2018

Investigating the Effectiveness of Using Part-Task or Whole-Task Training Methods to Facilitate Mindful Abstraction in Uncertain Tasks

John Killilea
University of Central Florida

INVESTIGATING THE EFFECTIVENESS OF USING PART-TASK OR WHOLE-TASK TRAINING METHODS TO FACILITATE MINDFUL ABSTRACTION IN UNCERTAIN TASKS

by

JOHN P KILLILEA
B.S. University of Central Florida, 2007
M.S. University of Central Florida, 2012

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

Fall Term
2018

Major Professor: Glenn Martin

ABSTRACT

As the global landscape changes and powers rise and fall, the Contested, Degraded, and Operationally Limited (CDO) environment is likely to be the new normal going forward. Uncertainty variables, such as missing, false, or extra information characterize the CDO environment. A key focus of this dissertation is optimizing training for recognizing these uncertainty variables when training time is limited. This was investigated by either exposing participants to multiple uncertainty variables at a time with low doses of each (whole-task training), by exposing singular variables at a time with high doses (part-task training) or using no variables throughout training (control). A key motivator behind this research was Cognitive Load Theory, as mindful abstraction can only occur if there are cognitive resources to spare. Dependent variables, such as time to complete, number correct, task workload, and uncertainty variables identified, were collected.

The results revealed that on the transfer task, the part-task condition recorded a significantly lower workload score than the whole-task (and control) condition, while the condition's workload scores were consistent across all training and transfer tasks. In contrast, the control and whole-task condition experienced significant increases in workload during the transfer task. Additionally, the part-task condition participants were able to significantly identify more uncertainty variables on the final task than the whole-task condition and control condition. The part-task condition found the transfer task to be the "easiest" in terms of workload, and as there is more opportunity for mindful abstraction if there are more cognitive resources available, it can be stated that based on these results, the part-task training schedule facilitated mindful

abstraction more than the other two training schedules. As this was a limited, abstracted, and laboratory experiment, future work should apply the same methodology to applied tasks in a controlled environment to gauge further usefulness of this research.

ACKNOWLEDGMENTS

I would like to thank the members of my committee: Dr. Glenn Martin, Dr. Sae Schatz, Dr. Valerie Sims, and Dr. Atsusi Hirumi. Thank you for serving on my committee and providing valuable input throughout my graduate career. I know that being a part of a committee is time consuming and primarily volunteer work, so I do appreciate every minute of your time. Your guidance and encouragement throughout the process has made it that much better. Glenn was exactly the type of advisor that I needed for this. He pushed me when necessary, and was laid back when we were just trying to make progress and push on. It is hard to imagine going through this process without someone like him on my team. I will always be thankful for the time and effort he provided for these years.

Additionally, I'm thankful for the help of my friends, Dr. Matthew Marraffino and Dr. Joseph Mercado. Both were able to provide helpful insights and suggestions on various topics and were always willing to take a call or answer a text on short notice. Lastly, I would like to thank my family for the encouragement and support throughout the years. My parents have been my greatest fans throughout graduate school, and they have always encouraged and motivated me to strive for greatness. I honestly could not have done this without your support every step of the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Statement of the Problem

As societies continue the pursuit of advanced technological progress, additional problem spaces are being created first, yet adapted to, second. Relatively new problem spaces, such as the electromagnetic spectrum and cyberspace, have seen a rise in interest for those wanting to operate in them, as well as understand them. Although these problem spaces certainly occur in civilian circles, the military services have been especially focused on expanding their understanding, capability sets, and training for spaces that primarily deal with uncertainty. For example, in recent years, the U.S. Navy reinstated celestial navigation instruction almost two decades after it was determined "outdated." The concern is that trusted systems like GPS, may not always be available. As near-peer nations edge closer to parity with U.S. power, the focus of warfighter training requires modifications to adapt to a new type of environment. In contrast to the past two decades of warfighting dominance, future conflicts may pair the U.S. against a nation that could significantly counter U.S. supremacy. In the spirit of employing an approach previously known as the Anti-Access/Area-Denial (A2/AD) strategy, near-peer nations will also force the US to operate in Contested, Degraded, or Operationally Limited (CDO) land, air, sea, and cyber battlespaces. These environments are characterized by trusted information systems displaying conflicting or inaccurate information due to enemy electronic or cyber intrusion. Additionally, some sensors or communication systems may be destroyed or unavailable. The concern is real and growing, as some in the service have publicly acknowledged that the Navy is not optimized to support a CDO environment.

Unfortunately, real-world constraints have limited live training opportunities in the past decade, and simulator time is coveted now, more than ever. Many of the U.S. Navy's flight and tactical training simulators operate from dawn until well past dusk. The question becomes how to train for the CDO environment, when training time and resources are already limited and services are seeing tight constraints on throughput. Although it is difficult to predict or reduce the uncertainties of operating in contested environments, the Navy can train warfighters to adapt and be flexible in order to achieve end-state objectives. To this end, research focuses on optimizing cognitive readiness training, its psychological sub-components, and improving transfer of training methods to ensure training is efficient and effective (Morrison and Fletcher, 2002; Fletcher, 2004). However, in a resource constrained training environment, efficient and effective transfer becomes a challenging prospect. Previous research into transfer of training has yielded conflicting and often unsuccessful results (Perkins & Salomon, 1992). Transfer, the ability to apply knowledge, skills, and abilities from the training context to a performance context, is not guaranteed and often requires much effort to attain it. To further complicate matters, training to operate in CDO environments requires more than simple repetition or increased practice. This "low-road" transfer method is practical for procedural, well-practiced, closed skills; however, it is an ineffective and inefficient prospect to train for cognitive flexibility and higher order thinking required for dynamic, complex, and ambiguous environments (Salomon and Perkins, 1989). It focuses on training a narrow range of examples, and by design, slowly increases the range of training capabilities.

In contrast, training for the CDO battlespace requires “high-road” transfer of training, which trains mindful, conscious abstraction that is especially useful for training complex open skills (Salomon and Perkins, 1989; Salas, Milham, and Bowers, 2003). Through this, the trainee is able to knowingly abstract concepts and principles from training. Additionally, the trainee is able to bridge situations together and apply his or her trained knowledge and skills to novel situations without having explicitly experienced them before. Although this method can potentially expand the trainees experience envelope in leaps rather than small increments, it requires significant effort of the learner, and the design of this training can be challenging.

Fortunately, transfer of training research is prevalent throughout the past century, starting with psychologists such as Thorndike in the early 1900s. This original research, combined with more recent transfer of training models, such as the Baldwin and Ford (1988) model, can be leveraged to start answering some of the more challenging and applied questions. Critical questions such as: how to promote high-road transfer of training, in a simulator-type environment, when time and resources are limited? Or, what can training designers do to facilitate the mindful abstraction necessary to quickly broaden a trainee’s realm of possible and increase their cognitive flexibility?

Even with decades worth of research, there are still marked gaps in the literature, especially when focused on transferring open skills, or skills based on principles and rules rather than procedures and checklists. Open skills, like mindful abstraction, provide the trainees the freedom to perform, as there is not a hard and set method to accomplish the task as there is in the more procedural closed skills that have a standardized way of getting from point A to point B (Yelon

and Ford, 1999). Although accounting for all components of a transfer model, like Baldwin and Ford's, is complex, by taking a specific applied training task, it is feasible to rule out many sources of variability and focus on one or two aspects of the model to fill much needed training gaps. The next section describes this dissertation's approach to investigating one of the transfer of training design aspects, the conditions of practice, which can affect the transfer of open skills, like mindful abstraction.

Through identifying small subsections of the transfer of training models, and using sound previous research to support it, this research can provide insight and guidelines to improve training that focuses on improving cognitive readiness.

Purpose of the Current Study

This dissertation lays out the basic and applied research need that calls for more specific and empirical research into open skills transfer of training research. Specifically, this effort aims to fill a research gap relating to one facet of training input design. Baldwin and Ford's (1988) transfer of training model lists three primary training inputs: trainee characteristics, training design, and work environment. This document focuses heavily on the training design aspect of the model, and more specifically, the "conditions of practice" element, that focuses on how the training material is presented (Baldwin and Ford, 1988). For example, is massed or distributed training better for a particular task? How and when should feedback be presented? Or is part-task or whole-task more suitable for a specific task? The focus of this effort investigated the last question, whether part-task or whole-task training better facilitates mindful abstraction necessary

for high-road training transfer, which is currently vague in research literature. These issues are examined within the context of an individual trainee progressing through scenario-based training.

To investigate this issue, a review of relevant literature is warranted. A first look will examine the environment itself that the training will be designed for, the Contested, Degraded, or Operationally Limited (CDO) environment. As the global landscape changes and powers rise and fall, the CDO environment is likely to be the new normal going forward, rather than "traditional" warfare environments. Next, the overarching solution, improving the training of cognitive readiness will be reviewed. With that as an impetus, the transfer of training literature will be examined with a specific focus on models of transfer, predictors of transfer, and low- and high-road transfer theories. Finally, the specific matter of whether condition of practice variables, like part-task training and whole-task training, will be examined in terms of how each may or may not contribute to facilitating the transfer of open skills, like mindful abstraction. This section also explores the underlying theoretical mechanisms that account for the anticipated gains, such as mental capabilities, workload, and Cognitive Load Theory. Existing literature is copious for many of these topics, and as such, this review will place special emphasis on the role each contributes towards conditions of practice transfer. At the conclusion of the literature review, the experimental study conducted to explore the effect of using part-task or whole-task training versus a control condition for a transfer task is discussed in terms of methodology, results, and a discussion of the findings.

Empirical testing of this topic is proposed to reveal if one method of presenting the training is 1) superior to the other and 2) a significant improvement over the control condition. The

experiment is a between-subjects design with three conditions: two experimental groups for a comparative evaluation and a control group. The two experimental groups include a part-task training condition and a whole-task training condition. This provides insight into the question of how the training design, or condition or practice, facilitates mental abstraction, an open skill required for high-road transfer necessary for complex, dynamic, and ambiguous environments.

Three primary CDO characteristics were identified: missing information, extra information, and stimulated/fake (spoofed) information. These CDO characteristics will be known throughout this dissertation as “uncertainty variables.” Each condition will take part in three training scenarios and one final transfer task. There will be little instructed training as the majority of the “training” will be incidental and organic. The trainees will be exposed to the CDO uncertainty variables in a controlled manner. In other words, the interventions will be implicitly incorporated into the scenarios, rather than explicitly stated to the participants. Pre-training will involve minimal instruction on the task itself as the task, puzzle assembly, is a common task that many should already know. However, previous experience with the task, along with other person-related variables (e.g., IQ and spatial ability) will be collected.

The part-task condition will receive a singular uncertainty variable (CDO characteristic) designed into each scenario. For example, the first part-task scenario will only include missing information (missing puzzle pieces); the second scenario will only include spoofed information (fake puzzle pieces replacing some original puzzle pieces), and so on. The whole-task condition will include all three of the uncertainty variables, designed into each scenario, although in smaller doses of each. The control condition will simply involve completing the task without any

of the CDO uncertainty variables included until the transfer task. In summary, the control task condition receives no uncertainty variable exposure during training, the part-task condition receives a high dose of one at a time, while the whole-task condition receives exposure to all three uncertainty variables in each training task, albeit at lower doses.

The non-comparative evaluation, or comparing the part-task or whole-task conditions with the control group, will answer the question whether the intervention is effective at all. The comparative evaluation, between the part-task condition and the whole-task condition will assess the relative effectiveness of the two interventions with the same goal. It will answer the question of which intervention is more effective.

Further, the literature is not solidified on which method is optimized for the variables that characterize the CDO environment. The knowledge and skills needed to overcome the CDO-type uncertainty variables are primarily open skills that require high-road transfer of training. The core notion of high-road transfer of training requires the trainee to mentally abstract the concept each training scenario presents. Determining which method promotes mental flexibility and mindful abstraction could be a key way to optimize transfer of training when training open skills.

Ultimately, this effort intends to provide empirically-based recommendations for designing and conducting training with an emphasis on transfer for cognitive readiness components to CDO environments. Training for these types of environments is complex and challenging, so providing evidence that one method or another (part-task vs. whole-task) is stronger at facilitating the mental abstraction necessary for optimized open skills transfer is the principal objective. The

results of the study are presented in Chapter 4 and analyzed and synthesized to advise training design recommendations in Chapter 5. Although a specific sample task will be used for the study, the implications can make an impact that stretch beyond the task itself, and likely extend to other similar domains.

Research Questions

1) Does part-task or whole-task training significantly promote mental abstraction necessary for operating in a CDO environment?

   1a) Will whole-task training (providing all uncertainty variables during each scenario at a lower dose) promote mental abstraction leading to greater knowledge and skill transfer?

   1b) Will part-task training (providing one unique uncertainty variable per scenario at a higher dose) promote mental abstraction leading to greater knowledge and skill transfer?

The Sample Applied Task

This effort stems from a relatively specific training focus. As such, it is important to describe the proposed trainee and applied training environment.

The envisioned task is that of a sensor operator in the U.S. Navy, although each service has these types of positions. A sensor operator typically sits at a control station and monitors one or two different types of sensors, either aboard an aircraft, ship, or submarine. For example, the sensor operator on a P-8A aircraft, might monitor the search and surface RADAR, the electronic support measures (ESM) dealing with enemy electronic emissions, and the electro-optical infrared camera. While monitoring each of these sensors, he or she also must monitor internal

and external communications to provide information or receive information to keep his or her station up to date. These operators do much of their training in individual trainers (high-fidelity simulators) on naval bases. They spend much of their time learning their individual role, and then participate in crew training as well to work on communication and coordination.

If operating in a CDO environment, this sensor operator would be hindered by a myriad of challenges. Characteristics of the CDO environment include: missing information, false information, additional information, as well as having incomplete big-picture situational awareness due to other platforms encountering the same issues. Specifically, his or her radar picture might be incomplete due to unknown enemy systems either jamming it, or evading it though stealth means. Additionally, through psychological operations or via other practical methods, the adversary may falsely increase the quantity or size of their forces, adding extra information to the operator's view. Also, recent advances in cyber intrusion and electronic warfare have increased the chances of enemy "spoofing," which is the manipulation of friendly systems to provide false information. An example of this is GPS spoofing. It is possible for an enemy force to stimulate U.S. satellites to ensure they output false positioning information of friendly and enemy forces. Finally, due to these issues, the overall operating and situational awareness "picture" the operator has is diminished and the "fog of war" extends to areas that he or she would normally have information on.

Many aircraft, ships, land platforms, and submarines have sensor operators who fill these roles. This effort replicates some of the psychological fidelity of this operator's task and determines if it is optimal to train them to deal with these issues incrementally by providing one at a time in a

high dose (part-task training) or to immerse them in a fully engaged CDO environment where a multitude of issues are happening all at once, albeit in low doses of each uncertainty variable. Although this effort explores a relatively scoped sample task, the implications extend to other tasks and domains, such as civilian intelligence analysts and counter-cyber threat operators.

