	1	.285	.594	.002
Error	131			
Total	143			

Reassembly

For the measure of cognitive load following the reassembly task, no statistically significant effects were found.

Table 28

ANOVA table for the measure of cognitive load following the reassembly task.

	df	*F*	*p*	η_p^2
Training Condition	2	.312	.733	.005
Input Modality	1	.620	.433	.005
Spatial Ability	1	2.770	.098	.021
Training Condition x Input Modality	2	.026	.975	.000
Training Condition x Spatial Ability	2	.946	.391	.014
Training Condition x Spatial Ability	1	1.175	.280	.009
Error	131			
Total	143			

Usability

The usability measure was administered once right after the multimedia training phase. The mean score (out of 7) was recorded. A higher score indicated a better rated usability. There were no statically significant effects found.

Table 29

ANOVA table for the usability measure following the multimedia training condition.

	df	*F*	*p*	η_p^2
Training Condition	2	.787	.457	.012
Input Modality	1	.020	.887	.000
Spatial Ability	1	1.00	.319	.008
Training Condition x Input Modality	2	2.19	.116	.032
Training Condition x Spatial Ability	2	.257	.774	.004
Training Condition x Spatial Ability	1	.000	.996	.000
Error	131			
Total	143			

Motivation

The average response for the IMI (out of 7) was recorded. A higher score indicated higher motivation. The IMI was administered after the multimedia training phase. For the IMI, no statistically significant effects were found.

Table 30

ANOVA table for the IMI.

	df	*F*	*p*	η_p^2
Training Condition	2	.514	.599	.008
Input Modality	1	.200	.655	.002
Spatial Ability	1	1.10	.297	.008
Training Condition x Input Modality	2	.027	.974	.000
Training Condition x Spatial Ability	2	.677	.510	.010
Training Condition x Spatial Ability	1	.012	.914	.000
Error	131			
Total	143			

CHAPTER NINE: EXPERIMENT 2 DISCUSSION

Experiment 2 had two primary research questions. The first research question was whether incorporating gestures that mimicked real-world action would increase learning outcomes relative to a simple tap interface. The second research question asked whether directly touching the screen would increase learning outcomes of a procedural-motor task relative to indirectly touching the screen with a mouse. To accomplish this, Experiment 2 examined six different conditions including three different variations of the RDS (i.e. RDS, RDS + Arrows, RDS + Gestures) and two different types of input (i.e. direct touch and indirect mouse). Spatial ability was also considered as a subject variable. The RDS represented the least amount of gesture in that it only required tapping on components to initiate animations. The RDS + Gesture took the RDS and added a gesture component. After selecting a component, a gesture arrow appeared that required the participant to trace in order to initiate the animation sequence. The RDS + Arrow condition was included to ascertain any learning effects that were due strictly to the addition of the gesture arrow that was also included in the RDS + Gestures. In the RDS + Arrow condition, the gesture arrow appeared but was not used to initiate the animation. Instead, tapping the component again completed the step. Overall, the results were inconclusive as no significant differences were found between the conditions. However, consistent with Experiment 1, spatial ability was a key predictor in performance, such that high spatials significantly outperformed low spatials in all measures of performance. The results are discussed in terms of the specific hypotheses.

Effects of Gesture on Learning from Multimedia

Hypothesis 1 predicted that individuals in the RDS + Gestures condition would have higher learning outcomes than those in the RDS condition because of the sensorimotor aspect of learning according to Embodied Cognition. Utilizing gestures was thought to promote better encoding based on the body movements associated with learning the material. Unfortunately, no significant results could support this hypothesis. I believe that a lack of significance in this experiment does not necessarily indicate that gestures do not support learning. Countless research studies have found evidence to the contrary (e.g., Cook et al., 2009; Lozano & Tversky, 2006). Several factors could have contributed to the lack of significance. For one, the gestures may not have been meaningful to the simulated action on the rifle. Evidence suggests that if gestures are not meaningful or relevant to the content being learned, they will not support meaningful learning (Cook, Yip, & Goldin-Meadow, 2013). In this experiment, the gestures may not have been meaningful representations of the animations on the screen and therefore did not serve to help learn the information above and beyond the non-gesture conditions. This conjecture is supported by the fact that the RDS + Arrows condition did not significantly differ from the RDS condition on any measures of performance. If the arrows provided meaningful information to help encode the steps, one would expect this condition to be significantly higher in performance outcomes.

Another potential reason for the lack of significance may have been a possible ceiling effect. Fifty-seven percent of participants were able to dissemble at least thirteen of the fourteen total steps across all the conditions. Although this lends support to the conjecture that tablets can be effective tools to train this particular task, it does not help distinguish differences between conditions.

Effects of Input on Learning from Multimedia

Hypothesis 2 predicted that direct touch input would result in higher learning outcomes than indirect mouse input across all conditions. Further, Hypothesis 2 (a) predicted that this effect would be most pronounced in the RDS + Gestures condition. Unfortunately, no significant results supported this hypothesis. The current experiment was unable to address whether direct or indirect input would affect embodied learning with gestures. However, anecdotally, several participants mentioned that their hands occluded the screen while making their selections and gestures. This may have contributed to the lack of significance. The mouse condition, on the other hand, allowed selection and gestures to occur without occluding visual information on the screen. So while there may have been a benefit to touching the screen, this effect may have been precluded by the inability to see the visual information on the screen. Unfortunately, this anecdotal evidence was not further verified by the usability measure. In spite of the lack of significance, future tablet application designs should still consider this when designing an interface. Gestures, as much as possible, should avoid interfering with visual information. As with the other hypotheses, the lack of significant results may have been due to a ceiling effect.

Effects of Spatial Ability

Spatial ability was a significant predictor across all the performance measures in Experiment 2. These results are consistent with those from Experiment 1. High spatials were able to perform the tasks more effectively than low spatials. However, unlike Experiment 1 and the prediction in Hypothesis 1b, spatial ability did not interact with either the multimedia training condition or the input. The lack of a significant interaction could be explained by the fact that all of the conditions utilized high levels of interactivity and dynamism. In Experiment 1, it was noted that high spatials excelled in this type of multimedia environment because of their ability to handle higher levels of spatial information. It appears however, that the inclusion of gestures was ineffective at improving learning outcomes for low spatials in order to compensate for their lack of spatial ability in Experiment 2.

Theoretical and Practical Implications

Based on the results of Experiment 2, it is unclear how the sensorimotor processes elicited during the RDS + Gestures condition contributed to learning through Embodied Cognition. Simply tapping the screen may have been enough to create a sensorimotor coupling attached to the information and that the addition of gestures did not add anything. Similarly, the input by which individuals manipulated the content on the screen did not seem to matter either. Perhaps the movement associated with the mouse, which is universally ubiquitous with interacting with computer interfaces, triggered similar sensorimotor coupling as the direct touch input. Future research should reinvestigate the use of gestures in multimedia on tablets using different tasks and gestures before any conclusions should be made regarding incorporating an Embodied Cognition paradigm into instructional designs for tablets.

From a practical perspective, it is still unclear if the added benefits of gestures warrant the additional programming required to incorporate them. Adding complex gestures to an application requires more programming effort than a simpler tap interface. Once again, more research should investigate the merit of including gestures in training applications on tablets.

CHAPTER TEN: GENERAL DISCUSSION AND LIMITATIONS

The two experiments described in this dissertation addressed the gap in the literature as how interactive and touch features inherent to tablets affect training a procedural-motor task. Both experiments investigated this from the perspective of the interactivity afforded by tablets and learning from a touchscreen.

Review of Results

For interactivity, Experiment 1 leveraged Cognitive Load Theory and Cognitive Theory of Multimedia Learning as a theoretical basis to determine the effectiveness of interactivity and dynamism to train a procedural motor task. The primary research question for Experiment 1 asked whether multiple forms of interactivity increase learning outcomes. Results indicated that more interactivity associated with dynamic multimedia was able to increase learning outcomes relative to less interactive multimedia. However, this result was strongly dependent on the spatial ability of the individual. Low spatials suffered in the 2D phase diagram and RDS conditions. On the other hand, high spatials excelled in the high levels of interactivity provided to them in the RDS condition and were able to compensate for the low levels of interactivity and animation in the 2D phase diagram condition. The secondary research question for Experiment 1 asked whether static or dynamic multimedia was superior in training a procedural-motor task. Results indicated that for learning to disassemble a Colt M4, dynamic multimedia was superior to equivalent static multimedia.

Experiment 2 examined the touchscreen in terms of touch input and gestures. Specifically, Experiment 2 considered two primary research questions related to the touchscreen: the touch component, and the gesture component. The touch component addressed whether touching the screen promoted better learning than indirectly manipulating the screen via a mouse. The gesture component questioned if gestures, under an Embodied Cognition paradigm, promoted learning better than a simple gesture-less tap interface. Results from Experiment 2 were inconclusive due primarily to a potential ceiling effect. However, consistent with Experiment 1, spatial ability was a large predictor of performance.

Theoretical Implications

The two experiments taken together addressed whether interactivity and gestures could promote learning outcomes on a procedural-motor task using a tablet using Cognitive Load Theory, Cognitive Theory of Multimedia Learning, and Embodied Cognition as a theoretical basis. Although Cognitive Load Theory, Cognitive Theory of Multimedia Learning are represented separately in the literature from Embodied Cognition, these two theories could be integrated to help advance the instructional design principles established by Cognitive Load Theory and Cognitive Theory of Multimedia Learning. For instance, Wong et al. (2009) and Ayers et al. (2009) alluded to a motor channel in addition to the auditory and pictorial channels described in Cognitive Load Theory and Cognitive Theory of Multimedia Learning. Embodied Cognition provides a sound basis of explanation why motor processes should be considered as a third channel. Embodied Cognition posits the coupling of the sensorimotor system to learning (Barsalou, 2005). This includes not only vision and hearing, but motor processes as well. This has ramifications for future research. For example, how does incorporating movement and gesture to Cognitive Theory of Multimedia Learning affect how selecting, organizing, integrating occur? Currently, the model only accounts for integrating images and sounds that are split in two channels as seen in Figure 1. A third motor channel via gestures should also be integrated into the model to incorporate the tenants of Embodied Cognition. However, under Cognitive Theory of Multimedia Learning, it is unclear how selecting, organizing and integrating would be affected by a third channel. The results from both experiments suggest that simply moving ones arm towards the screen was enough to elicit learning outcomes, specifically for high spatials. Furthermore, the additional movement may not be necessary possibly due to extraneous processing.