CHAPTER TWO: BACKGROUND

Setting the Stage - The Next Conflict

As the nature of conflict grows ever more complex, so does the technology that aids the Warfighter. The United States military has enjoyed decades of technological and information dominance over adversaries, with little opposition to the information flow and dissemination in battlespaces. However, future conflicts with near-peer adversaries adopting anti-access/area denial (A2/AD) style strategies will directly challenge the information dominance and technological systems that generate a common operating picture at multiple levels. Rising electronic, cyber, and kinetic warfare capabilities around the world can deny or degrade the information flow to and from U.S. Warfighters, diminishing their capability. This pseudo-asymmetric alternative approach to warfare enables nations that may not have the material or manpower to counter the U.S. in open battle, to gain parity. This enables future adversaries to deny or degrade the U.S.'s capability to collect, analyze, trust, and disseminate information across people, platforms, and commands.

Future conflicts with near-peer adversaries will be dynamic, complex, and ambiguous, but for dissimilar reasons than with recent conflicts. Although the enemy will wear uniforms and practice "conventional" warfare, the U.S. can no longer guarantee information dominance and the technological edge that it has enjoyed for decades. Communication and sensor systems will be targeted first to disrupt information flow. Next-generation Warfighters will have to operate in a Contested, Degraded, and Operationally Limited (CDO) environment. One where systems they were trained to rely on in training may or may not be offline. One where systems could be

compromised or stimulated with false information or "spoofing" designed to decrease situational awareness. One where the optimal scenario entails a sensor operator relying only on partial information. Although the military is developing redundancies to prevent such an occurrence, the likelihood of single platforms, such as an aircraft or ship, relying only on the systems and information they have organically (without input from other platforms), is increasing.

Although the current training envelope ensures Warfighters can troubleshoot faulty systems, less focus is given to training for the contested, degraded, or operationally limited environment. Military think tanks are warning of the need to thwart the "last war syndrome" and the over-reliance on high-bandwidth communication systems employed in the Middle-Eastern conflicts. Further, Rear Adm. Kraft acknowledged that "…maritime doctrine, organization, and training are not optimized to support operations in an environment in which command and control is denied or degraded" (Navy Warfare Development Command Public Affairs, 2012). Although some efforts to train for the CDO environment remain behind closed doors, others, such as the U.S. Naval Academy's decision to re-start the course on Celestial Navigation, in anticipation of the GPS system being destroyed or stimulated with false information, are more public and prominent (Prudente, 2015). It is clear that the technological race between nations will continue; however, the larger determinate is what Warfighters are able to do with that technology. There is a resource constraint tied to how much traditional procedural and checklist-type training can improve performance (Yelon & Ford, 1999). To truly sway future conflicts and exert command over the environment, higher-order cognitive readiness training is required.

Training as a Solution

Developing fail-safes and redundancy plans can mitigate risk at the systems and engineering levels, yet omit the human aspect of higher-order training required in the contested, degraded, and operationally limited environment. The concept of cognitive readiness, the mental preparation warfighters need to establish and sustain competent performance in complex and unpredictable environments (Morrison & Fletcher, 2002), has made an impact in recent years due to its focus on countering dynamic, complex, and ambiguous environments. Although the type of conflict and foe would likely be different, those same characteristics are prevalent in the CDO environment. Cognitive readiness and many of its psychological sub-components, such as transfer of training, problem solving, decision-making, mental flexibility, and creativity can be valuable in preparing and enhancing the knowledge, skills, and abilities of warfighters to counter the unexpected in a complex CDO environment (Morrison & Fletcher, 2002). Many of these sub-components fall under the umbrella of "open skills," or those that enable the trainee the freedom to perform a task in a multitude of ways. In other words, it is not a skill that has a traditionally correct answer, or absolute path to get there. This is in contrast to "closed skills", or those that are more procedural in nature, and have a standard way of performing the task (i.e., a checklist), giving the trainee little room to implement flexible strategies to adapt to variability (Yelon & Ford, 1999). Improving cognitive readiness, and its emphasis on open skills enhancement, is a promising way to enrich the Warfighters cognitive capacities for mental flexibility. Investigating means to improve open skills knowledge and skill transfer is certainly part of the solution; however, more empirical research is needed to identify which facets of transfer provide the most benefit.

Gaps in the Training Solution

One of largest gaps in the transfer of training literature involves training for highly variable open skills, or those tied to principles rather than specific procedures. Training open skills is challenging for several reasons; for example, the trainee has some freedom to perform as there is not a single correct way to complete the task. In contrast, training closed skills, or those tied to specific actions to be reproduced nearly identically in the transfer environment, is more prevalent, and the transfer literature is more solidified with respect to those optimization strategies. However, training closed skills is not always a practical solution. For instance, operators can learn to troubleshoot some common system faults through rote memorization until the knowledge becomes procedural and automatic; yet, expanding that method to encompass the full scope and variety of possible situations is inefficient and would require more time than is available in live, or even simulation-based, exercise. This often repetitive and time-consuming training is coined “low road transfer” training (Salomon & Perkins, 1989) and is very useful for closed skills and procedural training tasks. Although beneficial for conditional training, it is limited in training the higher-order cognitive skills required in CDO environments. “High road transfer,” which involves mindful abstraction, requiring the trainee to decontextualize and re-represent a concept from training, is more suited for open skills required to operate successfully in the CDO environment. High-road transfer is necessary to achieve the levels of flexibility required of the cognitive readiness concept. Yet further research is required to investigate the training methodology requirements to optimize high-road transfer.

Cognitive Readiness

As automation with machines and computers increases and lessens the load of what individuals have to do, it places the human in more of a decision-making and thoughtful role. Operators, especially sensor operators in the military, relegate more and more of the computational aspects to their software. This shift in responsibilities provides them the opportunity to look at the larger picture and determine what all the data points actually mean. As discussed previously, future conflicts with near-peers will be complex, dynamic, and ambiguous. Improving the skillset of an individual to be flexible, adjust, and discern the intent and truth behind the data is the next step in improving readiness overall (Morrison & Fletcher, 2002). Readiness, traditionally measured at the unit level, is the potential of units to perform well in a situation, or their potential effectiveness. Effectiveness is the summative outcome of a task or mission, measured after-the-fact. Conventionally, readiness is measured by aggregating four factors: personnel, training, equipment on hand, and equipment serviceability (Moore, Stockfisch, Goldberg, Holroyd, & Hildebrandt, 1991). What this lacks is an individual evaluation of personnel to assess if he or she is cognitively ready for the CDO environment. As the human remains the fundamental component of military operations, despite technological advances, cognitive readiness is owed its due attention (Thompson & McCreary, 2006).

Cognitive readiness describes the "…mental preparation an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (Morrison & Fletcher, 2002). The authors include knowledge, skills, abilities, motivations, and personal dispositions into this definition under mental preparation. The highlight of this definition is the prominence of performance, in a complex and unpredictable

environment. Cognitive readiness is not focused on every day, procedural type tasks. The concept is designed around (and for) personnel who must perform in quickly changing CDO-type environments. Battle plans drawn up to be executed precisely might be discarded within the first few minutes of contact. The success then falls to each individual to make sound but complex decisions, in little time. Although commander's intent, and the overall goal of the task or mission might stay the same, the environment may require personnel to execute it in a flexible and adaptable way. This is likely to get more difficult in the future, rather than easier as new fronts are opening in the battlespace. With cyber becoming a real threat across the spectrum and electronic warfare gaining prevalence in each branch of friendly (and adversary) militaries, CDO environments may soon be the norm rather than the exception. Even though technological advances are increasing the information flow and theoretically reducing the "fog of war," adversaries are developing comparable solutions to counter that information flow.

Fortunately, there are ways to train personnel in cognitive readiness, although some components are easier than others. The next sections discuss the fundamental way to transfer skills from training into the environment, transfer of training, and what components of it may be especially suited for improving cognitive readiness in preparation for CDO environments.

Transfer of Training

Transfer of training is a concept that has yielded prolific research opportunities over decades, and likely many more to come. It is not necessarily that the field is evolving and requiring new information; rather that the topic is a complex one with various factors that take a lot of experimentation and analysis to tease out and examine independently. Starting with Thorndike

and Woodworth's Principle of Identical Elements (1901), the transfer field launched and sustained careers of many practitioners. Many different variations of definitions appear throughout the literature, although most, seem to comprise similar core concepts. For this purpose, transfer of training is defined as the ability to extend what has been trained in a training environment, and apply it to a transfer environment (Byrnes, 1996; Ford & Weissbein, 1997). This is generally assumed to be positive transfer, although negative transfer certainly exists (Newstrom, 1984). Even broader, it is the application of knowledge, skills, and abilities learned in one situation to another (Perkins and Salomon, 1992). The transfer environment could be very similar to the original training environment (near transfer), or it could be vastly dissimilar, with only remote or conceptual similarities (far transfer) (Macaulay & Cree, 2000). Near transfer is especially suited for technical training where the training should closely mimic the performance desired (Laker, 1990). In contrast, far transfer provides the trainee more freedom to perform the transfer task. Although seemingly dichotomous, transfer could also fall somewhere in between near and far transfer. The Principle of Identical Elements states that transfer is dependent on how similar the training and transfer environment is (Thorndike and Woodworth, 1901). In other words, the more similarities between the two environments, the more able or likely transfer is apt to occur. Currently, this still holds true amongst transfer literature, although caveats certainly exist. It is also important to note that transfer is more about generalizing the knowledge, skills, and abilities learned in the training environment, rather than simply learning something during training (Baldwin and Ford, 1988). The true scope of transfer of training literature is difficult to comprehend, so it is best to discuss it in logical segments devised over the past century. The following sections will present an overview of the popular Baldwin and Ford (1988) model of transfer, discuss the various aspects of it, including input factors and predictor variables, and

then examine the Salomon and Perkins Low- and High-Road Transfer of Training theory, which is most relevant to the current effort.

Model of Transfer

One of the most cited transfer of training papers is Baldwin and Ford's 1988 review, "Transfer of training: A review and directions for future research." Within this is their transfer process model, shown in Figure 1 below.

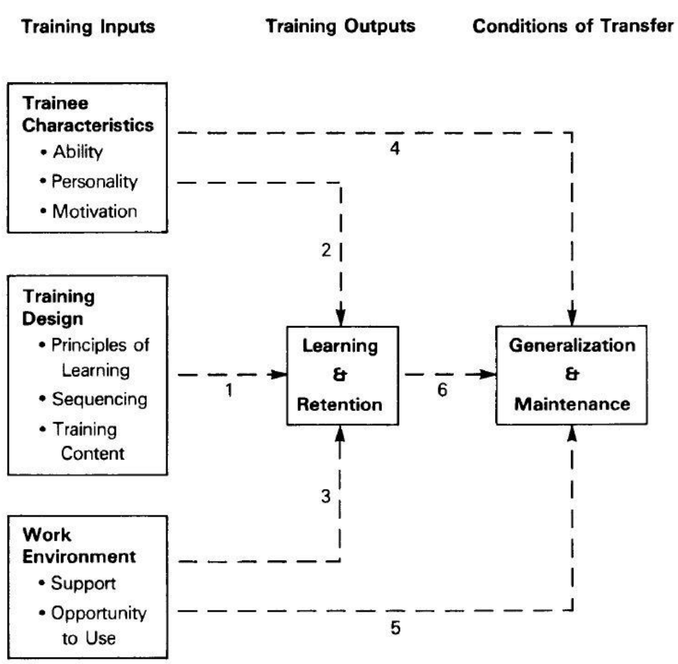


Figure 1. Baldwin and Ford's Model of the Transfer Process (1988)

Referencing the model in a chronological fashion, there are three Training Inputs: trainee characteristics, training design, and work environment. The lines with directional arrows imply that "trainee characteristics" and "work environment" have direct effects on ultimate transfer, regardless of learning and retention. They also have an impact on learning and retention, the

Training Output, but are dissimilar to training design that feeds directly into learning and retention. Kirkpatrick (1967) aptly notes that if a skill is going to transfer, it first must be learned and retained, hence the prior-to-transfer positioning of this portion of the model. Finally, the Conditions of Transfer is the last chronological step of the model. This pertains to generalization and maintenance, the key concepts of transfer. As previously noted, transfer is more than learning and retaining knowledge, skills, or abilities, it requires that one generalize the learned material, and then retain it for use later (Baldwin and Ford, 1988). Even with a seemingly clear model laying out the transfer process, a guarantee of transfer is not assured, especially with respect to cross-domain transfer. Pea and Kurland (1984) found that in general, adults have issues with transfer and it is not necessarily reasonable to expect it. Facilitating transfer is a difficult and complex process that requires attending to the many factors prevalent in the model and all of the sub-components of each of the Training Input factors (Ford & Weissbein, 1997). The following section will describe what each of the Input factors entail, and why the current effort looks to focus on the training design aspect of it.

Training Input Factors

<u>Training Design</u>

Training design focuses on how the training material is developed and presented to the trainee, and any interventions, such as feedback, that occurs during it. The original Baldwin and Ford model included sub-components of training design such as: identical elements, general principles, stimulus variability, and conditions of practice. Over time, there have been additions

and some re-naming of the subcomponents; however, these original components still stand (Burke and Hutchins, 2007).

Identical elements refers to the Thorndike theory referenced previously concerning similarity between training and transfer environments. If a training designer wants to maximize potential transfer, the more identical elements the environments share, the higher the probability for transfer (Thorndike and Woodworth, 1901). This concept has been tested with motor and cognitive skills and the premise remains true (Underwood, 1951; Gagne, Baker, & Foster, 1950).

General principles is the notion that if trainees are taught the underlying rules, principles, and concepts in addition to the step-by-step, if it exists for the task, the potential for transfer is increased (McGhee & Thayer, 1961). Training design involving this would detail the information required for the task, as well as the "why" underlying it all. As discussed later, it is the rules, principles, and general concepts that are abstracted and form the basis for high-road transfer and also allow for cross-domain transfer.

Stimulus variability describes the notion that training around a training objective, in the form of varied scenarios, enhances transfer more than simply training the one objective, over and over. The idea is to use several examples of a concept, allowing the trainee to see multiple sides of it and different applications. Through using relevant, but varied examples, trainees are more likely to transfer the material, and see the benefit of it in a transfer environment (Ellis, 1965).

Lastly, and the focus of this effort, is the “conditions of practice” facet. This entails training design and presentation variables, as well as sequencing. For example, the notion of massed vs distributed practice. Massed practice entails presenting the training session all at one time, whereas distributed training involves separating training over time, providing breaks into the practicing of the trained skill (Naylor & Briggs, 1963). Donovan and Radosevich (1999) noted considerable support for distributed practice for increasing learning, although measures for its impact on transfer were minimal. Distributed training is also essentially required for maintenance and retention of trained knowledge, skills, and abilities. Another example of practice condition is feedback, specifically what, when, and how it is provided. This topic alone can fill bookshelves of journal articles on the optimal feedback interventions per each training task. Additionally, overlearning is another aspect of practice conditions. This is when training continues beyond successful performance. This can take considerable extra resources to continue training to mastery and beyond; however, it has shown benefits to retention and maintenance (Driskell, Willis, and Copper, 1992).