A lack of understanding regarding how gestures fit into Cognitive Load Theory and Cognitive Theory of Multimedia Learning may have been central to the lack of findings in Experiment 2. Both Cognitive Theory of Multimedia Learning and Cognitive Load Theory assert that not overloading either channel (auditory or pictorial) facilitates learning by reducing cognitive overload. However, Cognitive Theory of Multimedia Learning notes that words and pictures must be integrated, which is a source of processing. It could be the case that adding a gesture component that is prompted by the instructional system required the individual to not only integrate images and sounds, but also gestures as well. Furthermore, if the gestures were not indicative of movements associated with the learners' existing schema, integrating the gesture with what is seen and heard may have placed extraneous processing demands on the learner causing potential cognitive overload. However, even if the gestures did line up appropriately with an individual's mental model, the additional processing required to integrate with the visual and pictorial stimuli may cause extraneous processing.

Another theoretical challenge in both experiments was the inability to identify which processing demands were affected by the multimedia presentations, which precluded any specific conclusions about the source of load in each experiment. For instance, Experiment 1 found differences in overall cognitive load for the multimedia training, but no indications if it was from extraneous load or germane load. In Experiment 2, differences in overall cognitive load were found only between high and low spatials for the disassembly task. Even though instructional designs on tablets should aim to address the loads and processing spelled out by Cognitive Load Theory and Cognitive Theory of Multimedia Learning, the inability to measure them after the fact makes further improvements to a system challenging even if they were designed to reduce extraneous load and increase germane load.

Practical Implications

Overall, the results support the use of tablets to train the disassembly procedures of a Colt M4 due to the overall high success rate of the transfer of training tasks. Although the ceiling effect in Experiment 2 precluded finding specific differences between input and gesture conditions, it did serve to indicate the effectiveness of the application to train the disassembly task.

In addition to interactivity and input, spatial ability proved to be a large predictor of performance on the task in both experiments. In an ideal, learner-centric situation, spatial ability could be taken into consideration when assigning which features to include in a multimedia training based on the individual. This could be accomplished by a series of spatial tests that would assign the appropriate training. Alternatively, given the ability to access tablet applications anywhere, low spatials could practice additional repetitions until mastery. Both experiments limited the number of times participants could run through the training. This limitation would not exist in real-world applications.

In terms of the Army's desire to incorporate tablets into training through ALM 2015, the two Experiments support their use and integration. It should be stated that applications, such as the one used for this dissertation, should not replace live training. However, specific training applications could be used in several instances. For example, tablet applications could be used as pre-trainers based on the Pretraining Principle that states people learn better from training when they already know the names and characteristics of essential components (Mayer, 2008). As an applied example, incoming recruits could be given a series of applications to download on their tablet that would prepare them prior to live training. Live training is costly to the Army, which was a major motivation for ALM 2015, and tablets offer an opportunity to make live training more effective. Furthermore, tablet applications could be used after live training to go over any topics that were not sufficiently acquired. A soldier who did not quite understand the disassembly procedures after a classroom session could revisit the information via a tablet application on their own time. Similarly, applications such as this could be used as "refresher training" to go over forgotten procedures with quick run-through on the tablet application. To this end, tablets can be used to potentially provide a supplemental curriculum, utilizing interactive and dynamic features, to train applied Army tasks. Once again, this aligns with a major goal of ALM 2015, which was to provide a more learner-centric curriculum for soldiers. Overall, the results of this study support incorporating tablet training as part of ALM 2015 so long as the individual differences of the learner are considered.

Limitations and Future Research

As with all research, the two experiments were met with limitations. The most apparent limitation was due to the lack of complexity of the two transfer of training tasks which resulted in a ceiling effect. To combat this, Experiment 2 reduced the number of times participants viewed the training multimedia from twice to once which still resulted in a ceiling effect. Even though all of the interactive conditions across both experiments resulted in ceiling effects, the multimedia training application proved to be an overall effective way to train novices the disassembly procedures of a M4 carbine in spite of the ability to discern difference between the conditions.

Another potential reason why the transfer of training tasks had such a high success rate could be that participants were given too long to perform the tasks. Participants were able to "figure out" how to perform the steps simply by tinkering with the rifle. Future studies should impose a more difficult time constraint that requires a substantial understanding of the material to finish the task. Alternatively, a more complex task could be utilized. Another limiting factor may have been the population from which participants were recruited. The experiments in this dissertation relied exclusively on college students and may not have been representative of the targeted soldier demographic. Unfortunately, no soldiers were available to participate in the study.

The two experiments described an initial attempt to examine the unique features inherent to tablets, interactivity and the touchscreen. Additional research should further this line of work by systematically investigating other types of interactivity. The current study only investigated a very specific procedural-motor task using controlling and manipulating as the interactivity. Other studies may investigate the effects of dialoguing, searching and navigating on other types of tasks and domains (Mayer & Moreno, 2003). Different domains may require different types of interactivity to elicit the highest learning outcomes. Furthermore, different levels of intrinsic load may also mediate the effectiveness of different types of interactivity.

Although this dissertation did not find any results stemming from gesture or touch, future research should continue to investigate the relationship between motor processes and learning. The current research was unsuccessful in demonstrating the effects of Embodied Cognition on learning through tablets, however this should not preclude future studies from examining different ways to incorporate the sensorimotor system into instructional designs for tablet.

Conclusion

Overall, this area of research will continue to become increasingly important as mobile devices and touchscreens become more and more pervasive and integrated into education and training. In line with ALM 2015, tablets, if utilized appropriately, can be an effective way to push anytime, anywhere training to Soldiers in the U.S. Army, and other branches. The technology offers interactivity that can facilitate learning from multimedia. The experiments conducted for this dissertation should be used to guide future instructional design elements by promoting the use of interactivity and gestures in tablet applications for training.

APPENDIX A: DEMOGRAPHICS

Age__________

Gender__________

Are you colorblind?________

Do you have normal or corrected to normal vision?____________

Highest level of education completed____________

Do you own a smartphone?__________

If yes, what kind of smartphone do you own (e.g., iPhone 5, Samsung Galaxy SII)

____

Do you own a tablet?________

If yes, what kind (e.g., iPad2, Asus Transformer Prime)?__________________

Please enter the average or typical number of hours per week that you use a computer:

Where do you currently use a computer? Please select all that apply:
- Home
- Work
- Library or Learning Center
- Other

Do you own a personal computer?:
- Yes
- No

How often do you play computer games?:
- Daily
- Weekly
- Monthly
- Less than once a month
- Never

How often do you play video games (run on a console, not a computer)?:
- Daily
- Weekly
- Monthly
- Less than once a month
- Never

How often do you use graphics or drawing features in software packages?: *

Daily
Weekly
Monthly
Less than once a month
Never

How often do you use email (at home or work)?: *

Daily
Weekly
Monthly
Less than once a month
Never

How often do you use the internet (not including email or gaming)?:

Daily
Weekly
Monthly
Less than once a month
Never

How much do you enjoy playing video games (either computer or console)? :

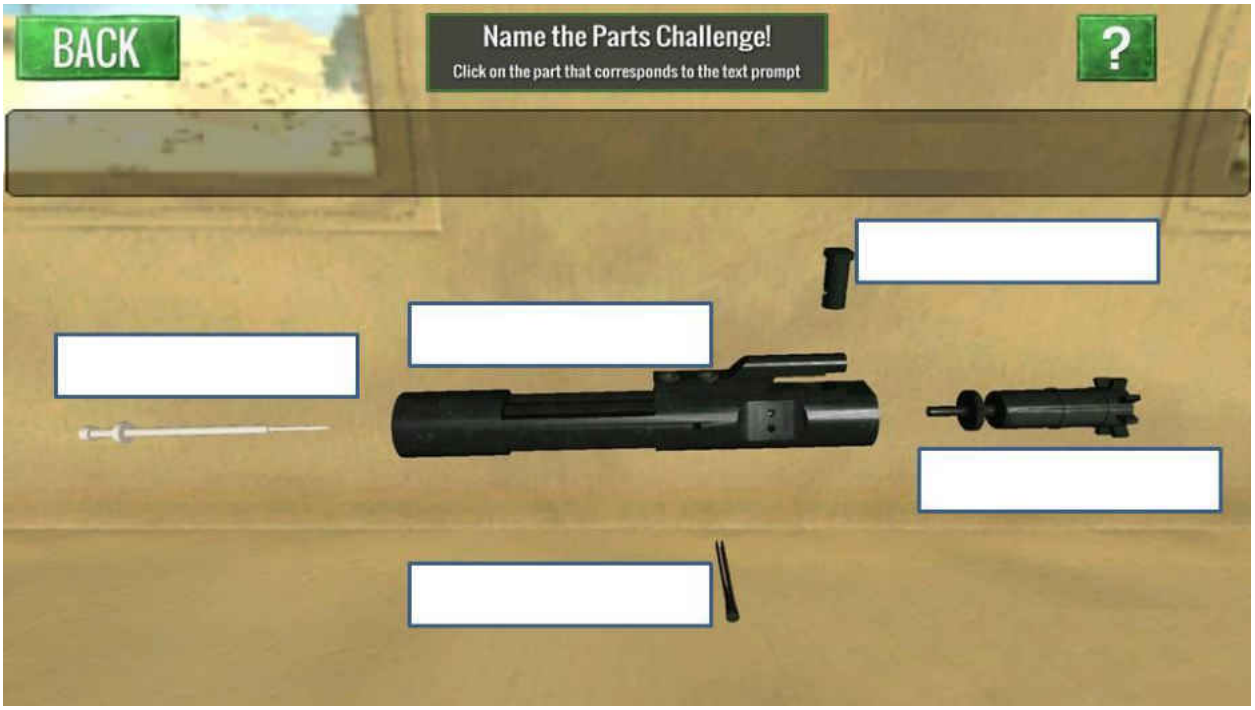
Not very much
Somewhat
Average enjoyment
A lot of fun
Most Fun in Life