Finally, a significant aspect of training design considers if training should be presented all at once, on a whole-task schedule, or if it should be separated into parts, on a part-task training schedule. Naylor and Briggs (1963) have delved into this topic and provided guidelines; however, in some instances, the empirical research is still wanting. There is a standalone section devoted to this slice of research presented later in this dissertation that examines the benefits and limitations of part-task training and whole-task training.

Trainee Characteristics

As the model shows, trainee characteristics not only impact learning and retention, but can impact conditions of transfer directly. Although research had not reached its full potential at the time of the Baldwin and Ford (1988) analysis, future work revealed some similar insights. Characteristics such as motivation, self-efficacy, perceived utility, locus of control, and cognitive ability have all been found to impact transfer potential in a moderate to strong way (Burke and Hutchins, 2010; Baldwin and Ford, 1988; Blume, Ford, Baldwin, and Huang, 2010). As some of these are mostly relevant to Industrial and Organizational (I/O) environments, the relevant trainee characteristics to this effort will be discussed in the Predictor Variables section below.

Work Environment Characteristics

Work environment plays a significant role in the transfer of training potential of trainees. Although mostly relevant to I/O focused training tasks, it is important to note that fostering a supportive environment comprising peers, supervisors, and the organization at large, has a strong positive impact on training transfer (Clarke, 2002; Facteau, Dobbins, Russell, Ladd, & Kudisch 1995). Foxon (1997) showed that an employee's perception of a supervisor's approval and support had a strong influence on if that worker would transfer that trained skill. If the skill appears important to the supervisor, the trainee also feels a stronger bond with it and is more motivated to learn and transfer that skill. Additional transfer facilitators in I/O type environments include opportunities to perform said trained skill (Clarke, 2002; Lim & Johnson, 2002), goal setting, encouragement, modeling of behaviors, and reinforcement (Huczynski & Lewis, 1980; Maddox, 1987). Again, fostering a positive work climate is essential for organizational settings,

although less so for the particular use case this effort is focusing on, as it relates to mostly individual training on a simulator. However, volumes of work have been produced on the importance of transfer climate in business settings and other I/O related environments (Burke and Hutchins, 2007).

Predictor Variables and Moderators of Transfer

One of the most recent meta-analyses of transfer examined empirically supported predictor variables and moderators (Blume, Ford, Baldwin, & Huang, 2010). For context, it is important to delineate how transfer is assessed in some instances. Some studies assess it as the use of the trained skill, while others assess it as the effectiveness of performing the trained skill, or based on the outcome. It should be required to first operationalize the assessment of transfer prior to each investigation. For example, the present effort will be assessing transfer as the effectiveness of performing the trained skill, although a post-study questionnaire will also try to elicit use as well. It is expected that there will be stronger predictor-transfer relationships when transfer is measured as use, versus effectiveness, due to the fact that use is required for effectiveness. Simply, one can use the skill, and score high in that regard, but still fail to employ the skill appropriately and not perform on the effectiveness side of assessment (Blume, Ford, Baldwin, & Huang, 2010).

The following section identifies empirically associated predictor-transfer variables. There are more identified in the literature; however, in an effort to scope the review to the relevance of this effort, some are excluded. For example, transfer climate (or work environment) has a high relationship with transfer ($\rho = .27$) but is less important in the current use case as it is not an I/O

type study. Additionally, motivation also had a high positive relationship ($\rho = .29$); however, motivation is not a requirement for many critical skills in all domains (such as the military), and likely plays less of a role. In other words, motivation is encouraged but not necessary.

<u>Cognitive Ability (Predictor)</u>

The Blume, Ford, Baldwin, & Huang (2010) meta-analysis found similar results to other research showing that cognitive ability has strong relationships to transfer (overall ($\rho = .37$)). However, as expounded upon in the later section on open and closed skills, there is a negative relationship concerning open skills, and a positive relationship concerning closed skills. This was the largest predictor in their analysis. Other trainee characteristics had either smaller relationships, or were not relevant to the current use case.

<u>Time Measurement (Moderator)</u>

Time measurement relates to when the measurement of transfer was taken. In the majority of lab settings, this is done right at the completion of the study. Blume, Ford, Baldwin, & Huang (2010) found, as expected, that transfer relationships are stronger when it is measured right after training and task completion. If there is lag, as there is in many field studies, the relationship between the predictor variables and transfer weakens. Taylor (2009) attributes this to training decay, and/or the fact that the trainee might not have had enough time to actually use the skill in a performance environment yet. Although the relationship generally weakens, it was found that some constructs (environment, pre-training self-efficacy, and motivation) were not affected by time measurement.

Low- and High-Road Theory on Transfer

Alluded to before, the theory put forth by Salomon and Perkins (1989) suggests that there are two methods to transfer training skills: low-road and high-road transfer. Both have benefits, limitations, and optimal use cases.

Low-Road Transfer

Low-road transfer involves repetitive, but varied practice, to automaticity (Salomon and Perkins, 1989). This repetitive, and often time-consuming training, is primarily used for closed skills and procedural training tasks. Typically, trainees learn a skill, and then practice it until it becomes "second nature." Optimally, there is some variance in the repetitive training to slowly expand the trainee's awareness of use cases for the skill, adding some flexibility. A key phrase used by Salomon and Perkins is that low-road transfer is "incremental" in that it slowly builds up broader applicability. Although beneficial for conditional training, it is limited in training the higher-order cognitive skills required in CDO environments. Further, if varied practice is all but removed, there is evidence that although the trainee will become quite proficient, even a master, at the task, a transfer paradox occurs, where transfer of the learning outcomes is bleak as the trainee is only seeing a very narrow view of the training task (van Merrienboer, Kester, and Paas, 2006).

High-Road Transfer

High-road transfer, in contrast, focuses on "mindful abstraction," requiring the trainee to decontextualize and re-represent a concept from training (Salomon and Perkins, 1989). It is a

deliberate process, and the trainee is largely in charge of what elements are decontextualized. The abstraction of core elements, acts as the bridge between the training environment and the transfer environment, whether it is similar to the original task, or completely different domains. This type of transfer is more suited for open skills required to operate successfully in the CDO environment. High-road transfer is necessary to achieve the levels of flexibility required of the cognitive readiness concept. Yet further research is required to investigate the training methodology requirements to optimize high-road transfer.

This effort focuses solely on optimizing high-road transfer in a sample use case. Low-road transfer is simply too resource consuming to yield the desired results. Gains can be made through varied practice in low-road training, but optimally, high-road gains would result in the largest "jumps" of skill. The next section examines mindful abstraction in more detail.

Mindful Abstraction

The essential concept of high-road transfer is that of mindful abstraction, the conduit from one context to another. This is a conscious, effortful, deep, awareness of potential connections that allows a person to bridge concepts or environments (Salomon and Perkins, 1987; Langer, 1989). Generally, this involves the deliberate search for patterns, rules, or principles, and a decontextualizing of core elements (Salomon and Perkins, 1989). The key idea is that the trainee is aware that he or she is trying to abstract ideas from a particular training session or task. The trainee is not a passive participant in the training, rather in a heightened alertness state and taking an active role in forming new connections and perhaps seeking them out. Even when cues or hints are not given to trainees explicitly, the connections found act as possible hints and prompts

(Gick and Holyoak, 1980). The researchers examined analogous problem solving, and although the study was limited, they suggested that comparable results would be found in future studies. They caution that the same might not be found with respect to procedural type situations with discrete steps.

Although the trainee might not need to be explicitly told about the connection or abstraction, Salomon and Perkins (1987) point out that the abstraction, itself, has to be understood, hence the mindful requirement of it. If the trainee is not aware that he or she has abstracted something, it is likely not to benefit that person much in different environments. In other words, it would be a superficial abstraction, rather than one that is genuinely comprehended. Yelon, Sheppard, Sleight, & Ford (2004) note that intention to transfer is a mindful process and conscious decisions are necessary for effective transfer.

In relation to the sample use case of a sensor operator preparing for entering the Naval Force, he or she has a specified amount of training time. Especially in a resource-constrained environment (fiscally), there may be limited opportunities to train in high-fidelity simulators. The time afforded is then restricted to training the basic and intermediate tasks for which the operator will be primarily responsible. These types of tasks are suitable for low-road transfer, where procedural and reflexive responses are desirable. However, when focusing on training for a dynamic, complex, and ambiguous situation, such as the CDO environment, operators require more than their checklist-type training. To train the wide variety of situations they might encounter through low-road transfer, it not only would require a significant amount of time, but it would limit them to the breadth of varied scenarios to which they have been exposed. In

contrast, high-road transfer could enable greater “jumps” or “leaps” in expanding their training envelope. Ideally, it would require less time, and greatly increase the ability of the trainee to respond to a wide variety of situations through abstracting concepts and ideas through training, rather than individual scenarios. Those concepts, principles and ideas can then be leveraged to new environments and contexts. As such, determining ways to optimize mindful abstraction facilitation during training is a potentially fruitful endeavor.

Open vs Closed Skills

Another way to look at low- and high-road transfer is by the type of skills typically trained. Low-road transfer comprises “closed skill” training, and high-road transfer is mostly concerned with “open skill” training. Although closed skill transfer is more straightforward in comparison, if able to focus on specific cases and “clear away the smoke” that makes transfer complex, one can systematically examine factors that impact transfer, even for open skills (Yelon and Ford, 1999). Blume, Ford, Baldwin, and Huang (2010) point out that the open or closed nature of the task being trained is typically neglected with respect to transfer, especially so if the task is dealing with open skills. These two types of skills are different enough, with one being rigid and the other being variable, that the factors that impact transfer of each are likely different (Baldwin, Ford, and Blume, 2009).

Closed Skills

Closed skills are those where performance and circumstances for use are standard, leading to one way of correctly acting to bring about the desired result (Yelon and Ford, 1999). These skills are procedural in nature and the training of certain skills is meant to be replicated in the transfer

environment according to a set of predetermined rules. Often, the trainee is trained on the task, and then given ample time, in a related environment, to then practice that task until proficiency. In practice, the majority of information regarding these studies come from lab-based research.

In a recent meta-analysis of moderators of transfer, researchers examined which had an impact, and how different it was with respect to closed or open skills (Blume, Ford, Baldwin, Huang, 2010). One of the more interesting moderators, cognitive ability, had a strong positive relationship with transfer for closed skills (ρ = .41). It is interesting because for open skills, the reverse is true, as cognitive ability showed a negative relationship with transfer (ρ = -.14). This provides evidence that cognitive ability plays a more important role in closed skill transfer, but less of a role in open skill transfer. In virtually all other moderators examined (e.g., motivation, trainee experience, work environment, etc.), open skills always had stronger positive relationships with transfer than did closed skills. Although the current effort is focused on transfer of an open skill, this one oddity warrants an extra look in the current effort to ensure cognitive ability is controlled for.

Open Skills

Open skills are those where performance and circumstances for use are flexible, and provide the trainee the leeway to decide when and where to use them. The training objectives are tied to principles, rather than specific skills giving the trainee the freedom to perform, and not be forced into doing it one single way (Yelon and Ford, 1999). In other words, the trainee has a say in, or has more choice in the way he or she accomplishes the training objective. He or she can also look for opportunities to use different skill sets to reach the training objective, assuming the job

environment allows for such (Ford, Quinones, Sego, & Sorra, 1992). As much of this research is Industrial / Organizational Psychology related, and open skills are generally field-based studies, one can see how these play an important role.

In the recent Blume, Ford, Baldwin, and Huang (2010) meta-analysis, there was a significant disparity between the relationship of cognitive ability and open and closed skills, as noted in the previous section. Although transfer had a negative relationship with open skill (ρ = -.14), this was based on only two studies and certainly warrants future investigation considering closed skills has a strong positive relationship with cognitive ability. The majority of the other variables examined in the meta-analysis showed higher positive correlations between open skills (e.g., pre-training self-efficacy, work environment, post-training self-efficacy, etc.). Although variables like motivation, environmental factors, and self-efficacy impact open skills more than closed skills, an effort to control these variables will likely mitigate their influence concerning the current effort.

The discussion of open and closed skills is relevant to this effort because the goal of facilitating mindful abstraction through part-task training or whole-task training falls into the open skill category. There is not a checklist that can be used for this task, in this type of complex, ambiguous, and dynamic environment. Imagine providing ample flexibility to someone completing a checklist that should be completed in order and one can see how that may not yield optimal results. Open skills are suited for rapidly changing situations, as they are inherently flexible, and enable the trainee to have more choice in the path forward. For example, in normal, everyday operations, a sensor operator would refer to Standard Operating Procedures (SOPs) that

have likely been ingrained in his or her brain. By training standard procedures, mostly closed skills, rigidly and often, the hope is that those fall under automaticity, and require little active thinking to perform when then stimulus arises. By training these to proficiency, the operator, hypothetically, has more resources to deal with changing environments or tasks. Salas, Milham, and Bowers (2003) note that with increased technology, changing task demands in military jobs, and volatile operating environments, open skills will be more prevalent and necessary going forward.

Conditions of Practice

The popular transfer of training model by Baldwin and Ford (1988), as previously discussed, has training inputs, training outcomes, and conditions of transfer. Under the umbrella of training inputs, they list trainee characteristics, work environment, and training design. The latter of which is the primary concern of this dissertation.

This effort's fundamental contribution is to provide insight to the benefits of employing part-task or whole-task training to facilitate mindful abstraction, the key behind successful high-road training transfer (Salomon and Perkins, 1989). The goal, being able to pronounce that, all things equal (or accounted for), part-task training involving CDO variables and characteristics facilitated mindful abstraction to the transfer task, or vice-versa. Part-task and whole-task training accompany other "conditions of practice" variables, such as: massed vs distributed practice, feedback prompting, and overlearning. Each of these are studied independently, as well as in the context of transfer of training. For example, if the question becomes how often to train a skill to optimize transfer, one might look to the literature on massed vs distributed training

(Briggs and Naylor, 1962). Addressing each of these conditions is beyond the scope of this effort, as the focus is providing insight into one small, but critical area: does part-task and/or whole-task training facilitate mindful abstraction necessary for high-road transfer?

Part-task and whole-task research, in general, is not a new field and has been the interest of many throughout the past century. One persisting theme that frequents the literature is that training design, is critical to training transfer (Teague, Gittelman, & Park, 1994). However, within training design, there are multiple facets, and training design itself, is only one part of the model, as trainee characteristics and the work environment are equally as important input factors. Although trainee characteristics cannot be fully controlled for, previous research examining predictor variables affecting transfer have exposed which characteristics are more important than others. For instance, cognitive ability (or IQ) had the highest positive relationship with transfer of training ($\rho = .37$) in a recent meta-analysis conducted by Blume and colleagues (Blume, Ford, Baldwin, & Huang, 2010). This, along with voluntary participation ($\rho = .34$) and conscientiousness ($\rho = .28$) had moderate relationships with training transfer, at the trainee characteristics level. Similarly, work environment factors also play a role in affecting positive transfer; however, many of these instances were examined through an I/O lens, rather than a pure training lens. For example, many of the factors affecting transfer in the work environment, such as supervisor support and having a positive transfer climate, are less relevant in a structured training task for someone in the military. If able to control for trainee characteristics and work environment, within a lab setting, it is possible to focus less on these inputs of the model, and more on the training design aspect.