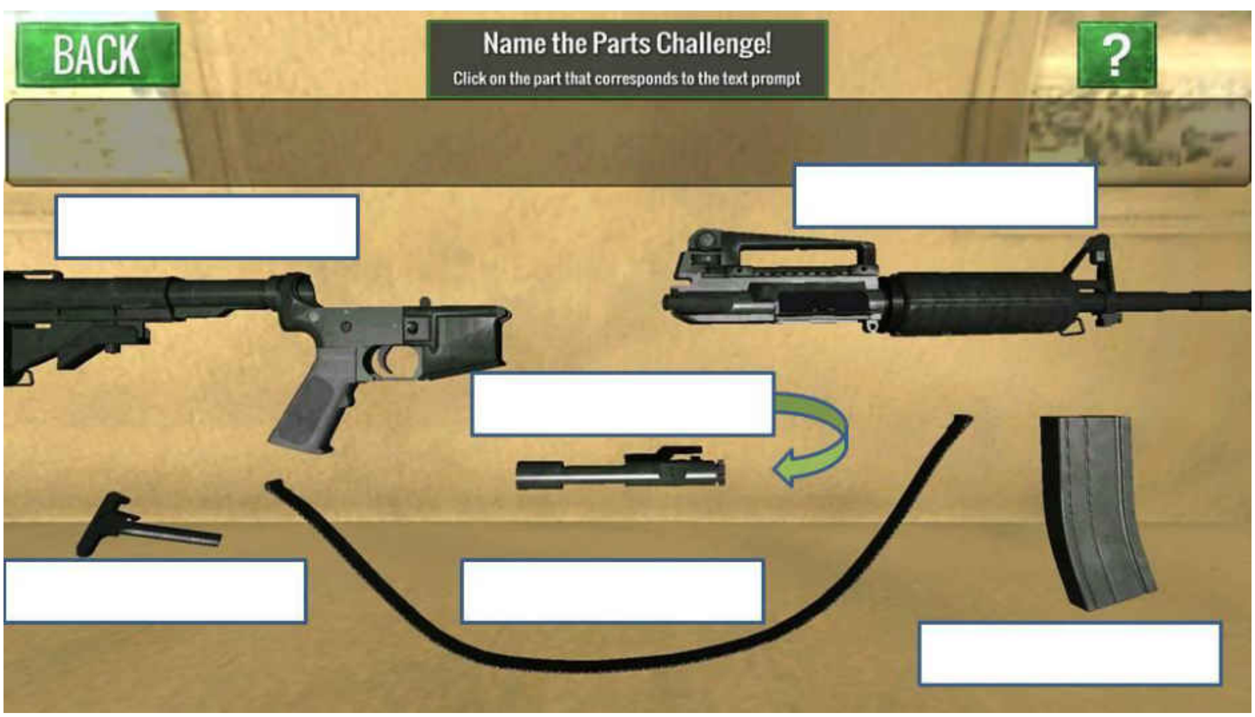
Please rate your skill at playing video games:

Bad
Poor
Average
Better than Average
Good

Please enter the number of hours per week that you play video games. Please enter whole digits, e.g. 8 for eight hours:

APPENDIX B: NOMENCLATURE QUIZ


BACK
Name the Parts Challenge!
Click on the part that corresponds to the text prompt
?


BACK
Name the Parts Challenge!
Click on the part that corresponds to the text prompt
?

APPENDIX C: FIREARM EXPERIENCE SURVEY

Do you have any military experience? Select all that apply.

- ☐ Army
- ☐ Navy
- ☐ Air Force
- ☐ Marines
- ☐ ROTC
- ☐ Law Enforcement
- ☐ National Guard

Are you a member of any of the following? Select all that apply.

- ☐ National Rifle Association
- ☐ Gun Club
- ☐ Sportsmen's Club

Do you have any formal training (e.g., military or law enforcement) in the use and operation of a firearm?

- ○ Yes
- ○ No

Do you have a concealed weapons permit?

- ○ Yes
- ○ No

Do you, or have you ever owned any of the following?

- ☐ Long Rifle
- ☐ Shotgun
- ☐ Pistol

Regardless of ownership, have you ever handled (held and fired), any of the following?

- ☐ Long Rifle
- ☐ Shotgun
- ☐ Pistol

Regardless of ownership, have you ever disassembled or assembled any of the following?

- ☐ Long Rifle
- ☐ Shotgun
- ☐ Pistol

Home comfortable were you handling the replica rifle?

- ○ Not at all
- ○
- ○
- ○
- ○
- ○
- ○ Completely

APPENDIX D: CARD ROTATIONS TEST

Name ____________________

CARD ROTATIONS TEST — S-1 (Rev.)

This is a test of your ability to see differences in figures. Look at the 5 triangle-shaped cards drawn below.

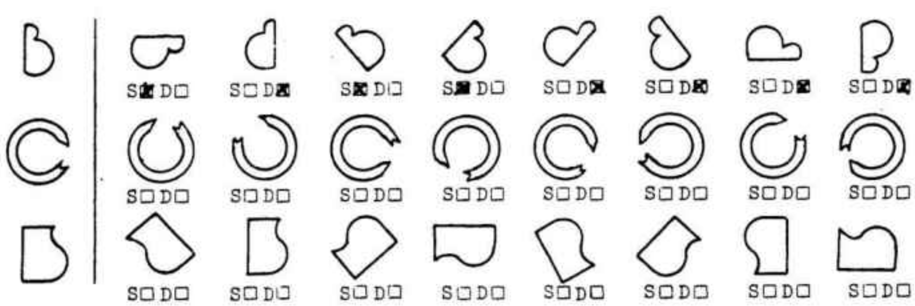
All of these drawings are of the same card, which has been slid around into different positions on the page.

Now look at the 2 cards below:

These two cards are not alike. The first cannot be made to look like the second by sliding it around on the page. It would have to be flipped over or made differently.

Each problem in this test consists of one card on the left of a vertical line and eight cards on the right. You are to decide whether each of the eight cards on the right is the same as or different from the card at the left. Mark the box beside the S if it is the same as the one at the beginning of the row. Mark the box beside the D if it is different from the one at the beginning of the row.

Practice on the following rows. The first row has been correctly marked for you.

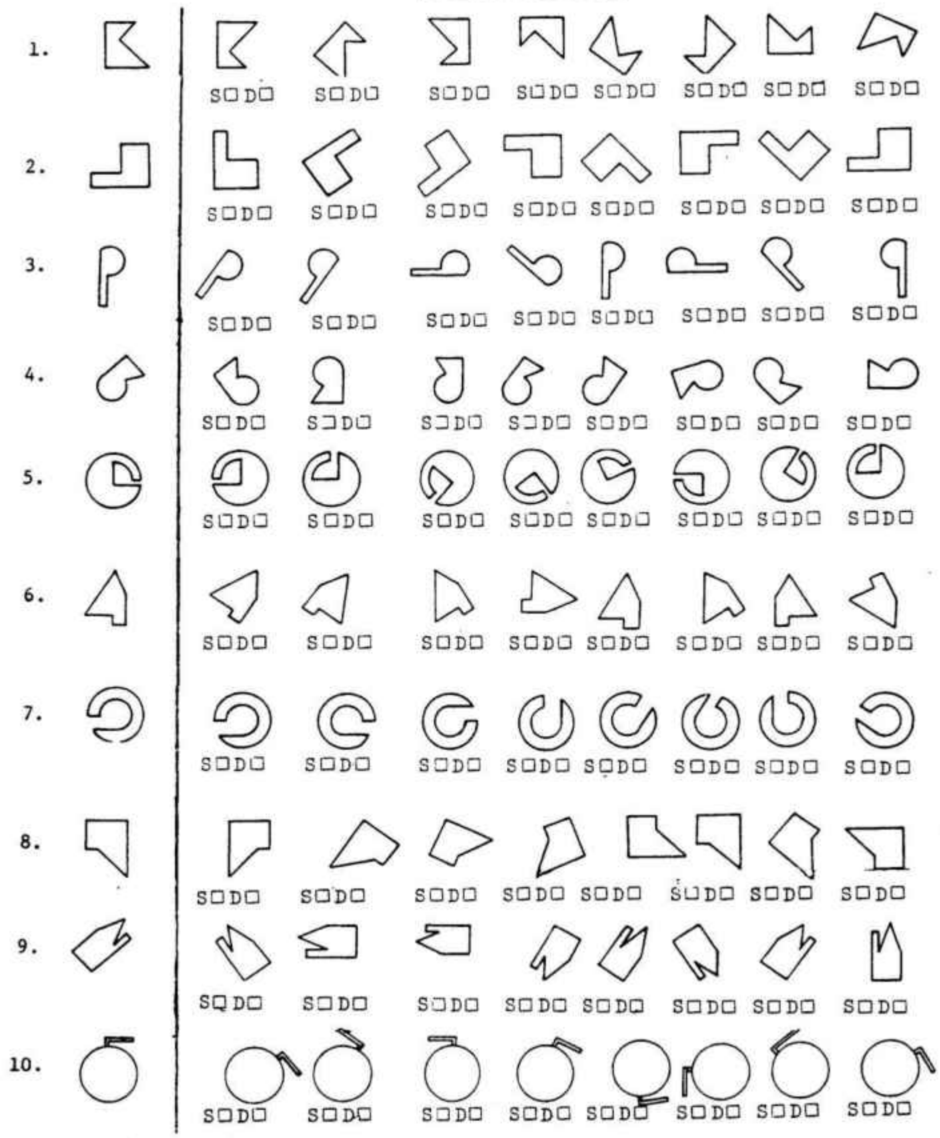


Your score on this test will be the number of items answered correctly minus the number answered incorrectly. Therefore, it will not be to your advantage to guess, unless you have some idea whether the card is the same or different. Work as quickly as you can without sacrificing accuracy.