Early work featuring part-task and whole-task research favored the whole-task training method, although the results were small to moderate, and later researchers attributed much of the early results to flawed methodological issues (Teague, Gittelman, & Park, 1994). Later work found that part-task training was beneficial, but relatively contingent on the trainee understanding the whole task, or else the "benefits of part-task training are short lived" (Newell, Carlton, Fisher, and Rutter, 1989). Naylor and Briggs (1963) examined the effect of task complexity and task organization on part-task and whole-task training schedules. Task complexity was defined as "…demands placed on a trainee's information processing and memory storage capacities." Task organization was defined as "…demands imposed on the trainee due to the nature of the interrelationship existing among task dimensions." They found that for high organization, high complexity tasks, whole-task training resulted in the highest performance. This sentiment is still echoed in recent literature as complex and highly organized are best trained via whole-task methods (Van Merriënboer, Kirschner, & Kester, 2003). However, the more unorganized a task becomes, the more part-task training becomes beneficial, especially as the complexity (difficulty) rises. This is chiefly because part-task training methods are able to prevent or reduce that chances of cognitive overload due to parts of a task requiring less load than what is associated with the whole task (Van Merriënboer, Kirschner, & Kester, 2003). In relation to the current effort, an assumption is that normal military operations for sensor operators can be classified as high organization, low-to-medium complexity tasks. The operators are adhering to their standard operating procedures (high organization) and they have been well trained in their roles, resulting in relatively low difficulty (low-to-medium complexity). Yet, as CDO variables start to encroach on everyday procedures, the task becomes more disorganized, and the task complexity rises. When this happens, and trainees start being unable to clearly understand the

full picture, and how it all fits together, the part-task training schedule should be a more viable alternative. The Naylor-Briggs theory is organized as a dichotomy with high and low task organization and complexity, but in practice, categorizing tasks proves to be challenging and many fall on a continuum between their anchor points. While their research provides guidelines, it also contributes to conflicting research regarding this topic. Amidst the early disparity and unclear results, additional researchers pressed forward to provide insight into the part-task vs whole-task training schedule question. The overall pros and cons of each are presented below.

Semantically, this dissertation uses part-task and whole-task to describe the presentation of the variables to participants, as does much of the conditions of practice literature. Although the same presentation of variables, depending on the task, can also be described as simple presentation, or complex presentation. There is significant overlap in the literature and wording of these two types of presentation.

Part-Task

A part-task training schedule involves dividing a training task into sub-tasks, or smaller components when presenting the training to the trainee (Naylor and Briggs, 1963; Teague, Gittelman, & Park, 1994). For complex, dangerous, and/or difficult tasks, it is likely easier to practice small segments of the task, in isolation first, and then move towards the next part. For instance, when learning how to disassemble a car engine, the trainee would not be shown how to do it from start to finish and expected to practice that over and over. Instead, the training would likely involve training the individual how to complete smaller, naturally occurring sub-tasks and learning to disassemble each component of the engine, one at a time. When natural sub-units of

training exist or “small wholes,” part-task training is more beneficial than whole-task training (Holding, 1965). This is also the case when training the whole task, might be unreasonable, or expensive even. Additionally, part-task training excels when focusing only on re-training (or refresher training) one critical element of a task (Knerr et al., 1985).

Part-task training, itself, has been broken up into a few “types” as there is more than one way to divide up a task. Wightman and Sistrunk (1987) examined three different ways to develop a part-task training schedule: segmentation, fractionation, and simplification. Segmentation involves scheduling training in the reverse order by presenting the final sub-task, first, and then presenting each preceding sub-task. They found that this schedule excels particularly when training perceptual motor tasks, when tasks have an organized fashion. Fractionation is a way to divide a training task into sub-tasks that are typically done simultaneously. The goal is to train each, in isolation, and then combine them so that the trainee is doing them all at once. Lastly, simplification involves actually changing the difficulty of the task. Early training would be significantly simplified with the whole task being decomposed into the very basic elements. Once the basic elements are trained, additional sub-tasks are added back on top of the training, each time raising the difficulty some. Once all the parts are reintegrated, the trainee would optimally be performing the whole task. It is up to the training designer to discern which is the optimal technique of breaking the task down when scheduling part-task training.

Although part-task training is beneficial and optimal in many instances, it also has limitations in that it takes extra time and resources to divide a training task into smaller sub-components. Further, focus must also be given to re-integrating the disparate training tasks at the end of

training to ensure that trainees have a grasp of the complete and whole task. Additionally, most part-task training focus has been on closed skill tasks (procedural training), rather than open skills (e.g., decision making, mindful abstraction, leadership). Research into cognitive skills have found that part-task training is less suitable for training these specific skills (Spector and Anderson, 2000; van Merriënboer, Kirschner, & Kester, 2003). Lim, Reiser, and Olina (2009) studied the effects of part-task and whole-task design approach on the acquisition and transfer of a complex cognitive skill and found that the whole-task group performed significantly better on the transfer task, as well as the acquisition task. Van Merriënboer's work also raised concerns that part-task training focuses too much on reaching separate, disparate objectives, and less on integrating them at the end to form a whole task.

With respect to the current effort, the general finding is that part-task training is optimal when trainee cognitive ability is low, when training is massed (vs distributed over time), and when the task is low in task organization, but high in complexity (Naylor and Briggs, 1963). However, this effort hopes to provide clarity to the issue that Lim, Resier, and Olina (2009) raise regarding the lack of empirical evidence concerning practice schedules for the transfer and acquisition of cognitive skills. Some initial evidence points to whole-task training being more suited for cognitive skills, although research, to date, is not conclusive.

Whole-Task

Whole-task training consists of presenting a complete or whole training task at one time. Using the car engine example from before, rather than learning how to take apart different sub-components of the engine, the trainee would be told how to disassemble the engine, from start to

finish, including all the sub-component disassembly. In general, this is optimal for simple tasks, where the trainee can logically imagine the whole task as he or she is completing it (Teague, Gittelman, & Park, 1994). The more complex, difficult, and disorganized the task becomes, the less able trainees are able to approximate the whole task at a high level.

Previous whole-task training work has also been largely examined through the closed skill lens. Many of the tasks in studies consist of tangible knowledge, skills, and abilities that are procedural in nature, rather than attempting to measure open skills, that are cognitively more complex. Lim, Reiser, and Olina (2009) note that "…little empirical evidence exists with regard to the effects of the application of whole-task approaches on the acquisition and transfer of complex cognitive skills." It may be less represented in the literature because it is harder to measure; however, some have pointed out that whole-task training is better for teaching cognitive skills, so it does warrant extra focus (Spector and Anderson, 2000; van Merrienboer, Kirschner, & Kester, 2003). As suggested previously, Lim and colleagues (2009) found the whole-task instructional approach to be superior in the transfer and acquisition of a complex cognitive skill. Yet, they remain unsure as to why and were not able to discern if the learners developed (or used) an underlying schema and applied that to the transfer task. The whole-task group could have also performed better simply because they engaged in whole-task varied training, and research is clear that varied training is key to improving transfer (Cormier and Hagman, 1987; Singley and Anderson, 1989). In other words, their participants completed a whole-task for multiple different scenarios, whereas their part-task participants completed only one whole scenario, but in parts. The whole-task group was exposed to more variability, hence a greater prospect of transfer.

With respect to the current effort, the general finding is that whole-task training is optimal when trainee cognitive ability is high, when training is distributed (vs massed at once), and when the task is high in task organization, but low (or varying) in complexity (Briggs and Naylor, 1962; Naylor and Briggs, 1963). Naylor and Briggs have shown that task organization (interrelationship existing among task dimensions) is critical, with less emphasis placed on difficulty. Whole-task training seems most appropriate when the trainee understands the high-level picture, can understand it cognitively, is easily able to make the connections in their brain regarding task relationships, and is able to practice over many sessions.

Regarding the Contested, Degraded, and Operationally Limited (CDO) environment, its definition is almost the antithesis of organized; typically, it is described as complex, dynamic, and ambiguous. A task focused on training under these conditions likely falls in the low task organization, medium or high difficulty realm on the Naylor/Briggs continuum introduced above. In a perfect operational setting, mission tasks might be somewhat related and organized; however, in the CDO environment that all changes, and increases the demand on trainees. Although there is evidence that points to whole-task training being more suited for cognitive skills, there is also a possibility that due to the low task organization of the training, part-task scheduled training may prove to be better. Research, to date, is inconclusive regarding this set of circumstances, and is one of the catalysts for this research effort.

The Role of Workload

To build on the previous section relating to the work and ideas of Naylor and Briggs, the next logical and relatable topic entails the "why." Why is it that one type of training schedule (e.g., part-task) would do better in a complex and disorganized task, than say a whole-task schedule? The answers to these types of questions are found in the cognitive workload literature.

As previously discussed, mindful abstraction takes effort and focus, as it does not happen without awareness. While participants are engaging in a task that is trying to foster it, it is vital that the participants have enough left-over cognitive resources to attend to abstracting. Otherwise, participants will likely be too overburdened by the task and their finite amount of cognitive resources will be used in an effort to complete the task, rather than abstracting concepts from it to apply to later instances.

Anytime someone is doing something, such as completing a task, or participating in training, there is a cost to performing that task. Some of it is a physical cost, such as moving or lifting items. Other tasks require primarily mental capabilities. Some, require both to various levels. This cost imposed on the trainee, whether physical or cognitive, needs to be measured and is done so chiefly, by measuring workload is to determine the mental cost of performing a task (Cain, 2007). Defining workload precisely is a subject for debate amongst researchers, although core ideals are usually agreed upon. It typically is the result of taskload and performance on a task. It is the amount of "work" someone has to do while engaging in a task, whether physical or cognitive. Practitioners measure workload in order to identify any areas of overload that could be problematic to the user completing the task. They measure this by subjective or objective

measures (or sometimes a mixture of both). Subjective measurement involves perceived workload and is measured post-task (or mid-task). Objective measurement usually uses physiological measures to try and assess how much strain is put on the human, physiologically. Both types endeavor to approximate the overall mental effort required of task.

Two early theories explained workload in both contrasting and complementary ways. One of the first, unitary resource theory, as the name suggests, understands mental resources as a singular pool (Kahneman, 1973). When completing a task, the difficulty of it, or the taskload, determines how quickly those resources are expended. Processing after the resource limit proves to be difficult and imposes costs on the user or operator. This theory retains all resources, be it verbal, or visual, audible, in the same pool. Another idea sought to tease out and expand on unitary resource theory, as some claimed mental processing was more complex than originally stated. Multiple Resource Theory (MRT), put forth by Christopher Wickens (1984), postulated that there are multiple pools of resources at a human's disposal with varying capacities that can be utilized simultaneously, as long as the task is not drawing from the same resource pool. These pools chiefly comprise the different senses, such as visual, auditory, tactile, and olfactory. The takeaway from MRT, is that one can sufficiently perform a task that requires visual and auditory resources, but not one that requires two visual or two auditory needs. When too much is required of one resource pool, then the trainee will become overloaded and processing and task performance will likely degrade. This is because the tasks that are competing at the same level are inhibiting each other, which would not be the case if the tasks were using separate pools or dimension levels (Wickens, 2002).

In terms of this dissertation, the significant points from the above resource theories is that both discuss cognitive workload with the recognition that the human system possesses a finite amount of resources. Additionally, both theories note that once the resources are depleted as cognitive load increases, performance can suffer (Friedenberg & Silverman, 2006).

Another theoretical interest for this effort is the role of working memory and workload. For this aspect, the work of Baddeley and Hitch (1974) and Sweller (1988) will be examined. The Working Memory Model of Baddeley and Hitch (1974) describes how different types of information is processed through short-term memory, and if warranted, into long-term memory. Chiefly, working memory is comprised of a few different subsystems, as shown in dual-task studies (Hitch and Baddeley, 1976). One component is the phonological loop, responsible for spoken and written material. Another component, and the one that is more important for this dissertation effort, is the visuo-spatial scratch pad. This stores and processes information in visual or spatial form. The last component is the central executive, which makes decisions on which issues deserve attention, and which should be ignored (Baddeley, 1986). Baddeley notes that the central executive is less of a storage system, like the phonological loop or visuo-spatial sketchpad, and more of a system that controls attentional processes. These three subsystems form the working memory model that describes how humans are able to remember and process information at the same time, for short periods. This information will be forgotten if there is not a conscious effort to retain it into long-term memory. Once it is consciously retained, it can then be fit into an existing schema, or placed in a new one for more permanent storage.

Determining the conditions for if a piece of information is stored for later during training and/or instruction is the subject of Cognitive Load Theory (CLT) (Sweller, van Merriënboer, & Paas, 1998; Paas, Renkl & Sweller, 2003). Cognitive Load theory was proposed as a way to describe the information pathway from working memory, to long-term storage (Baddeley & Hitch, 1974; Baddeley, 1986). As people come across endless streams of new information daily, CLT envisions how individuals sort through the working memory information and process it for more permanent memory. The following discussion of CLT will expand on the three types of cognitive load, intrinsic, extraneous, and germane, as well as touch on the assumptions that underlie CLT, especially focusing on schemas.

In psychology, schemas are described as a mental framework for organizing and discerning where to place new information, and at what complexity level. They allow someone to receive a new piece of information and compare it with their current knowledge structure. For example, seeing a new breed of dog in the world. If the person has encountered other dog breeds in the past, he or she likely has a schema foundation to build on. An animal, with four legs, that is furry, has a tail, sixty pounds, and generally found around humans being friendly. Now the new animal might not fit directly into their existing schema, but rather than generating a new "place" in his or her brain from the ground up, CLT posits that the person will generally align the new breed of dog into their existing schema and simply expand their definition of "dog." In other words, it helps humans to organize incoming information with preconceived ideas and concepts and lets humans combine a few pieces of information into a single component for storage (Paas, Renkl, & Sweller, 2004). The schemas exist as an efficiency in the brain in processing new information, as well as sorting and structuring old information and are able to do this quickly. A

key element of schema efficiency is that when new information comes in (learning or training), an individual can take a cluster of lower-level schemas about a topic, and combine them into a higher level schema. This results in a complexity structure that enables creation of increasingly complex schemas.

Once information is processed from working memory into long-term memory and is placed or sorted into the appropriate schema, there is virtually no limit to the amount of information that can be stored. New information in the working memory is simply incorporated (if possible) into existing schemas. If someone is an expert or very knowledgeable in a domain and has a well-constructed schema, less load is placed on working memory to incorporate the new information. The less load imposed on a trainee or learner during information intake, the more load available to process novel information into their schemas.

The seemingly limitless long-term memory lies in contrast with the working memory, that has limitations when handling novel information (Baddeley and Hitch, 1974). Additionally, as mentioned in the previous section, the working memory handles visual information and auditory information separately. Keeping the assumptions of CLT in mind when designing new training is core to limiting unnecessary burden, as too much cognitive load will lessen the chance that new information will be retained and build into existing schemas. The types of cognitive load that one can intentionally or unintentionally overload are the subject of the next section.

Cognitive load, as defined by CLT, is broken down in to three sub-components. There is intrinsic load, extraneous load, and germane load. Each are summative to the burden placed on an

individual during learning or training. There are some sources of cognitive load that one can account for, but not adjust, such as intrinsic load. A subject that is complex inherently would be an example or high intrinsic load. Essentially, a topic or subject that is "hard" to begin with. This type of load is difficult to adjust through traditional methods of instructional design and is more of a placeholder load that simply exists and must be reasoned with (Ayres, 2006). The obvious exception to simply accepting high intrinsic load, is for performers to become more acquainted with the material. Someone proficient in a traditionally "hard" topic, will encounter less intrinsic load naturally as he or she has the expert advantage and needs to process far less new information, and has a considerable schema already constructed. This is further discussed in the section on the differences between novice and experts in a subject matter field. Not having to process loads of new information frees up cognitive resources that novices would not have available. Due to this, further schema creation is possible, whether it be expansion or modification of their existing structure (Paas, Renkl, and Sweller, 2004).

In contrast to intrinsic load, extraneous load is a burden on the working memory that is manageable and is typically unrelated to the learning task at hand, or in other words, a distraction from the learning task. As the three types of load are additive, anytime there is extraneous load, there are less resources for the other forms, such as the last type, germane load. Germane load is the mental processing effort that supports the development of processing relevant incoming information to long-term memory, or more simply, the development and modification of schemas. Germane load is the effort it takes to create the actual connections between existing knowledge and novel information. This type of load should take priority in learning and training situations, and instructional designs can stimulate the actions of schema development. So in

summary, intrinsic is largely unchangeable but should be minimized, extraneous should also be limited, and germane load should be maximized. By doing this, the learner or trainee is able to create new connections between new and existing ideas or concepts and store them for long-term memory. If one does not foster germane load and instead allows for more intrinsic or extraneous load, the working memory capacity may be exceeded, resulting in a lack of new processing power (Sweller, van Merriënboer, & Paas, 1998).

With respect to this research effort, the CDO environment characteristics, being complex and disorganized, will likely stress the cognitive abilities of trainees. Using the research outlined above, and specifically cognitive load theory, this dissertation looks to reduce the cognitive load imposed on the trainees through using the optimal training schedule (part-task or whole-task). By presenting the training in a part-task fashion, the participants should experience minimal cognitive overload (as measured by the TLX scores), that should enable trainees to use their free cognitive abilities in order to process and abstract concepts for use later. Whereas participants in the whole-task condition are predicted to have a higher cognitive load imposed on them, and will not have the germane load required to mindfully abstract core concepts to transfer later.

Research Hypotheses

Hypothesis 1: There will be a significant difference in performance between the control condition (no intervention), and both experimental conditions (part-task training and whole-task training).

Hypothesis 2: There will be a significant difference in the performance (time, accuracy, uncertainty variables identified, and workload) between the part-task training and whole-task training conditions.

Hypothesis 3: Participants in the part-task training condition will outperform (time, accuracy, uncertainty variables identified, and workload) participants in the whole-task training condition and the control condition.

CHAPTER THREE: METHODS

Methodology

Participants

Participants for this experiment included undergraduate and graduate students from the University of Central Florida between the ages of 18 and 42 (M = 21.9, SD = 4.1). Eighty-one participants (55 females, 26 males) were recruited for the study using online participant systems (SONA) curated by the University of Central Florida Psychology Department and the Institute for Simulation and Training (IST). A power analysis was conducted with G-Power software to determine a sufficient sample size (n = 81) using an alpha of 0.05, a power of 0.80, and a small-to-medium effect size (d = 0.35) (Faul, Erdfelder, Lang, & Buchner, 2007; Cohen, 1988). The 81 participants were randomly assigned to one of three training conditions: control, experimental part-task training, experimental whole-task training. Each group comprised approximately twenty-seven participants. Outliers were screened and accounted for at the conclusion of the data collection and prior to primary analysis. Participants were compensated for their time through monetary means ($10 per hour).

All participants were screened to have normal or corrected-to-normal vision, as well as being tested for colorblindness due to the discriminatory nature of the task.

Materials

Stimuli

Four custom puzzles were created for use in this experiment. The puzzle pieces were square shaped, and not jigsaw puzzles. Each was designed using copyright free high-resolution images from photography repositories. The three training task puzzles were designed to be a standard "moderate" difficulty puzzle for the age group. Puzzle difficulty is determined by complexity of the image, number of pieces, number of discernable sub-zones that can be solved mostly independent of the rest of the puzzle, and shape of the puzzle pieces. Each puzzle included 36 pieces before condition intervention. Additional stimuli were gathered in the same fashion (e.g., in the additional information task that includes extra puzzle pieces, closely related photographs were chosen to the original so that the pieces may look like they fit, however, under scrutiny, it will be possible to discriminate the additional pieces from the original puzzle piece set). Each of the training condition puzzle pictures were made from natural environment photography. The images themselves varied; however, each was essentially split into thirds, with the upper third being noticeable discernable from the middle third, which is also discernable from the lower third. Imagery of the final puzzle solutions are shown below:

Figure 2. An ocean themed training task puzzle image

Figure 3. A lake and mountain themed training task puzzle image

Figure 4. A forest themed training task puzzle image

The transfer task puzzle was more complex with less discernable regions throughout. For example, an urban scene with mostly buildings containing lines and right angles including very similar coloring patterns was used. This is more difficult to solve due to image differentiation issues. Participants should find it harder to solve because it has fewer discernable and distinct zones. Strategies that involve solving “mini puzzles” within the main puzzle are anticipated to prove difficult and time consuming. This left the trainee with fewer automatic reference points (i.e., this piece is blue, it is part of the upper third that makes up the sky). The transfer puzzle solution image is included below:

Figure 5. A city themed transfer task puzzle image

Demographic Form

A demographics survey was administered to collect information regarding the participant's age, gender, vision status (normal or corrected vision), colorblind status, approximate GPA, and past puzzle experience and confidence in solving puzzles. This information was used for descriptive purposes, to aid in screening out possible outliers, and ensure overall equivalence across conditions. One reason to collect prior-experience data is because we know there to be significant differences in novices and experts. Increasing experience and knowledge in a specific field such as physics, chess, or electricity, has the effect that objects or properties that originally had to be abstracted or contextualized, no longer have to be; these properties can be perceived by experts and problems can be solved more accurately than by novices. Essentially, experts solve problems with different strategies. For example, in a research study that evaluated the ability of novices and experts on solving physics problems, it was reported that experts work from the quantities given to the unknown while novices work from the unknown to the given values (Larkin, McDermott, Simon, & Simon, 1980). Experts can also quickly recognize the category of

problem, while the novice will try to get factual and procedural information that relates to the problem and then try to formulate an answer using that knowledge (Green, McCloskey, & Caramazza, 1982). Experts and novices also differ in not only the amount of information that they have access to, but the kind of information as well (e.g. morphological information that can be observed is equal to both expert and novices, however, only experts will have information through cultural transmission or experience) (Boster, and Johnson, 1989). Another issue between experts and novices is that novices tend to form half-truths and misconceptions about a subject before they are formally trained in the subject, while experts learn over time that these half-truths and misconceptions are false and can explain natural phenomena using fundamental principles (Green, McCloskey, and Caramazza, 1982). For these reasons, demographic questions related to puzzle solving experience were collected to attempt to have experts self-identify prior to the task. The demographics survey is found in Appendix A.

Ishihara Colorblind Test

A simple six-panel Ishihara colorblind test was administered to ensure equality across all participants. Although the demographics survey questioned participants about color vision abnormalities, the Ishihara test acted as an additional check for some that may not know they are colorblind. As the task involves color association and discrimination, an objective measure of color vision was appropriate. A copy of this test is found in Appendix B.

Wonderlic Personnel Test - Basic IQ Test

All participants took the Wonderlic Personnel Test (WPT-Q), which is a 30 question, quick, measure of intelligence, verbal, numerical, and spatial abilities (www.wonderlic.com). It is a

widely used psychological assessment that can be administered in eight minutes. Questions progress from easy to difficult as the tester moves through the battery. Although there is no direct conversion to IQ scores, it is possible to make approximations using the Wonderlic scores (e.g., a 21 Wonderlic score is approximately a 100 IQ score (average)).

Cognitive ability has shown to be a significant moderator in previous transfer of training research. Specifically, there have been positive predictor-transfer relationships between closed skill transfer (ρ = .41) (Blume, Ford, Baldwin, Huang, 2010). This posits that the smarter someone is, the more likely that they will positively transfer their training. However, the same researchers found equally significant findings that cognitive ability has a small, but negative relationship between the transfer of open skills (ρ = -.14). Due to possible moderating effects, the Wonderlic scores will serve as an approximation of participant's cognitive ability that can then be controlled for during analysis. The Wonderlic scores were examined as a possible covariate during analysis. Example questions from the Wonderlic Personnel Test - Q are found in Appendix C.

NASA-TLX (Task Load Index)

Each participant completed the pen and pencil version of the NASA-TLX (Task Load Index) at the conclusion of each of their training and transfer tasks. The NASA-TLX, primarily developed for the aviation domain, has seen wide use across domains and tasks and is the most commonly used subjective workload rating scale (Hart & Staveland, 1988). The NASA-TLX evaluates global workload, as well as six subscales of workload that include mental demand, physical demand, temporal demand, performance, effort, and frustration. During administration,

participants marked their level of perceived cognitive load on a scale of 1-100. The TLX is a quick and effortless way to attain perceived global and subscale workload information.

Attempting to infer perceived workload is necessary as some researchers have warned that individual differences in learning styles can affect whether part-task training or whole-task training is optimal for each individual trainee (Teague, Gittelman, & Park, 1994). The TLX measured approximate levels of workload across training tasks and for the transfer task to provide insight as to if one condition is inducing higher levels of workload than the other. Additionally, there should be marked changes in perceived workload between the control condition's training tasks and when they complete the relatively complex transfer task with variables that they have not been exposed to yet. Post-study analysis uses workload to examine the perceived "cognitive cost" of employing one condition of practice versus the other (part-task versus whole-task). As discussed in the Role of Workload section, the more free cognitive abilities, the better chance one has to mindfully abstract concepts to use in later tasks. The paper and pencil version of the NASA-TLX is found in Appendix D.

Paper Folding Test

The Paper Folding Test is a measure of spatial working memory (Ekstrom, French, Harman, & Dermen, 1976). Participants are shown a series of paper folds demonstrated on the first image. After the folds are simulated, a pencil is pushed through a portion of the folded paper, creating a hole that goes all the way through. The participant's task is to view the remaining images in the series that show possible solutions of what the paper looks like unfolded. The key marker is determining where the holes will end up once the paper is unfolded. There are two sides to the

test, both with ten questions. Participants are given three minutes for each side. The resulting score is graded out of twenty possible points.

The Paper Folding test was selected as it has been used in previous research efforts to measure spatial ability and has demonstrated significant relationships with puzzle-solving ability. Verdine and colleagues investigated a proposed link between jigsaw puzzle assembly and spatial ability and found a high positive correlation ($r(50) = .45$, $p < .01$) between the two on spatial abilities tests (Verdine, Troseth, Hodapp, & Dykens, 2008). As spatial ability is linked to performance on puzzle solving, it is an individual difference worth controlling for during analysis as it is likely a covariate. A copy of the Paper Folding test, and instructions is found in Appendix E.

Post-Participation Survey

A post-participation survey was developed and distributed to participants. This allowed them to convey if any common puzzle solving strategies were used during the study. Common solving strategies include working from the edges to the center, placing pieces in approximate locations based on the box cover image, and solving "mini-puzzles" with similar color markings. It is also possible that no particular strategy was used, and the participant just "solved it." This survey also included a free-response section for participants to list any and all puzzle abnormalities they encountered while engaging in the training and transfer tasks. The uncertainty variables noted at the conclusion of the study are valuable indicators of mental abstraction. Each participant's free-response was graded based on how many of the four uncertainty variables they identified in the transfer task. The results from this section will provide insight into the comparison of part-task and whole-task training methods. A concern of training design literature is that whole-task

methods can be overwhelming and does not focus on specific tasks (Wightman and Sistrunk, 1987; Teague, Gittelman, & Park, 1994). It is possible that participants in the whole-task condition may be overwhelmed by the multiple variables included at once, and may not mentally abstract each individual one. The free-response section allowed participants to note if they we able to discern the CDO characteristics included in their tasks. It was important to construct this section as a free-response entry, as other selection methods, such as multiple choice, or check all that apply tables, might lend credence to hindsight bias. All responses needed to come directly from the participant, in their own words, to get a clear idea of what exactly he or she noticed during the trials, not if "X uncertainty variable" was noticed. A copy of the post-participation survey is found in Appendix F.

Experimental Design

The research plan includes a one-way between-subjects design with three conditions: two experimental groups for a comparative evaluation and a control group. The two experimental groups comprise a part-task training condition and a whole-task condition. This provides insight into the question of how the training design, or condition of practice, facilitates mental abstraction, an open skill required for high-road transfer necessary for complex, dynamic, and ambiguous environments. The condition participants were in was randomized, the order of the puzzles they solved was randomized, and for the part-task condition, the variables they were exposed to first, second, and third, was also randomized.

The non-comparative evaluation (comparing the part-task or whole-task conditions with the control group) will answer the question if the intervention is effective at all. The comparative

evaluation, between the part-task condition and the whole-task condition, will assess the relative effectiveness of the two interventions with the same goal. It will answer the question of which intervention is more effective.

Table 1. Training conditions

	Training Task 1	**Training Task 2**	**Training Task 3**	**Transfer Task**
Control Condition	No Variables	No Variables	No Variables	All + Transfer Variable
Part-Task Condition	Missing Information	False/Spoofed Information	Extra Information	All + Transfer Variable
Whole-Task Condition	All Variables	All Variables	All Variables	All + Transfer Variable

Three primary Contested, Degraded, or Operationally Limited (CDO) environment characteristics were identified through literature review and discussion with Naval Flight Officer subject matter experts: missing information, extra information, or stimulated (spoofed/fake) information. These three variables are termed “uncertainty variables” throughout the remainder of this document. Participants in each condition completed three training tasks and one final transfer task. There was little instructional training as the majority of the “training” was incidental and organic. In other words, the interventions were implicitly incorporated into the tasks, rather than explicitly stated to the participants. Pre-training involved minimal instruction on the task itself. Puzzle completion is a common task of which many should already be familiar at a novice to intermediate level. However, previous experience with the puzzle completion, along with other trainee characteristic variables, was collected and accounted for.