You will have 3 minutes for each of the two parts of this test. Each part has 1 page. When you have finished Part 1, STOP. Please do not go on to Part 2 until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

Part 1 (3 minutes)



DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

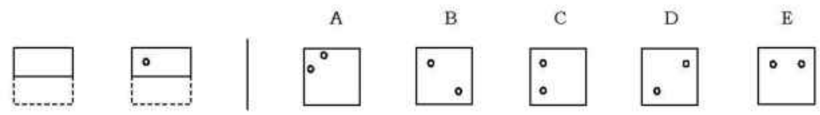
STOP.

APPENDIX E: PAPER FOLDING TEST

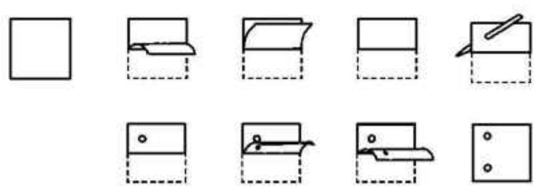
Paper Folding Test—Vz-2-BRACE

In this test you are to imagine the folding and unfolding of pieces of paper. In each problem in the test there are some figures drawn at the left of a vertical line and there are others drawn at the right of the line. The figures at the left represent a square piece of paper being folded, and the last of these figures has one or two small circles drawn on it to show where the paper has been punched. Each hole is punched through all the thicknesses of paper at that point. One of the five figures on the right of the vertical line shows where the holes will be when the paper is completely unfolded. You are to decide which one of these figures is correct and draw an X through that figure.

Now try the sample problem below. (In this problem only one hole was punched in the folded paper).



The correct answer to the sample problem above is C and so it should have been marked with an X. The figures below show how the paper was folded and why C is the correct answer.



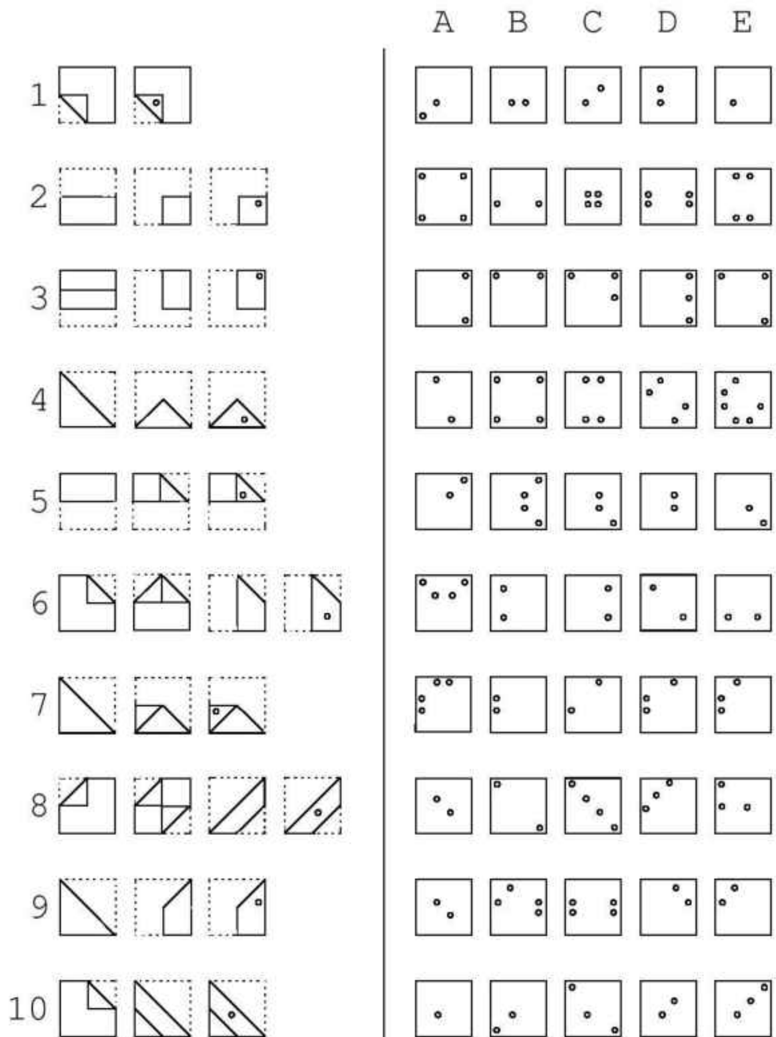
In these problems all of the folds that are made are shown in the figures at the left of the line, and the paper is not turned or moved in any way except to make the folds shown in the figures. Remember, the answer is the figure that shows the positions of the holes when the paper is completely unfolded.

Some of the problems on this sheet are more difficult than others. If you are unable to do one of the problems, simply skip over it and go on to the next one.

You will have three minutes for each of the two parts of this test. Each part has one page. When you have finished Part One, STOP. Please do not go on to Part Two until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

PART ONE (3 MINUTES)



STOP

DO NOT PROCEED TO THE NEXT PAGE UNTIL ASKED TO DO SO

APPENDIX F: NASA TLX

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

APPENDIX G: THE COMPUTER SYSTEM USABILITY QUESTIONNAIRE

Administration and Scoring. Use the CSUQ rather than the PSSUQ when the usability study is in a non-laboratory setting. Appendix Table 1 contains the rules for calculating the CSUQ and PSSUQ scores.

Appendix Table 1. Rules for Calculating CSUQ/PSSUQ Scores

Score Name	Average the Responses to:
OVERALL	Items 1 through 19
SYSUSE	Items 1 through 8
INFOQUAL	Items 9 through 15
INTERQUAL	Items 16 through 18

Average the scores from the appropriate items to obtain the scale scores. Low scores are better than high scores due to the anchors used in the 7-point scales. If a participant does not answer an item or marks "N/A," then average the remaining item scores.
Instructions and Items. The questionnaire's instructions and items are:

This questionnaire (which starts on the following page) gives you an opportunity to express your satisfaction with the usability of your primary computer system. Your responses will help us understand what aspects of the system you are particularly concerned about and the aspects that satisfy you. To as great a degree as possible, think about all the tasks that you have done with the system while you answer these questions.
Please read each statement and indicate how strongly you agree or disagree with the statement by circling a number on the scale. If a statement does not apply to you, circle N/A.

Thank you!

1. Overall, I am satisfied with how easy it is to use this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

2. It is simple to use this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

3. I feel comfortable using this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

4. It was easy to learn to use this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

5. Whenever I make a mistake using the system, I recover easily and quickly.

STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

6. It is easy to find the information I need.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

7. The information provided with the system is easy to understand.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

8. The information is effective in helping me complete my work.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

9. The organization of information on the system screens is clear.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

Note: *The interface includes those items that you use to interact with the system. For example, some components of the interface are the keyboard, the mouse, the screens (including their use of graphics and language).*
10. The interface of this system is pleasant.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

11. I like using the interface of this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

12. This system has all the functions and capabilities I expect it to have.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

13. Overall, I am satisfied with this system.
STRONGLY STRONGLY
AGREE 1 2 3 4 5 6 7 DISAGREE

APPENDIX H: INTRINSIC MOTIVATION INVENTORY

For each of the following statements, please indicate how true it is for you.

Interest/Enjoyment

	not true at all			somewhat true			very true
I enjoyed doing this activity very much.	○	○	○	○	○	○	○
This activity was fun to do.	○	○	○	○	○	○	○
I though this was a boring activity.	○	○	○	○	○	○	○
This activity did not hold my attention at all.	○	○	○	○	○	○	○
I would describe this activity as very interesting.	○	○	○	○	○	○	○
I thought this activity was quite enjoyable.	○	○	○	○	○	○	○
While I was doing this activity, I was thinking about how much I enjoyed it	○	○	○	○	○	○	○

Effort/Importance

	not true at all			somewhat true			very true
I put a lot of effort into this.	○	○	○	○	○	○	○
I didn't try very hard to do well at this activity.	○	○	○	○	○	○	○
I tried very hard on this activity.	○	○	○	○	○	○	○
It was important to me to do well at this task.	○	○	○	○	○	○	○
I didn't put much energy into this.	○	○	○	○	○	○	○

APPENDIX I: RETENTION TEST CODIING INSTURCTIONS

Correct Number of Steps:

Point awarded for each listed step if…

- Step is clear and understandable &
- Step is correct

No point given if…

- Rater cannot understand step
- Step is incorrect
- Steps for the bolt assembly occur prior to finishing the exterior steps
- Exterior steps are listed after describing steps towards the bolt assembly

APPENDIX J: STEPS FOR DISASSEMBLY AND REASSEMBLY TASKS

Step	Disassembly Steps: Exterior	Completed
1	Remove Strap	
2	Remove Handguards	
3	Remove Takedown Pin	
4	Remove Receiver Pin and	
5	Separate Upper and Lower Receivers	
6	Remove Carrying Handle	
7	Pull back Charging Handle	
8	Remove Bolt carrier and Charging Handle	
	Disassembly Steps: Interior Bolt Assembly	
9	Remove Firing Pin retaining pin	
10	Push in Bolt Assembly to locked position	
11	Take Firing Pin out	
12	Turn Bolt Cam Pin 1/4 turn	
13	Take Cam Pin out	
14	Remove Bolt Assembly	
	Assembly Steps: Interior Bolt Assembly	
1	Slide Bolt back into Carrier	
2	Insert Bolt Cam	
3	Turn Bolt Cam Pin 1/4 Turn	
4	Drop Firing Pin back into Carrier	
5	Pull Bolt out	
6	Insert Firing Pin Retaining Pin into bolt carrier	
7	Slide Charging Handle into Upper Receiver	
8	Slide Bolt Assembly into Upper Reciever	
	Assembly Steps: Exterior	
9	Replace Carrying Handle	
10	Place Upper and Lower Receiver together	
11	Insert Pivot Pin	
12	Replace Takedown Pin	
13	Replace Handguards	
14	Replace Strap	

Note: "Clearing the Weapon" was also included for the recall test coding under disassembly.