Conditions and Uncertainty Variables

The study design included three conditions: two experimental (part-task, whole-task), and one control. Of the three previously described uncertainty variables (missing information, false information, and extra information), the part-task training condition was exposed to one of them at a time per task, while the whole-task training condition was exposed to all three simultaneously in each task, although in lower doses. The control condition was not exposed to any uncertainty variables until the final transfer task.

Table 2. Contested, Degraded, and/or Operationally Limited (CDO) / uncertainty variable descriptions

Missing information	X number of the original puzzle pieces were removed
False information	X number of the original puzzle pieces were replaced with similar looking replacements
Extra information	X number of the original puzzle pieces were duplicated, and included in addition to the original pieces

The part-task condition introduced the participant to a sole uncertainty variable designed into each of the three training tasks (see Table 1). Approximately 15% of the task was manipulated in each training task. In raw puzzle piece terms, six of the total thirty-six pieces was modified. For example, if the first uncertainty variable was "missing information," 15% (or six pieces) of the puzzle were not included for the participant to assemble. For the second variable, 15% of the puzzle pieces were false or spoofed. For this, six pieces were removed from the original puzzle and replaced with similar, but not identical looking pieces from another puzzle. For the last training task, 15% of the puzzle pieces were exact replicas and included in addition to the

original thirty-six pieces. As this variable is "extra information," the participants had a total of 40 pieces.

The whole-task condition included all three of the uncertainty variables, designed into each scenario. The level of variable intervention remained at 15%, but was split between the three uncertainty variable characteristics. For example, 5% (2 pieces) were missing entirely, 2 were replaced with similar but fake pieces, and 2 were exact copies of the original pieces. Participants in the whole-task condition were exposed to all variables, but in smaller doses, whereas the part-task participants were exposed to high doses of a singular variable per task.

The control condition simply had participants completing the training tasks without any of the uncertainty variables included until the transfer task. In essence, these participants solved the puzzles as normal.

After completing the initial three training tasks specific to the participant's condition, all participants completed the transfer task. The transfer task included all three original uncertainty variables, similar to the whole task condition, (missing information, false information, and extra information) as well as an extra transfer uncertainty variable. The final uncertainty variable is a cropping of the solution picture they use to solve the puzzle, or what the "picture on the box" would look like. The participants had access to the puzzle solution image to help them solve the puzzle, however, the picture was cropped so that they were missing the full view of the final result. An example of this is shown in Figure 5. This was not indicated prior to them engaging in the task. This variable is simulating the CDO environment characteristic of not having a full

representation of the overall operating picture, or going into a situation with a faulty intelligence brief. All training tasks in each condition were randomized in the order they are presented and counterbalanced to avoid asymmetric transfer effects.

Figure 6. An example of what the participant will see (top picture) versus the real solution image (bottom picture)

Independent Variables

The independent variable used in this experiment were task type: no training, part-task training, or whole-task training.

Dependent Variables

Four performance-based dependent variables were collected: time to complete, number correct (accuracy), subjective workload score, and number of variables identified.

Participants were instructed to complete each training and transfer task as quickly as they could. The researcher then started the timer until they indicated they were finished.

Number correct (or task accuracy) was scored based on if the correct pieces were in the correct places. A fully correct transfer task puzzle had 32 pieces correctly placed with 4 empty squares.

Perceived workload scores from the NASA-TLX were also used as a subjective dependent variable. Additional demographic data were used as dependent variables during post-study analysis to investigate equality across conditions (e.g., previous puzzle experience by condition ANOVA).

Lastly, a free-response section at the end of the transfer task asked the participants to list all puzzle abnormalities. This was scored out of 4, as there were 4 total uncertainty variables that all participants were exposed to.

Possible Covariates

Cognitive ability has shown to be a significant moderator in previous transfer of training research (Blume, Ford, Baldwin, and Huang, 2010). As stated previously, the Wonderlic Personnel Test results were used to control for variance related to cognitive ability. Other possible covariates

concerning transfer of open skills include puzzle experience and visuospatial ability (Dykens, 2002; Verdine, Troseth, Hodapp, & Dykens, 2008).

Procedure

Participants were recruited through UCF's SONA systems to participate in the study that lasted approximately 90 minutes. After obtaining consent, participants read a brief paragraph concerning the outline of the study, and what they would be doing for the next 90 minutes or so. If the participants did not have any questions, they were given the brief demographic survey to complete, as well as the Ishihara colorblind test.

Following the Ishihara test, participants were given instructions on the Wonderlic Personnel Test. This cognitive assessment has a hard time constraint of eight minutes. Participants were told to complete as many questions as they could in the given time. After that was complete, participants were given the Paper Folding Test instructions and sample questions. Once they acknowledged that they understood, they were given three minutes for the first part, and three minutes for the second part.

At the conclusion of the pre-test questionnaires, participants were randomly assigned to one of the three training conditions. The three training task puzzles and the transfer task puzzle were pre-setup according to condition and puzzle order. Participants were then told that their task is to solve a puzzle using the solution image provided. They were asked to complete the task as quickly as possible, as they were being timed. Lastly, they were instructed to let the experimenter know when they were done, or they felt that they could not complete any more of the puzzle.

This was repeated for all four puzzle tasks. All experimentation took place in the same testing environment under the same environmental conditions.

At the conclusion of each training task, and the transfer task, participants were instructed on how to fill out the NASA-TLX measure of workload, and were given a few moments to complete the paper and pencil version of it based on the puzzle task they just completed.

At the conclusion of the transfer task and the final NASA-TLX, participants were given the post-participation survey. After which, they were debriefed on the experiment and allowed to ask any questions, and then they were free to go.

CHAPTER FOUR: EXPERIMENT RESULTS

Preliminary Analysis

Prior to experiment analysis, a few preliminary analyses were conducted to identify outliers, as well as to ensure equality across conditions for various measures. An outlier analysis was conducted to remove any participants that scored uncharacteristically low or high on the performance task (specifically, time to complete). Two participants were identified and removed from future analyses with one clocking approximately thirty minutes to complete the final transfer task, while the other clocked close to five minutes. These two participants were outside two standard deviations, on the high and low end respectively. The mean score of all participants was 1065 seconds (~17 minutes) with a standard deviation of 305 seconds (~5 minutes). The removal of these two outliers left 79 cases to be analyzed (27 in the control group, 26 in the part-task group, and 26 in the whole-task group). An a priori power analysis indicated that 81 participants were required across the three groups to have 80% power for detecting a medium sized effect when employing the traditional .05 criterion of statistical significance, which was met. Analyses were performed using IBM SPSS 25 for Windows with an alpha level of .05 used for all statistical analyses (unless otherwise stated).

One-way ANOVAs were conducted to ensure equality across each condition for a variety of demographic and prior experience related measures. These measures included: grade point average (G.P.A.), puzzles completed in the last year, self-noted puzzle experience, puzzle solving confidence, Wonderlic IQ score, and the Paper Folding test of spatial ability. No significant differences were found, except for the Paper Folding test. Spatial ability was

identified as a possible covariate during literature review and is included as such in later analyses.

Analysis

Descriptives

The final participant pool comprised of 79 individuals recruited from the University of Central Florida (UCF) Institute for Simulation and Training (IST) SONA recruitment system. The average age was 22, as the clear majority were undergraduate students at the university, although the ages ranged from 18 to 42. Of these, 54 were female, while the remaining 25 were male. Table 3 breaks down participant gender by condition. All participants noted that he or she had normal or corrected vision, as well as all but one scoring perfect on the Ishihara colorblind test.

Table 3. Participants by gender
Participants by Gender in Each Condition

Condition	Male	Female	Total
Control	8	19	27
Part-Task	7	19	26
Whole-Task	10	16	26
	25	54	79

Covariates

Two possible covariates were identified in the literature review that could impact the results of this study: spatial ability, and general intelligence (cognitive ability). Both were examined by running correlations between them, and the four primary performance measures.

To account for general intelligence, two measurements, grade point average (G.P.A) and the Wonderlic Exam scores were recorded and used in the correlations. Table 4 shows that G.P.A. was not significantly correlated with any other measure, and interestingly, not correlated with the Wonderlic Score, which is claimed to measure general cognition. As neither G.P.A. nor the Wonderlic Score were significantly correlated with any performance measures, they were excluded as covariates.

However, as expected, spatial ability, as measured by the Paper Folding test, was significantly correlated with three of the four performance measurements (Time to Complete, Number Correct, and Variables Noted). Pearson's partial correlation was run to assess the relationship between spatial ability and the primary dependent variables. A bivariate Pearson's correlation established that there was a strong, negative, statistically significant relationship between spatial ability and Transfer Time (or Time to Complete), $r(79) = -.521$, $p < .01$. There was also a positive significant relationship between spatial ability and Number Correct, $r(79) = .328$, $p < .01$. Finally, another positive significant relationship was found between spatial ability and the Variables Noted, $r(79) = .387$, $p < .01$. These moderate and strong significant relationships between spatial ability and the performance measures indicated that spatial ability should be controlled for as a covariate in follow-on analyses. There were no significant differences between genders on the spatial ability scores.

Table 4. Table of Correlations for Main Variables
Table of Correlations for Main Variables

	G.P.A	Wonderlic Score	Paper Folding	Transfer (Time)	Number Correct	TLX Score	Variables Noted
G.P.A	--	.017	.089	-.109	-.194	-.033	-.126
Wonderlic Score	.017	--	.187	-.204	-.038	.064	.191
Paper Folding	.089	.187	--	-.521**	.328**	-.122	.387**
Transfer (Time)	-.109	-.204	-.521**	--	-.200	.193	-.299**
Number Correct	-.194	-.038	.328**	-.200	--	.151	.411**
TLX Score	-.033	.064	-.122	.088	-.163	--	-.225*
Variables Noted	-.126	.191	.387**	-.299**	.411**	-.225*	--

Note: $^{*}p < .05$, $^{**}p < .01$, two-tailed. $N = 79$.

Performance Measures

The participants were randomly assigned to a condition (control, part-task, or whole-task) and the order in which they were exposed to the puzzles was also randomized. Further, the uncertainty variables were randomized in the part-task condition (e.g., the ocean themed puzzle was not always the missing pieces puzzle, sometimes it had extra pieces, and sometimes fake pieces). The following analyses reports on the participants' performance measures and the subjective measures of the experiment.

Completion Time (Transfer Time)

One of the primary dependent variables considers how long it took the participant to finish the final puzzle (or transfer task). Table 5 shows the raw completion time in seconds by experimental condition. Participants in the control condition completed the transfer task with a

mean time of 1245 seconds (*SD* = 321). The participants in the part-task condition completed the transfer puzzle with a mean time of 967 seconds (*SD* = 235). Finally, the whole-task participants completed the transfer task with a mean time of 978 seconds (*SD* = 275).

Table 5. Raw Time (in seconds) to Complete the Performance Task by Condition
Raw Time (in seconds) to Complete the Performance Task by Condition

Condition	*M*	*SD*
Control	1245	321
Part-Task	967	235
Whole-Task	978	275

Figure 7 visualizes Table 5 as a stem-and-leaf boxplot, accompanying the means with quartile measurements.

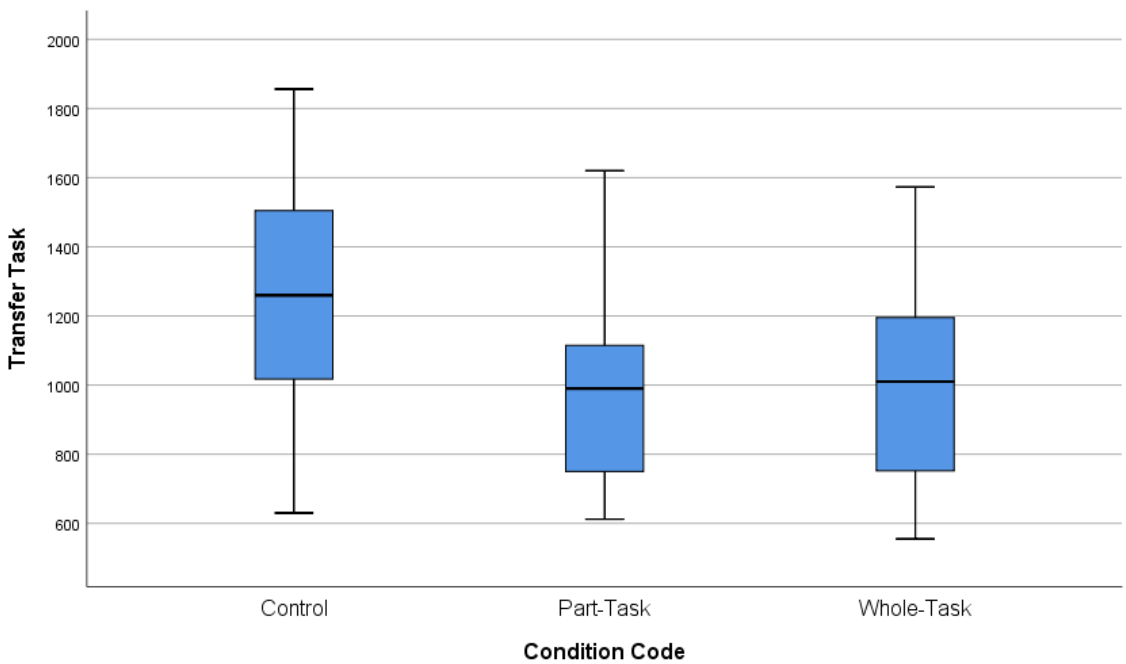


Figure 7. Stem-and-leaf boxplot of transfer task time by condition code

Using time to complete as the performance variable, there was a statistically significant main effect found for training condition $F(2,76) = 8.444$, $p = .001$, $\eta_p^2 = .182$. Planned comparisons using the Bonferroni correction revealed that the control group statistically differed from the both the part-task group ($p = .002$) and the whole-task group ($p = .003$). There was no statistically significant difference between the part-task and the whole-task group ($p = 1.000$).

However, the above results are reported in terms of unadjusted times and scores. Due to the impact of spatial ability as a covariate, it is necessary to control for it in further analyses. As such, an Analysis of Covariance (ANCOVA) will be used to determine adjusted scores of time to complete, controlling for spatial ability.

The Pearson's correlation established that there was a strong, negative, statistically significant relationship between spatial ability and Transfer Time (or Time to Complete), $r(79) = -.521$, $p < .01$. As a negative relationship, as spatial ability scores increase (better), time to complete is lower (better). The following mean and standard deviation reporting are adjusted, controlling for spatial ability (Table 6).