APPENDIX K: MEANS AND STANDARD DEVIATIONS FOR EXPERIMENT 1 SUBJECTIVE MEASURES

Table K1

NASA TLX: Multimedia Training Condition

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	33.9	13.5	17
	High	34.0	14.2	10
	Total	33.9	13.5	27
Non-Interactive Animation	Low	35.1	19.3	15
	High	33.7	11.1	13
	Total	34.4	15.7	28
Interactive Animation	Low	30.5	15.8	10
	High	28.3	16.2	16
	Total	29.1	15.7	26
RDS	Low	32.3	14.6	14
	High	29.9	13.6	14
	Total	31.1	13.9	28
Total	Low	33.2	15.5	56
	High	31.1	13.8	53
	Total	32.2	14.7	109

Table K2

NASA TLX: Disassembly Task

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	58.08	15.99	17
	High	51.45	21.74	10
	Total	55.62	18.21	27
Non-Interactive Animation	Low	53.38	20.15	15
	High	55.19	17.00	13
	Total	54.22	18.43	28
Interactive Animation	Low	58.18	20.17	10
	High	45.95	18.06	16
	Total	50.65	19.47	26
RDS	Low	53.01	19.70	14
	High	52.40	23.67	14
	Total	52.71	21.37	28
Total	Low	55.57	18.50	56
	High	50.96	19.89	53
	Total	53.33	19.24	109

Table K3

NASA TLX: Reassembly Task.

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	56.07	17.10	17
	High	55.57	23.27	10
	Total	55.88	19.17	27
Non-Interactive Animation	Low	60.33	18.83	15
	High	47.83	19.55	13
	Total	54.53	19.85	28
Interactive Animation	Low	49.98	19.97	10
	High	52.44	22.26	16
	Total	51.49	21.03	26
RDS	Low	59.92	18.81	14
	High	54.63	18.26	14
	Total	57.27	18.38	28
Total	Low	57.09	18.40	56
	High	52.48	20.41	53
	Total	54.85	19.45	109

Table K4

Measure of Cognitive Load: Multimedia Training Phase

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	3.94	1.60	17
	High	3.40	1.26	10
	Total	3.74	1.48	27
Non-Interactive Animation	Low	3.87	1.19	15
	High	3.92	1.12	13
	Total	3.89	1.13	28
Interactive Animation	Low	4.10	1.37	10
	High	5.06	1.34	16
	Total	4.69	1.41	26
RDS	Low	4.57	1.28	14
	High	4.86	1.23	14
	Total	4.71	1.24	28
Total	Low	4.11	1.37	56
	High	4.42	1.38	53
	Total	4.26	1.38	109

Table K5

Measure of Cognitive Load: Disassembly.

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	2.88	1.80	17
	High	3.00	1.49	10
	Total	2.93	1.66	27
Non-Interactive Animation	Low	3.47	1.81	15
	High	3.23	1.88	13
	Total	3.36	1.81	28
Interactive Animation	Low	2.40	1.84	10
	High	4.13	1.60	15
	Total	3.44	1.87	25
RDS	Low	3.71	1.82	14
	High	3.86	2.11	14
	Total	3.79	1.93	28
Total	Low	3.16	1.83	56
	High	3.62	1.81	52
	Total	3.38	1.82	108

Table K6

Measure of Cognitive Load: Reassembly.

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	2.82	1.29	17
	High	2.90	1.60	10
	Total	2.85	1.38	27
Non-Interactive Animation	Low	3.07	1.33	15
	High	3.92	1.85	13
	Total	3.46	1.62	28
Interactive Animation	Low	4.20	1.32	10
	High	3.25	1.39	16
	Total	3.62	1.42	26
RDS	Low	2.36	1.08	14
	High	3.64	1.55	14
	Total	3.00	1.47	28
Total	Low	3.02	1.37	56
	High	3.45	1.59	53
	Total	3.23	1.49	109

Table K7

Usability Means (out of 7).

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	4.99	1.45	17
	High	6.09	1.11	10
	Total	5.40	1.42	27
Non-Interactive Animation	Low	5.61	0.83	15
	High	5.94	0.85	13
	Total	5.76	0.84	28
Interactive Animation	Low	6.04	0.87	10
	High	5.83	0.81	16
	Total	5.91	0.82	26
RDS	Low	5.78	0.98	14
	High	6.07	0.81	14
	Total	5.92	0.89	28
Total	Low	5.54	1.13	56
	High	5.97	0.86	53
	Total	5.75	1.03	109

Table K8

IMI (out of 7).

Condition	Spatial Ability	*M*	*SD*	*n*
2D Phase Diagram	Low	5.29	1.21	17
	High	4.59	0.77	10
	Total	5.03	1.11	27
Non-Interactive Animation	Low	4.82	0.98	15
	High	5.27	0.91	13
	Total	5.02	0.96	28
Interactive Animation	Low	4.87	0.89	10
	High	5.40	0.86	16
	Total	5.20	0.89	26
RDS	Low	5.47	1.08	14
	High	5.28	0.69	14
	Total	5.38	0.90	28
Total	Low	5.13	1.08	56
	High	5.18	0.85	53
	Total	5.16	0.97	109

APPENDIX L: SUPPLEMENTAL ANALYSIS FOR EXPERIMENT 1

Supplemental Analysis of Gender

Although Gender was not initially predicted to have an effect on any performance measures, during experimentation, researchers noted anecdotal differences in performance between males and females. To that end, a 2 (gender) x 4 (multimedia training condition) ANOVA was conducted for each of the performance variables. Although spatial ability did not significantly differ between genders in Experiment 1, it was not included in analysis due to previous research indicating spatial ability differences between the genders (Voyer, Voyer, & Bryden, 1995).

Recall Test

For the recall test, a significant main effect for gender was found, $F\ (1,101) = 18.164$, $p < .001$, $\eta_p^2 = .152$, such that males ($M = 10.44$, $SD = 2.62$) recalled more correct steps than females ($M = 8.24$, $SD = 2.89$). A statistically significant main effect for condition was also found, $F(3,101) = 2.70$, $p = .050$, $\eta_p^2 = .074$. However, follow up post-hoc tests using Tukey's HSD failed to reach significance. There was no statistically significant gender by condition interaction.

Disassembly

For the disassembly task, a 2 (gender) x 4 (multimedia training condition) ANOVA was conducted for both the number of steps completed and completion time.

Steps Completed

For number of steps completed, a a statistically significant main effect for gender was found, $F(1,101) = 22.85$, $p < .001$, $\eta_p^2 = .184$. such that males ($M = 12.65$, 2.13) completed more steps than females ($M = 10.11$, $SD = 3.59$). A statistically significant main effect for condition was also found, $F(3,101) = 4.64$, $p = .004$, $\eta_p^2 = .121$. Follow-up post-hoc tests using Tukey's HSD mirrored the results in the primary analysis such that the RDS significantly differed from both the phase diagram condition ($p = .006$) and the non-interactive animation condition ($p = .032$). More steps were completed for individuals in the RDS condition ($M = 12.82$, $SD = 2.07$) than the non-interactive animation condition ($M = 10.75$, $SD = 3.78$), and the 2D phase diagram condition ($M = 10.39$, $SD = 3.54$). No significant gender by condition interaction was found.

Completion Time

For completion time, a statistically significant main effect was found gender, $F(1,101) = 17.643$, $p < .001$, $\eta_p^2 = .117$, such that males completed the disassembly task faster ($M = 323$sec, $SD = 127$) than females ($M = 432$sec $SD = 147$). No statistically significant main effect for condition or condition by gender interaction was found.

Reassembly

For the reassembly task, a 2 (gender) x 4 (multimedia training condition) ANOVA was conducted for both the number of steps completed and completion time.

Steps Completed

For number of reassembly steps completed, a significant main effect for gender was found, $F(1,101) = 24.487$, $p < .001$, $\eta_p^2 = .195$, such that males completed more steps ($M = 10.04$, $SD = 2.27$) than females ($M = 7.91$, $SD = 2.31$). No statistically significant main effect for condition was found nor was a significant gender by condition interaction.

Completion Time

For completion time, a statistically significant main effect for gender was found, $F(1,101) = 21.521$, $p < .001$, $\eta_p^2 = .176$, such that males ($M = 467$sec $SD = 116$) completed the reassembly task faster than females ($M = 549$sec, $SD = 81.7$). A statistically significant main effect for condition was also found, $F(3,101) = 6.949$, $p < .001$, $\eta_p^2 = .171$, was also found. Follow-up post-hoc tests using Tukey's HSD revealed that the 2D phase diagram condition significantly differed from both the interactive animation condition ($p = .001$) and the RDS ($p = .001$). Completion times were faster in the interactive animation condition ($M = 475$sec, $SD = 107$) and the RDS ($M = 471$sec, $SD = 107$) as compared to the 2D phase diagram condition ($M = 571$sec, $SD = 67.3$). These results mirrored those found in the primary analysis. No significant gender by condition interaction was found.

Exploratory Analysis of Gender

Two steps were taken to further investigate the strong gender effect. First, correlations were conducted, to find variables that were significantly related to task performance. Second, independent-samples t-tests, were conducted to see if the genders significantly differed in any of those variables that correlated with performance. Four individual difference measures were selected that were thought to be related to performance. First was comfort handling the replica rifle which was measures on a 7-point likert (anchored with "not at all" and "completely") scale which asked participants how comfortable they were handling the replica rifle. A higher rating indicated a higher comfort level. Another variable selected was a self-rated skill with first-person shooters (FPS). Using a 7-point likert scale ("Very Bad," "Bad," "Poor." "Neither Good nor Bad," "Fair," "Good," "Very Good") participants were asked their level of skill at playing FPS. A higher score indicated a higher level of skill. The last two individual difference variables selected were tablet and smartphone ownership. As seen in Table 11, comfort with the replica rifle and FPS skill were significantly correlated with all performance variables. Smartphone, and tablet ownership did not significantly correlate with performance. However, tablet ownership was significantly correlated with disassembly performance (steps and time) in the RDS condition. Additionally, FPS experience was correlated with comfort, $r(109) = .341$, $p < .001$.

Table 31

Correlation table: performance variables by individual difference measures.

Performance Variable	Comfort	FPS Skill	Tablet Ownership	Smartphone Ownership
Recall Test	.351**	.311**	.038	-.016
Disassembly: Steps	.427**	.280**	-.096†	-.039
Disassembly: Time	-.443**	-.291**	.048†	-.101
Reassembly: Steps	.216*	.337**	-.023	.066
Reassembly: Time	-.218*	-.303**	-.001	-.103

Note. *$p < .05$. **$p < .01$. $^{\dagger}p < .05$ in RDS condition.

Follow up t-tests with the variables that significantly correlated with performance using gender as the independent variable were conducted. Statistically significant gender differences were found for both comfort, $t(107) = 3.37$, $p = .001$, $d = .65$, and FPS skill, $t(107) = 10.75$, $p < .001$, $d = 2.07$. Females ($M = 4.46$, $SD = 2.18$) rated themselves as being less comfortable at handling the replica rifle than males ($M = 5.65$, $SD = 1.44$). Furthermore, females ($M = 2.63$, $SD = 1.67$) rated themselves as having less skill in FPS than males ($M = 5.58$, $SD = 1.15$). No significant differences were found for tablet ownership in the RDS condition. These findings are consistent with the predictor variables in that comfort and FPS skill correlated with performance and females rated themselves lower in both variables.

APPENDIX M: MEANS AND STANDARD DEVIATIONS FOR EXPERIMENT 2 SUBJECTIVE MEASURES

Table M1

NASA TLX: Multimedia Training

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	41.10	15.37	10
		High	33.67	11.56	13
		Total	36.90	13.55	23
	Touch	Low	36.56	9.53	9
		High	37.29	20.67	15
		Total	37.01	17.08	24
	Total	Low	38.95	12.80	19
		High	35.61	16.86	28
		Total	36.96	15.29	47
RDS + Arrows	Mouse	Low	42.02	22.38	16
		High	36.90	19.74	8
		Total	40.31	21.24	24
	Touch	Low	41.27	15.94	10
		High	38.52	11.11	14
		Total	39.67	13.08	24
	Total	Low	41.73	19.80	26
		High	37.93	14.38	22
		Total	39.99	17.45	48
RDS + Gestures	Mouse	Low	36.97	15.32	13
		High	42.15	13.69	12
		Total	39.46	14.50	25
	Touch	Low	41.44	17.37	13
		High	34.15	14.83	10
		Total	38.27	16.38	23
	Total	Low	39.21	16.21	26
		High	38.52	14.46	22
		Total	38.89	15.27	48
Total	Mouse	Low	40.10	18.25	39
		High	37.54	14.63	33
		Total	38.93	16.62	72
	Touch	Low	40.01	14.79	32
		High	36.93	15.96	39
		Total	38.32	15.41	71
	Total	Low	40.06	16.66	71
		High	37.21	15.26	72
		Total	38.62	15.98	143

Table M2

NASA TLX: Disassembly.

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	54.45	21.00	10
		High	45.58	17.86	13
		Total	49.43	19.36	23
	Touch	Low	49.30	20.73	9
		High	46.11	18.87	15
		Total	47.31	19.20	24
	Total	Low	52.01	20.46	19
		High	45.86	18.07	28
		Total	48.35	19.10	47
RDS + Arrows	Mouse	Low	56.74	19.72	16
		High	47.02	20.54	8
		Total	53.50	20.10	24
	Touch	Low	69.30	12.27	10
		High	45.15	18.46	14
		Total	55.22	19.99	24
	Total	Low	61.57	18.07	26
		High	45.83	18.78	22
		Total	54.36	19.85	48
RDS + Gestures	Mouse	Low	56.88	10.03	13
		High	55.72	14.05	12
		Total	56.33	11.88	25
	Touch	Low	60.09	23.65	13
		High	49.88	19.66	10
		Total	55.65	22.13	23
	Total	Low	58.49	17.87	26
		High	53.07	16.67	22
		Total	56.00	17.36	48
Total	Mouse	Low	56.20	17.05	39
		High	49.62	17.38	33
		Total	53.18	17.40	72
	Touch	Low	59.93	20.79	32
		High	46.74	18.52	39
		Total	52.68	20.53	71
	Total	Low	57.88	18.78	71
		High	48.06	17.94	72
		Total	52.93	18.95	143

Table M3

NASA TLX: Reassembly.

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	54.75	16.61	10
		High	53.55	13.79	13
		Total	54.07	14.73	23
	Touch	Low	45.57	21.92	9
		High	45.92	16.55	15
		Total	45.79	18.27	24
	Total	Low	50.40	19.33	19
		High	49.46	15.54	28
		Total	49.84	16.97	47
RDS + Arrows	Mouse	Low	60.82	21.86	16
		High	43.10	17.28	8
		Total	54.92	21.80	24
	Touch	Low	59.48	16.76	10
		High	45.62	19.88	14
		Total	51.40	19.54	24
	Total	Low	60.31	19.70	26
		High	44.70	18.59	22
		Total	53.16	20.56	48
RDS + Gestures	Mouse	Low	57.33	8.20	13
		High	58.61	20.07	12
		Total	57.95	14.79	25
	Touch	Low	57.62	22.84	13
		High	52.58	14.06	10
		Total	55.43	19.29	23
	Total	Low	57.47	16.82	26
		High	55.87	17.47	22
		Total	56.74	16.95	48
Total	Mouse	Low	58.10	16.78	39
		High	52.86	17.65	33
		Total	55.70	17.26	72
	Touch	Low	54.81	21.04	32
		High	47.52	17.09	39
		Total	50.81	19.18	71
	Total	Low	56.62	18.75	71
		High	49.97	17.43	72
		Total	53.27	18.34	143

Table M4

Measure of Cognitive Load: Multimedia Training

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	3.20	1.40	10
		High	4.15	0.99	13
		Total	3.74	1.25	23
	Touch	Low	3.56	1.33	9
		High	3.33	1.35	15
		Total	3.42	1.32	24
	Total	Low	3.37	1.34	19
		High	3.71	1.24	28
		Total	3.57	1.28	47
RDS + Arrows	Mouse	Low	3.63	1.82	16
		High	3.63	1.06	8
		Total	3.63	1.58	24
	Touch	Low	3.40	0.84	10
		High	4.21	1.85	14
		Total	3.88	1.54	24
	Total	Low	3.54	1.50	26
		High	4.00	1.60	22
		Total	3.75	1.55	48
RDS + Gestures	Mouse	Low	3.92	1.19	13
		High	3.42	1.31	12
		Total	3.68	1.25	25
	Touch	Low	3.54	1.39	13
		High	4.20	1.62	10
		Total	3.83	1.50	23
	Total	Low	3.73	1.28	26
		High	3.77	1.48	22
		Total	3.75	1.36	48
Total	Mouse	Low	3.62	1.52	39
		High	3.76	1.15	33
		Total	3.68	1.35	72
	Touch	Low	3.50	1.19	32
		High	3.87	1.63	39
		Total	3.70	1.45	71
	Total	Low	3.56	1.37	71
		High	3.82	1.42	72
		Total	3.69	1.40	143

Table M5

Measure of Cognitive Load: Disassembly.

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	2.70	1.95	10
		High	3.85	1.91	13
		Total	3.35	1.97	23
	Touch	Low	3.44	2.07	9
		High	3.67	1.72	15
		Total	3.58	1.82	24
	Total	Low	3.05	1.99	19
		High	3.75	1.78	28
		Total	3.47	1.87	47
RDS + Arrows	Mouse	Low	2.81	1.52	16
		High	3.13	1.36	8
		Total	2.92	1.44	24
	Touch	Low	1.70	0.48	10
		High	4.21	1.53	14
		Total	3.17	1.74	24
	Total	Low	2.38	1.33	26
		High	3.82	1.53	22
		Total	3.04	1.58	48
RDS + Gestures	Mouse	Low	2.54	1.33	13
		High	3.92	1.73	12
		Total	3.20	1.66	25
	Touch	Low	2.69	1.80	13
		High	3.70	1.95	10
		Total	3.13	1.89	23
	Total	Low	2.62	1.55	26
		High	3.82	1.79	22
		Total	3.17	1.75	48
Total	Mouse	Low	2.69	1.54	39
		High	3.70	1.70	33
		Total	3.15	1.68	72
	Touch	Low	2.59	1.70	32
		High	3.87	1.69	39
		Total	3.30	1.80	71
	Total	Low	2.65	1.60	71
		High	3.79	1.69	72
		Total	3.22	1.74	143

Table M6

Measure of Cognitive Load: Reassembly.

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	3.30	1.70	10
		High	3.23	1.42	13
		Total	3.26	1.51	23
	Touch	Low	3.56	1.94	9
		High	3.53	1.60	15
		Total	3.54	1.69	24
	Total	Low	3.42	1.77	19
		High	3.39	1.50	28
		Total	3.40	1.60	47
RDS + Arrows	Mouse	Low	2.94	1.57	16
		High	4.13	1.25	8
		Total	3.33	1.55	24
	Touch	Low	3.50	1.65	10
		High	3.93	1.54	14
		Total	3.75	1.57	24
	Total	Low	3.15	1.59	26
		High	4.00	1.41	22
		Total	3.54	1.56	48
RDS + Gestures	Mouse	Low	2.85	0.90	13
		High	3.83	1.47	12
		Total	3.32	1.28	25
	Touch	Low	3.46	1.45	13
		High	3.50	1.27	10
		Total	3.48	1.34	23
	Total	Low	3.15	1.22	26
		High	3.68	1.36	22
		Total	3.40	1.30	48
Total	Mouse	Low	3.00	1.40	39
		High	3.67	1.41	33
		Total	3.31	1.43	72
	Touch	Low	3.50	1.61	32
		High	3.67	1.47	39
		Total	3.59	1.53	71
	Total	Low	3.23	1.50	71
		High	3.67	1.43	72
		Total	3.45	1.48	143

Table M7

Usability (out of 7)