Table 6. Adjusted Time to Complete the Performance Task by Condition, Controlling for Spatial Ability
Adjusted Time (in seconds) to Complete the Performance Task by Condition, Controlling for Spatial Ability

Condition	*M*	*SD*
Control	1198	321
Part-Task	1050	235
Whole-Task	943	275

Figure 8 visualizes adjusted data as a bar chart, accompanying the means with error bars. Participants in the control condition completed the transfer task with an adjusted time of 1198 seconds ($SE = 48$). The participants in the part-task condition completed the transfer puzzle with an adjusted mean time of 1050 seconds ($SE = 50$). Lastly, the whole-task participants completed the transfer task with an adjusted mean time of 943 seconds ($SE = 48$).

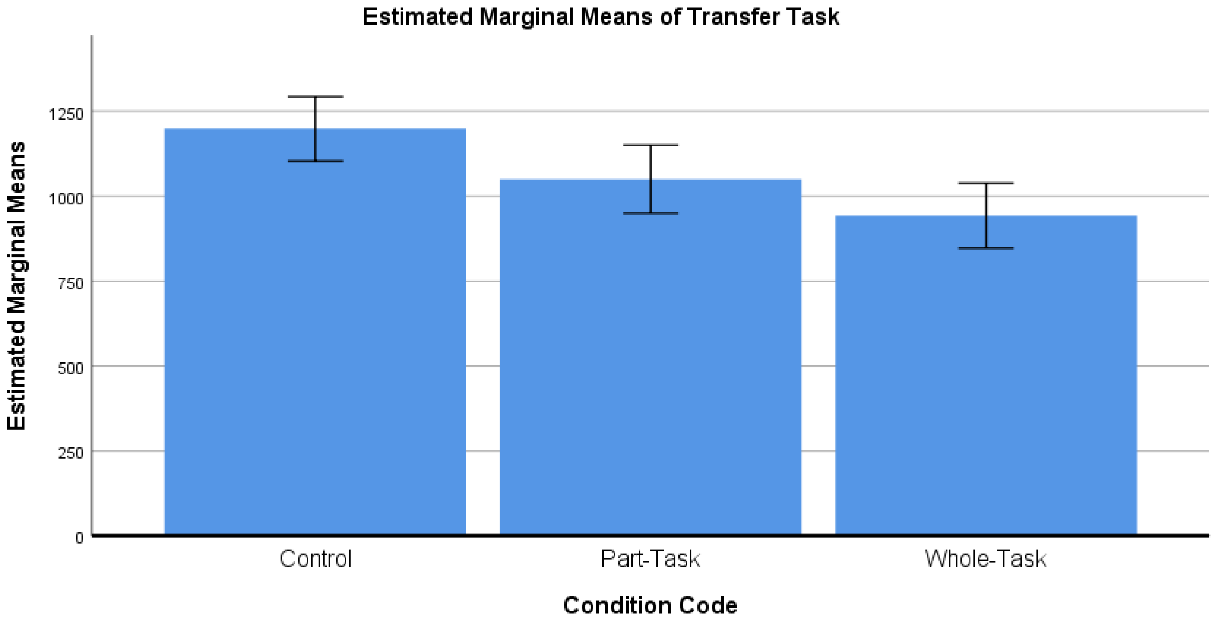


Figure 8. Adjusted transfer task time by condition code, with standard error bars

An ANCOVA was run to determine the effect of condition code (control, part-task, whole-task) on time to complete the transfer task, after controlling for spatial ability. After controlling for spatial ability, there was a statistically significant difference in time to complete on the transfer task between the conditions, $F(2,76) = 7.365$, $p = .001$, $\eta_p^2 = .164$. Post hoc analysis was performed with a Bonferroni adjustment revealing that time to complete was significantly greater in the control condition vs the whole-task condition (M_{diff} = 255.3 secs, 95% CI [91.8, 418.6], $p < .001$). There was no statistically significant difference between the control condition and the part-task condition (M_{diff} = 148.0 secs, 95% CI [-26.9, 322.9], $p = .125$), or between the two experimental, part-task and whole-task, groups (M_{diff} = 107.2 secs, 95% CI [-67.2, 281.6], $p = .409$). Table 7 (next page) shows the adjusted and unadjusted means for time to complete, using spatial ability as a covariate.

Table 7. Adjusted and Unadjusted Time to Complete (in seconds) by Condition, Controlling for Spatial Ability as a Covariate

Adjusted and Unadjusted Time to Complete (in seconds) by Condition, Controlling for Spatial Ability as a Covariate

		Unadjusted		*Adjusted*	
Condition	*N*	*M*	*SD*	*M*	*SE*
Control	27	1245	321	1198	48
Part-Task	26	967	235	1050	50
Whole-Task	26	978	275	943	48

As participants in the whole-task condition significantly outperformed (lower time to complete) participants in the control condition, Hypothesis 1, that there will be a significant difference in performance between the control condition and both experimental conditions, is partially supported. However, Hypothesis 2, that there will be a significant difference in the performance (time) between the part-task training and whole-task training conditions was not supported. Finally, Hypothesis 3 predicted that participants in the part-task condition will outperform participants in the whole-task condition and the control condition. Although the part-task condition participants had better mean scores than the control condition, it was not statistically significant. Hypothesis 3 remains unsupported.

Additional analyses looked at the time to complete scores, by condition, for each of the training tasks, as well as the transfer task. The results are shown in table 8 below. This is further discussed in the discussion section on time to complete.

Table 8. Time by condition by training task
Time to complete scores of each Task, by Condition, adjusted for spatial ability (in seconds)

		Task 1		*Task 2*		*Task 3*		*Transfer Task*	
Condition	*N*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Control	27	722	62	677	61	633	52	1198	48
Part-Task	26	853	60	771	65	681	55	1050	50
Whole-Task	26	1015	63	839	62	789	52	943	48

<u>Number Correct</u>

Another measure of task performance is the number of correct answers a participant provided, or in this case, how fully they completed the final puzzle with all the correct pieces in the correct assigned spaces. On the final task, a 100% score was possible by placing all 32 pieces in the correct spots. Table 9 shows the number of correct pieces placed at the end of the transfer task.

Table 9. Number correct on transfer task by condition
Number of Correct Pieces Placed on the Transfer Task by Condition, Unadjusted

Condition	*M*	*SD*
Control	28.1	3.9
Part-Task	30.2	2.5
Whole-Task	28.0	5.0

Note. Means were out of 32 possible correct answers.

Figure 9 visualizes Table 9 as a stem-and-leaf boxplot, accompanying the means with quartile measurements.

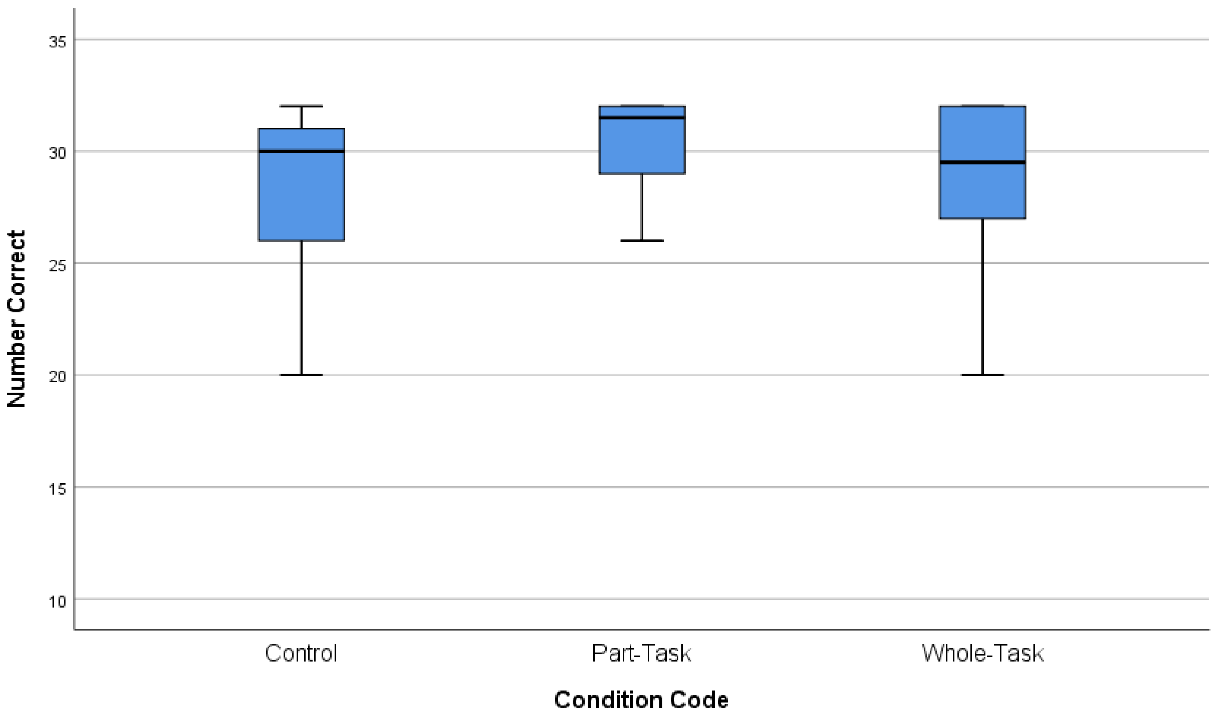


Figure 9. Number correct on transfer task by condition code

As spatial ability was identified as a possible covariate, a 3x1 ANCOVA was conducted to assess the effects of condition (control, part-task, and whole-task) on overall task performance (number correct). Table 10 (next page) shows the number of correct pieces, adjusted for spatial ability, placed at the end of the transfer task.

Table 10. Number correct on the transfer task by condition, adjusted for spatial ability
Number of Correct Pieces Placed During the Transfer Task by Condition, Adjusted for Spatial Ability

Condition	*M*	*SD*
Control	28.7	3.9
Part-Task	29.6	2.5
Whole-Task	28.2	5.0

Note. Means were out of 32 possible correct answers.

Figure 10 visualizes Table 10 as a bar chart, accompanying the means with error bars. Participants in the control condition completed the transfer task with an adjusted number of correct pieces of 28.7 (*SE* = .75). The participants in the part-task condition completed the transfer puzzle with an adjusted number of correct pieces of 29.6 (*SE* = .79. Finally, the whole-task participants completed the transfer task with an adjusted number of correct pieces of 28.2 (*SE* = .76).

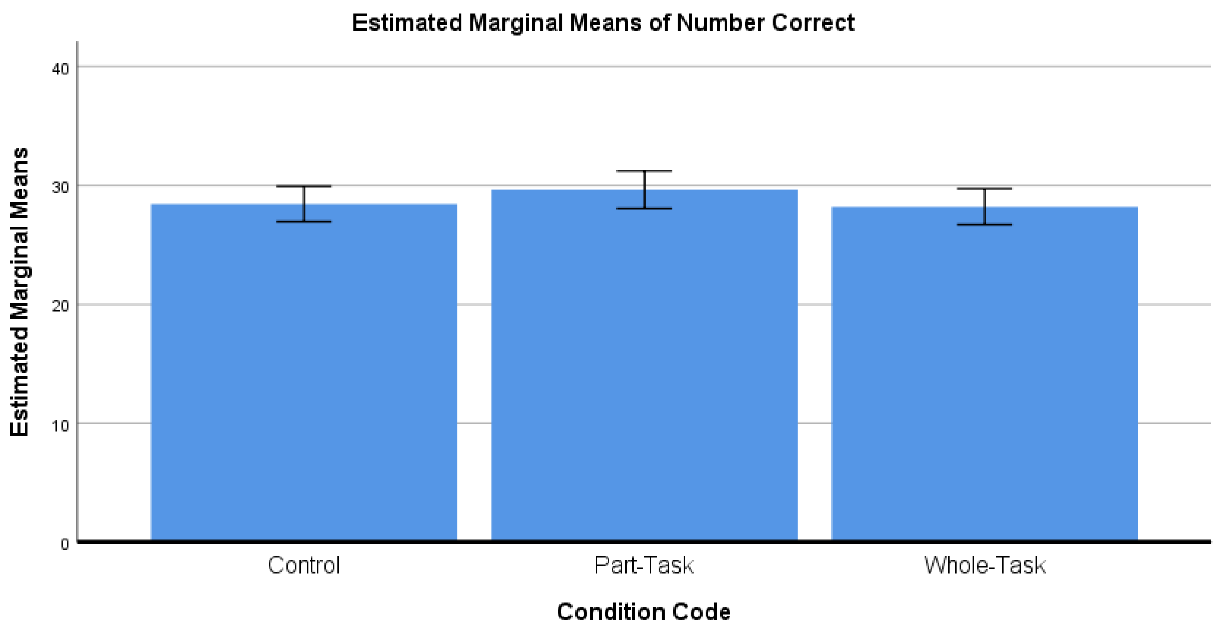


Figure 10. Adjusted number correct on the transfer task with standard error bars

The ANCOVA revealed no significant main effect for training condition, $F(2,76) = .907$, $p = .408$ $\eta_p^2 = .024$.). Table 11 shows the adjusted and unadjusted means for number correct, using spatial ability as a covariate.

Table 11. Adjusted and unadjusted number correct by condition
Adjusted and Unadjusted Number Correct by Condition, Controlling for Spatial Ability as a Covariate

		Unadjusted		*Adjusted*	
Condition	*N*	*M*	*SD*	*M*	*SE*
Control	27	28.1	3.9	28.7	.75
Part-Task	26	30.2	2.5	29.6	.79
Whole-Task	26	28.0	5.0	28.2	.76

For the transfer task, no statistically significant main effect for condition was found. Although the participants in the part-task group trended better by placing, on average, more correct pieces, in both adjusted and raw scores, it was not a statistically significant difference. Neither Hypothesis 1, 2, or 3 was supported when examining number correct as a performance variable.

The next section examines another performance variable, Variables Identified, and the role training condition plays on it.

Variables Identified

Another dependent variable, uncertainty Variables Identified, relates to how many, of the four, variables participants were able to list in a free-response section after the transfer task. As the participants were never explicitly told what any of the uncertainty variables were (missing

pieces, false pieces, extra pieces, cropped image), the participants were asked to note any abnormalities they noticed with the final puzzle. Recall, that the final puzzle, or transfer task, included all four uncertainty variables, regardless of condition, making the maximum score being four. This section explores how well each condition's participants did in comprehending and identifying the four uncertainty variables.

Table 12 shows the raw number of variables identified by condition. Participants in the control condition identified a mean of 2.4 variables (*SD* = 1.0). Participants in the part-task condition identified a mean of 3.5 variables (*SD* = .58). Lastly, the whole-task participants identified a mean of 2.6 variables (*SD* = .98).

Table 12. Number of variables identified by condition
Number of Variables Identified by Condition, Unadjusted

Condition	*M*	*SD*
Control	2.4	1.0
Part-Task	3.5	.58
Whole-Task	2.6	.98

Note. The highest number identified possible was 4.

Figure 11 visualizes Table 12 as a stem-and-leaf boxplot with unadjusted data, accompanying the means with quartile measurements.