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	5.28	1.10	10
		High	5.78	0.73	13
		Total	5.57	0.93	23
	Touch	Low	5.27	1.47	9
		High	4.73	1.66	15
		Total	4.93	1.59	24
	Total	Low	5.28	1.26	19
		High	5.22	1.40	28
		Total	5.24	1.33	47
RDS + Arrows	Mouse	Low	5.11	1.50	16
		High	4.54	1.18	8
		Total	4.92	1.40	24
	Touch	Low	5.29	0.97	10
		High	5.43	0.96	14
		Total	5.37	0.95	24
	Total	Low	5.18	1.30	26
		High	5.11	1.11	22
		Total	5.15	1.20	48
RDS + Gestures	Mouse	Low	5.64	1.00	13
		High	5.09	1.14	12
		Total	5.38	1.08	25
	Touch	Low	5.56	1.09	13
		High	5.35	1.27	10
		Total	5.47	1.15	23
	Total	Low	5.60	1.03	26
		High	5.21	1.18	22
		Total	5.42	1.10	48
Total	Mouse	Low	5.33	1.24	39
		High	5.23	1.10	33
		Total	5.29	1.17	72
	Touch	Low	5.40	1.14	32
		High	5.14	1.35	39
		Total	5.25	1.26	71
	Total	Low	5.36	1.19	71
		High	5.18	1.23	72
		Total	5.27	1.21	143

Table M8

IMI (out of 7)

Training Condition	Input	Spatial Ability	*M*	*SD*	*n*
RDS	Mouse	Low	5.00	0.66	10
		High	5.47	0.84	13
		Total	5.27	0.79	23
	Touch	Low	4.62	1.39	9
		High	5.20	0.81	15
		Total	4.98	1.07	24
	Total	Low	4.82	1.05	19
		High	5.33	0.82	28
		Total	5.12	0.94	47
RDS + Arrows	Mouse	Low	4.92	1.08	16
		High	4.92	1.27	8
		Total	4.92	1.12	24
	Touch	Low	4.83	1.37	10
		High	4.98	1.09	14
		Total	4.92	1.18	24
	Total	Low	4.89	1.17	26
		High	4.95	1.13	22
		Total	4.92	1.14	48
RDS + Gestures	Mouse	Low	5.42	1.15	13
		High	5.19	0.89	12
		Total	5.31	1.02	25
	Touch	Low	5.08	1.19	13
		High	5.16	0.83	10
		Total	5.11	1.03	23
	Total	Low	5.25	1.16	26
		High	5.17	0.84	22
		Total	5.21	1.02	48
Total	Mouse	Low	5.11	1.02	39
		High	5.23	0.97	33
		Total	5.17	0.99	72
	Touch	Low	4.87	1.28	32
		High	5.11	0.91	39
		Total	5.00	1.09	71
	Total	Low	5.00	1.14	71
		High	5.17	0.93	72
		Total	5.08	1.04	143

APPENDIX N: SUPPLEMENTAL ANALYSIS FOR EXPERIMENT 2

Supplemental Analysis

As done in Experiment 1, gender, in lieu of spatial ability, was also investigated as an independent predictor of performance during the recall test, disassembly task and reassembly task. However, unlike Experiment 1, males (M = .276, SD = 1.09) had significantly higher, $t(140) = 3.52$, $p = .001$, $d = .594$ spatial ability than females (M = -.406, SD = 1.21). To that end, a 2 (gender) x 3 (multimedia training condition) x 2 (input) ANOVA was conducted for each of the performance variables.

Recall Test

For the recall test, a significant main effect for gender was found, $F(1,130) = 11.73$, $p = .001$, $\eta_p^2 = .083$, such that males (M = 8.61, SD = 2.97) recalled more correct steps than females (M = 6.93, SD = 2.89). No other statistically significant effects were found.

Disassembly

Steps Completed

For number of steps completed, a a statistically significant main effect for gender was found, $F(1,130) = 7.73$, $p = .006$, $\eta_p^2 = .056$. such that males (M = 11.3, SD = 2.43) completed more steps than females (M = 9.98, SD = 3.37). No other statistically significant effects were found.

Completion Time

For completion time, a statistically significant main effect was found gender, $F(1,130) = 7.70$. $p = .006$, $\eta_p^2 = .056$, such that males completed the disassembly task faster ($M = 356$sec, $SD = 157$) than females ($M = 423$sec $SD = 137$). No other statistically significant effects were found.

Reassembly

Steps Completed

For number of reassembly steps completed, a significant main effect for gender was found, $F(1,130) = 10.5$, $p = .001$, $\eta_p^2 = .075$, such that males completed more steps ($M = 11.1$, $SD = 2.62$) than females ($M = 9.64$, $SD = 2.91$). No other statistically significant effects were found.

Completion Time

For completion time, a statistically significant main effect for gender was found, $F(1,130) = 12.5$, $p < .001$, $\eta_p^2 = .176$, such that males ($M = 456$sec $SD = 129$) completed the reassembly task faster than females ($M = 519$sec, $SD = 94.2$). No other statistically significant effects were found.

Analysis of Tablet Interaction

In order to explore differences in how participants interacted with the tablet

Exploratory Analysis of Gender

As was done in Experiment 1, two steps were taken to further investigate the strong gender effect. First, correlations were conducted to find variables that were significantly related to task performance. Second, independent-samples t-tests, were conducted to see if the genders significantly differed in any of those variables that correlated with performance. The same four individual difference measures as Experiment 1 were selected that were thought to be related to performance. As seen in Table 14, comfort with the replica rifle and FPS skill were significantly correlated with all performance variables. Smartphone, and tablet ownership did not significantly correlate with performance which is surprising considering these variables significantly correlated with the RDS condition in Experiment 1. Additionally, FPS experience was correlated with comfort, $r(143) = .243$, $p = .004$.

Table 31

Correlation table: performance variables by individual difference measures.

Performance Variable	Comfort	FPS Skill	Tablet Ownership	Smartphone Ownership
Recall Test	.230**	.269**	-.007	.037
Disassembly: Steps	.182*	.268**	-.079	.052
Disassembly: Time	-.277**	-.299**	.102	-.026
Reassembly: Steps	.266**	.336**	.028	.026
Reassembly: Time	-.215*	-.263**	.021	-.016

Note. $^{*}p < .05$. $^{**}p < .01$.

Follow up t-tests with the variables that significantly correlated with performance using gender as the independent variable were conducted. The analysis revealed similar results to Experiment 1. Statistically significant gender differences were found for both comfort, $t(139) =$

2.36, $p = .020$, $d = .65$, and FPS skill, $t(139) = 6.39$, $p < .001$, $d = 2.07$. Females ($M = 4.61$, $SD = 2.10$) rated themselves as being less comfortable at handling the replica rifle than males ($M = 5.39$, $SD = 1.80$). Furthermore, females ($M = 3.20$, $SD = 1.85$) rated themselves as having less skill in FPS than males ($M = 5.07$, $SD = 1.61$). These findings are consistent with the predictor variables in that comfort and FPS skill correlated with performance and females rated themselves lower in both variables.

APPENDIX O: UCF IRB OUTCOME LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Matthew D. Marraffino** and Co-PI: **Glenn A. Martin**

Date: **June 24, 2013**

Dear Researcher:

On 6/24/2013, the IRB approved the following human participant research until 06/23/2014 inclusive:

Type of Review:	UCF Initial Review Submission Form Expedited Review Category # 7
Project Title:	Examining the Effects of Interactive Dynamic Multimedia on performance of a procedural motor task
Investigator:	Matthew D Marraffino
IRB Number:	SBE-13-09444
Funding Agency:	Army Research Institute(ARI)
Grant Title:	
Research ID:	1053672

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 06/23/2014, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a signed and dated copy of the consent form(s).

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 06/24/2013 10:02:05 AM EDT

Joanne Muratori

IRB Coordinator

REFERENCES

Alley, L. (2010, March). TRADOC releases new Warrior Tasks and Battle Drills. *www.army.mil*. Retrieved June 4, 2014, from http://www.army.mil/article/36496/tradoc-releases-new-warrior-tasks-and-battle-drills/

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps. *Journal of Educational Psychology*, *95*(4), 774.

Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *16*, 389–400.

Ayres, P., Kalyuga, S., Marcus, N., & Sweller, J. (2005, August). The conditions under which instructional animation may be effective. In *an International Workshop and Mini-conference, Open University of the Netherlands: Heerlen, The Netherlands*.

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior*, *25*(2), 348-353.

Ayres, P., & Paas, F. (2007). Can the cognitive load approach make instructional animations more effective?. *Applied Cognitive Psychology*, *21*(6), 811-820.

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning*, (pp. 135-146). New York, NY: Cambridge University Press.

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *The Journal of the Learning Sciences*, *11*(1), 25-61.

Baddeley, A. D. (1992). Working memory. *Science, 255*, 556–559.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation*, *8*, 47-89.

Baek, Y. K., & Layne, B. H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15*(4), 131-135.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577-660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645.

Barsalou, L. W., Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84-91.

Bassanesi, M., & Tindall, G. (2011). Heads Up, Heads Down: The iPad and its use in the General Aviation Cockpit. Retrieved from. http://www. aviationplatform.com/attachments/article/62/Bassanesi_Ipad.pdf

Beilock, S. L., & Goldin-Meadow, S. S. (2010). Gesture Changes Thought by Grounding It in Action*. Psychological Science, 21(11)*, 1605-1610.

Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning*, (pp. 287-296). New York, NY: Cambridge University Press.

Borghi, A. M., Scorolli, C., Caligiore, D., Baldassarre, G., & Tummolini, L. (2013) The embodied mind extended: Using words as social tools. *Frontiers in Psychology, 4*(214), 1-10.

Boucheix, J. M., & Guignard, H. (2005). What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, *20*(4), 369-388.

Brachman, L. (2003). Ronald J. Brachman, Hector J. Levesque. In *Knowledge representation and reasoning*.

Brill, J. C., Mouloua, M., Gilson, R. D., Rinalducci, E. J., & Kennedy, R. S. (2008). Effects of secondary loading task modality on attentional reserve capacity. In *Proceedings of the 52nd Human Factors and Ergonomics Society Annual Meeting* (18), 1219-1223.

Broaders, S. C., Cook, S., Mitchell, Z., & Goldin-Meadow, S. (2007). Making Children Gesture Brings Out Implicit Knowledge and Leads to Learning. *Journal Of Experimental Psychology. General, 136(4),* 539-550.

Bush, M. H., & Cameron, A. H. (2011). *Digital course materials: A case study of the Apple iPad in the academic environment*. (Doctoral Dissertation).

Byrne, M. D., Catrambone, R., & Stasko, J. T. (1999). Evaluating animations as student aids in learning computer algorithms. *Computers & Education*, *33*(4), 253-278.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-analytic Studies*. Cambridge University Press.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149-210.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*(2), 233-246.

Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. *Advances in the Psychology of Human Intelligence 1,*7–75.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal Of Experimental Psychology: General, 140(1),* 102-116.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047-1058.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*(2), 105-134.

de Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179-201.

de Koning, B. B., & Tabbers, H. K. (2011). Facilitating understanding of movements in dynamic visualizations: an embodied perspective. *Educational Psychology Review*, *23*(4), 501-521.

Dror, I. E., Kosslyn, S. M., & Waag, W. L. (1993). Visual-spatial abilities of pilots. *Journal of Applied Psychology, 78*(5), 763-773.

Edwards, L. D. (2009). Gestures and conceptual integration in mathematical talk. *Educational Studies in Mathematics*, *70*(2), 127-141.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests* (pp. 109-113). Princeton, NJ: Educational Testing Service.

Evans, C., & Gibbons, N. J. (2007). The interactivity effect in multimedia learning. *Computers & Education*, *49*(4), 1147-1160.

Evans, C., & Sabry, K. (2003). Evaluation of the interactivity of web-based learning systems: Principles and process. *Innovations in Education and Teaching International*, *40*(1), 89-99.

Ferguson, E. L., & Hegarty, M. (1995). Learning with real machines or diagrams: application of knowledge to real-world problems. *Cognition and Instruction*, *13*(1), 129-160.

Forlines, C., Wigdor, D., Shen, C., & Balakrishnan, R. (2007). Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 647-656). ACM.

Francaviglia, M., & Servidio, R. (2011). Gesture as a cognitive support to solve mathematical problems. *Psychology, 2(2),* 91-97.

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3-4), 455-479.

Garbarini, F., & Adenzato, M. (2004). At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition*, *56*(1), 100-106.

Gillespie, R., & O'Modrain, S. (2011). Embodied cognition as a motivating perspective for haptic interaction design: A position paper. In *World Haptics Conference*. Istanbul: IEEE.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*(4), 313-331.

Goksun, T., Goldin-Meadow, S., Newcombe, N., & Shipley, T. (2013). Individual differences in mental rotation: What does gesture tell us? *Cognitive Processing, (2),* 153-162.

Guay, F., Vallerand, R. J., & Blancdifficult, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*(3), 175-213.

Handy, B., Suter, T., & Hooper, B. K. (2011). Executive Summary: Oklahoma State University/Apple iPad Pilot Program.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, *1*, 139-183.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, *21*(6), 713-729.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301-307.

Hedman, L., Ström, P., Andersson, P., Kjellin, A., Wredmark, T., & Felländer-Tsai, L. (2006). High-level visual-spatial ability for novices correlates with performance in a visual-spatial complex surgical simulator task. *Surgical Endoscopy*, *20*(8), 1275-1280.

Hegarty, M. (2004). Dynamic visualizations and learning: getting to difficult questions. *Learning and Instruction*, *14*, 343-351.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, *21*(4), 209-249.

Höffler, T., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, *17*, 722-738.

Hostetter A. B., & Alibali M. W.(2008). Visible embodiment: Gestures as simulated action. Psychonomic. *Bulletin and Review, 15,* 495–514.

James, K. H., Humphrey, G. K., & Goodale, M. A. (2001). Manipulating and recognizing virtual objects: Where the action is. *Canadian Journal of Experimental Psychology, 55*(2), 111-120.

Jetter, H. C., Leifert, S., Gerken, J., Schubert, S., & Reiterer, H. (2012). Does (multi-) touch aid users' spatial memory and navigation in 'panning' and in 'zooming & panning' UIs?. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 83-90).

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*(2), 178.

Johnson, P. T., Zimmerman, S. L., Heath, D., Eng, J., Horton, K. M., Scott, W. W., & Fishman, E. K. (2012). The iPad as a mobile device for CT display and interpretation: Diagnostic accuracy for identification of pulmonary embolism. *Emergency Radiology*, *19*(4), 323-327.

Jones, M. G., Minogue, J., Tretter, T. R., Negishi, A., & Taylor, R. (2006). Haptic augmentation of science instruction: Does touch matter?. *Science Education*, *90*(1), 111-123.

Kaiser, M. K., Proffitt, D. R., & Anderson, K. (1985). Judgments of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 795.

Kalet, A. L., Song, H. S., Sarpel, U., Schwartz, R., Brenner, J., Ark, T. K., & Plass, J. (2012). Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. *Medical Teacher*, *34*(10), 833-839.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*(4), 509-539.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*,126–136.

Kalyuga, S., Chandler, P., & Sweller, J. (2001). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*, 5–23.

Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., & Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science*, *32*(7), 1099-1132.

Killilea, J. P., Marraffino, M. D., & Singer, M. J. (2013, January). Tablets for Institutional Army Education: Existing Skills & Expected Utility. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEmbodied Cognition)* (Vol. 2013, No. 1). National Training Systems Association.

Kin, K., Agrawala, M., & DeRose, T. (2009). Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. In *Proceedings of Graphics interface 2009* (pp. 119-124). Canadian Information Processing Society.

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, *65*(11), 911-930.

Kühl, T., Scheiter, K., Gerjets, P., & Edelmann, J. (2011). The influence of text modality on learning with static and dynamic visualizations. *Computers in Human Behavior*, *27*(1), 29-35.

Large, A., Beheshti, J., Breuleux, A., & Renaud, A. (1998). Effect of animation in enhancing descriptive and procedural texts in a multimedia learning environment. *Journal of the American Society for Information Science*, *47*(6), 437-448.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, *25*(6), 943-951.

Learning Technology Office. (2010, Oct 21). U.S. Army launches army learning concept 2015. . Retrieved from http://www.army.mil/article/46920/U_S__Army_launches_Army_Learning_Concept_2015/

Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction*, *13*(2), 177-189.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), 57-78.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*(2), 225-244.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, *13*(2003), 157-176.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*(3), 257-274.

MacKenzie, I. S., & Buxton, W. (1992). Extending Fitts' Law to two-dimensional tasks. In *ACM CHI'92: Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 219–226). New York, NY, USA: ACM.

Marmarelli, T., & Ringle, M. (2011). The Reed College iPad study.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning*, (pp. 31-48). New York, NY: Cambridge University Press.

Mayer, R.E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist, 63*(8), 760-769.

Mayer, R. E. (2009). *Cambridge handbook of multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages?. *Journal of Educational Psychology*, *93*(2), 390-397.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, *11*(4), 256-265.

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380-386.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43-52.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389-401.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48-58.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments. *Educational Psychology Review*, *19*, 309-326.

Paas, F., & Sweller, J. (2011). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*(1), 27-45.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design. *Educational Psychologist*, *38*(1), 1-4.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, *32*, 1-8.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63-71.

Paas, F., & Van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. *Learning and Instruction*, *16*(2), 87-91.

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*(4), 351-371.

Paivio, A. (1990). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.

Park, S. I., Lee, G., & Kim, M. (2009). Do students benefit equally from interactive computer simulations regardless of prior knowledge levels?. *Computers & Education*, *52*(3), 649-655.

Pew Internet. (2012, December). *Adult gadget ownership over time*. Retrieved from http://www.pewinternet.org/Trend-Data-(Adults)/Device-Ownership.aspx

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2011). The mechanics of embodiment: A dialog on embodiment and computational modeling. *Frontiers in Psychology*, *2*(5), 1-21.

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education*, *21*(1), 31-61.

Ploetzner, R., & Lowe, R. (2004). Dynamic visualisations and learning. *Learning and Instruction*, *14*(3), 235-240.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction*, *16*(1), 1-10.

Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Technology Research and Development*, *38*(1), 77-86.

Rossing, J. P., Miller, W. M., Cecil, A. K., & Stamper, S. E. (2012). iLearning: The Future of Higher Education? Student Perceptions on Learning with Mobile Tablets. *Journal of the Scholarship of Teaching and Learning*, *12*(2), 1-26.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43,* 450-461.

Ryan, R. M., Mims V., & Koestner, R. (1983). The relationship of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology. 45*, 736-750.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, *19*(4), 469-508.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, *14*(3), 293-305.

Sears, A., & Shneiderman, B. (1991). High precision touchscreens: Design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, *34*(4), 593-613.

Sims, V. K., & Mayer, R. E. (2001). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, *16*(1), 97-115.

Strobel, J. (2010). *All the better to see you with: A comparison of approaches to delivering instructions for Lego construction tasks* (Doctoral dissertation, Bowling Green State University).

Stull, A. T., Hegarty, M., & Mayer, R. E. (2009). Getting a handle on learning anatomy with interactive three-dimensional graphics. *Journal of Educational Psychology*, *101*(4), 803-816.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Emerging Themes in Cognitive Load Theory: The Transient Information and the Collective Working Memory Effects. *Cognitive Load Theory*, 219-233.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*(3), 185-233.

Sweller, J., Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251-296.

Tan, D. S., Pausch, R., Stefanucci, J. K., & Proffitt, D. R. (2002). Kinesthetic cues aid spatial memory. In *CHI'02 extended abstracts on Human factors in computing systems* (pp. 806-807). ACM.

Tucker, J. S. (2010). Mobile learning approaches for U.S. Army Training. (ARI Research Note 2010-07). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate?. *International Journal of Human-Computer Studies*, *57*(4), 247-262.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*(2), 250-270.

Wang, P. Y., Vaughn, B. K., & Liu, M. (2011). The impact of animation interactivity on novices' learning of introductory statistics. *Computers & Education*, *56*(1), 300-311.

Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in a small scale assembly task?. *Learning and Instruction*, *20*(1), 84-93.

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *The American Journal of Psychology*. *114*(14), 591-600

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*(1), 77-94.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625-636.

Wisher, R. A., Sabol, M. A., & Ellis, J. A. (1999). Staying sharp: Retention of military knowledge and skills. (Special Report 39). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G. A., Paas, F., & Sweller, J. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior*, *25*(2), 339-347.

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, *21*(3), 317-331.

Zwaan, R. A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, *28*(1), 81-88.