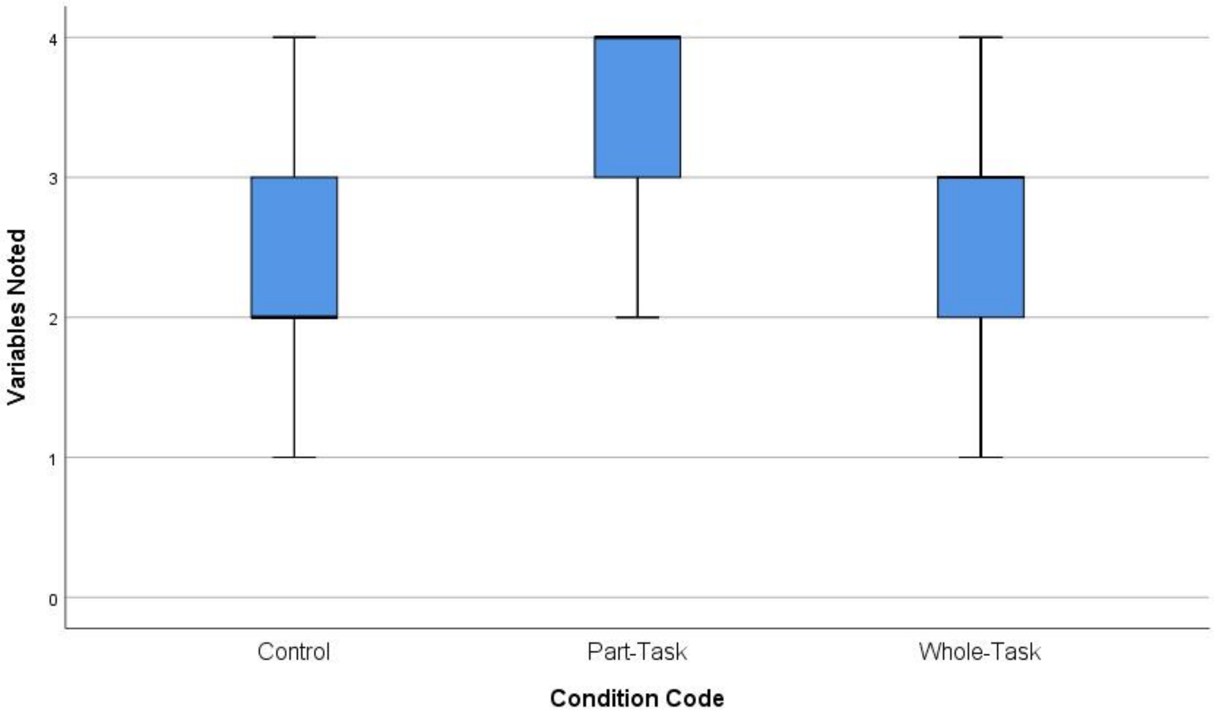


Figure 11. Number of variables identified by condition

However, the above results are reported in terms of unadjusted variables reported. Due to the impact of spatial ability as a covariate, it is necessary to control for it in when looking further at this performance measurement. As such, an Analysis of Covariance (ANCOVA) was used to determine adjusted variables reported, controlling for spatial ability.

The Pearson's correlation established that there was a positive, statistically significant relationship between spatial ability and Variables Reported, $r(79) = .387$, $p < .01$. Considering this positive relationship, as spatial ability scores increased (improved), the number of reported

uncertainty variables also increased (improved). The following mean and standard deviation reporting are adjusted, controlling for spatial ability and are shown in Table 13.

Table 13. Number of variables identified by condition, adjusted for spatial ability
Number of Variables Identified by Condition, Adjusted, Controlling for Spatial Ability

Condition	*M*	*SE*
Control	2.4	1.0
Part-Task	3.4	.58
Whole-Task	2.6	.98

Figure 12 (next page) visualizes Table 13 as a bar chart, accompanying the means with error bars. Participants in the control condition identified approximately 2.4 variables (adjusted) after the transfer task (*SE* = .17). The participants in the part-task condition identified approximately 3.4 variables (adjusted) after the transfer task (*SE* = .18). Finally, the whole-task participants identified approximately 2.6 variables (adjusted) after the transfer task (*SE* = .17).

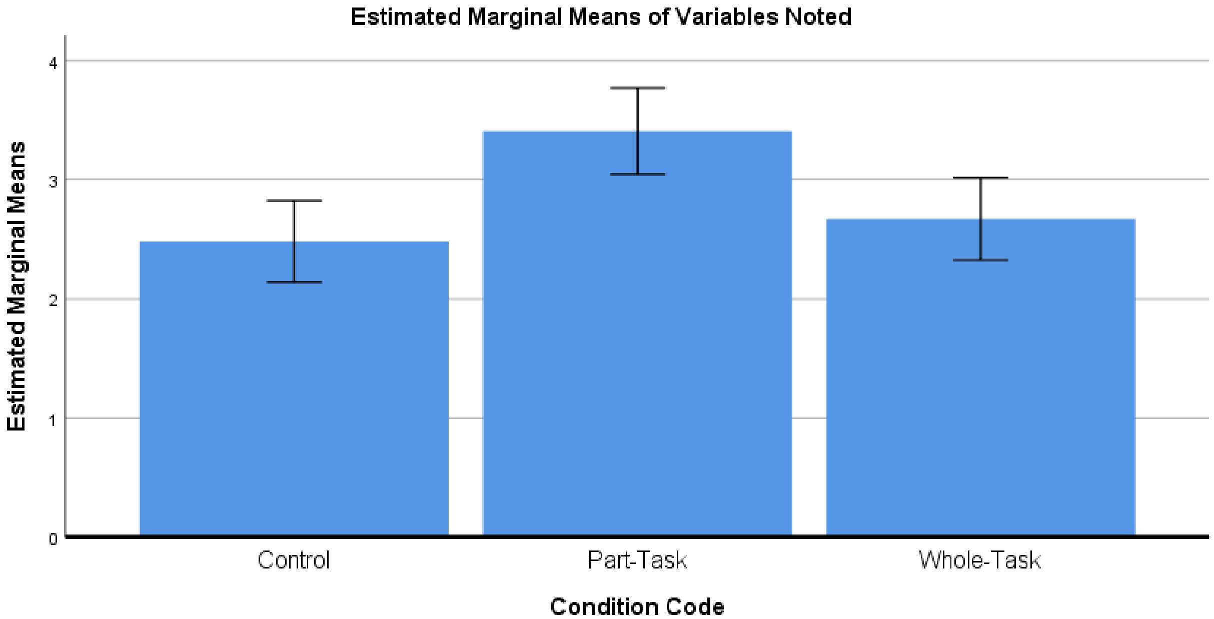


Figure 12. Variables noted by condition, adjusted for spatial ability, with standard error bars

An ANCOVA was run to determine the effect of condition code (control, part-task, whole-task) on Variables Identified, after controlling for spatial ability. There was a statistically significant difference in Variables Identified on the transfer task between the conditions, $F(2,76) = 7.018$, $p = .002$, $\eta_p^2 = .158$. A post hoc analysis performed with a Bonferroni adjustment revealed that Variables Identified was statistically significant and greater between the control condition and the part-task condition, with the part-task participants reporting significantly more variables than their control condition counterparts ($M_{\text{diff}} = .93$ variables, 95% CI [.28, 1.6], $p = .002$). Additionally, there was a statistically significant difference between the two experimental groups (part-task and whole-task), with the part-task participants reporting significantly more variables than their whole-task participant counterparts ($M_{\text{diff}} = .74$ variables, 95% CI [.11, 1.4], $p = .016$). Post hoc analysis showed no statistically significant difference between the control condition and

the whole-task condition (M_{diff} = .19 variables, 95% CI [-.40, .78], p = 1.00). Table 14 shows the adjusted and unadjusted means for Variables Identified, using spatial ability as a covariate.

Table 14. Adjusted and unadjusted variables identified by condition, controlling for spatial ability
Adjusted and Unadjusted Variables Identified by Condition, Controlling for Spatial Ability as a Covariate

		Unadjusted		*Adjusted*	
Condition	*N*	*M*	*SD*	*M*	*SE*
Control	27	2.4	1.0	2.4	.17
Part-Task	26	3.5	.58	3.4	.18
Whole-Task	26	2.6	.98	2.6	.17

As participants in the part-task condition significantly outperformed (more variables identified) participants in the control condition, Hypothesis 1, that there will be a significant difference between the control condition and both experimental conditions, is partially supported. Partial, in that only the part-task condition was significantly different than the control condition (the whole-task condition, was not). Hypothesis 2, that there will be a significant difference in the performance (variables identified) between the part-task training and whole-task training conditions was supported, as participants in the part-task condition identified significantly more variables than their whole-task counterparts. Finally, Hypothesis 3 predicted that participants in the part-task condition will outperform participants in the whole-task condition and the control condition. This was also supported, as the part-task condition participants performed significantly better than both the control, and whole-task conditions.

The above Variables Identified section looks at how many total variables were listed in the free-response section. This analysis explores the number of participants in each condition that noted

the “cropped” uncertainty variable. This is of significance as this was a “transfer task only” variable that was only exposed to them in the final task. Table 15 shows the cross-tabulation of the number of participants in each condition that noticed and wrote down that the “cropped” uncertainty variable was present in the final task.

Table 15. Participants in Each Condition Reporting the "Cropped" Uncertainty Variable
Crosstabulation of Participants in each Condition Reporting the “Cropped” Uncertainty Variable

Condition	*Yes*	*Total*
Control	15	27
Part-Task	18	26
Whole-Task	13	26

In the control condition, 15 out of the 27 reported noticing that the final solution image was cropped or manipulated in some way. In the part-task condition, 18 participants out of 26 noted the “cropped” uncertainty variable in the transfer task. Lastly, 13 participants out of 26 in the whole-task condition noted the “cropped” uncertainty variable in the transfer task.

The last section will examine the role of the final performance variable, the participant’s reported NASA TLX scores, and the effect training condition has on those scores.

NASA TLX Workload Scores

A final, but important performance variable to capture, is the participants’ subjective NASA TLX workload scores. This workload scale measures and records reported mental demand, physical demand, temporal demand, performance, effort, and frustration. Although each can be

measured independently, when calculated together, they result in a measure of overall workload. It is this measure of overall workload that is presented in this section. A lower number on the TLX scores indicates a lower perceived workload, while a higher number, indicates that the participant was experiencing a higher perceived workload. Recall that participants were encouraged to finish the puzzles as accurately and quickly as possible, so they likely perceived time and accuracy pressure while completing the task. This section explores the perceived workload scores of the participants in each of the conditions.

Table 16. Perceived overall workload TLX scores by condition on the Transfer Task
Perceived Overall Workload TLX Scores by Condition

Condition	*M*	*SD*
Control	47.4	14.9
Part-Task	32.9	15.7
Whole-Task	45.2	13.1

Note. The TLX scores range from 0 – 100, with 100 being highest workload.

Table 16 shows the perceived overall workload TLX scores reported by condition. Contrary to prior analyses, no adjusted scores are presented to control for spatial ability as this pertains to subjective workload, and due to spatial ability not being correlated with the TLX scores. Participants in the control condition reported a mean workload score of 47.4 (*SD* = 14.9) out of 100. Participants in the part-task condition reported a mean workload score of 32.9 (*SD* = 15.7). Participants in the whole-task condition reported a mean workload score of 45.2 (*SD* = 13.1).

Figure 13 visualizes Table 16 as a stem-and-leaf boxplot, accompanying the means with quartile measurements.

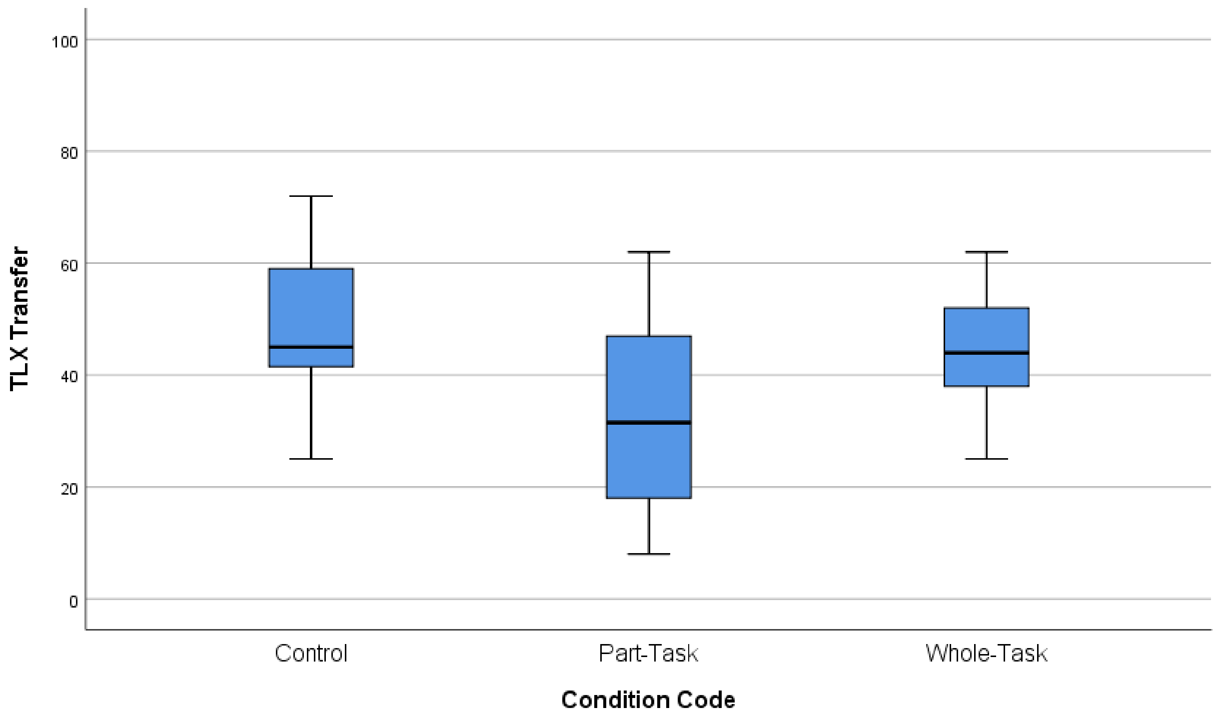


Figure 13. Perceived overall TLX workload scores by condition

An ANOVA was run to determine the effect of condition code (control, part-task, whole-task) on overall workload scores. There was a statistically significant difference in overall workload scores on the transfer task between the conditions, $F(2,76) = 7.454$, $p = .001$, $\eta_p^2 = .164$. A post hoc analysis was performed with a Bonferroni adjustment revealing that workload scores were statistically significantly lower in the part-task condition vs the control condition ($M_{\text{diff}} = -14.5$ TLX score, 95% CI [-24.3, -4.64], $p = .002$). Additionally, workload scores were significantly lower in the part-task condition vs the whole-task condition ($M_{\text{diff}} = -12.2$ TLX score, 95% CI [-

22.6, -2.30], p = .010). There was no statistically significant difference between the control condition and the whole-task condition (M_{diff} = 2.25 TLX score, 95% CI [-7.59, 12.1], p = 1.00).

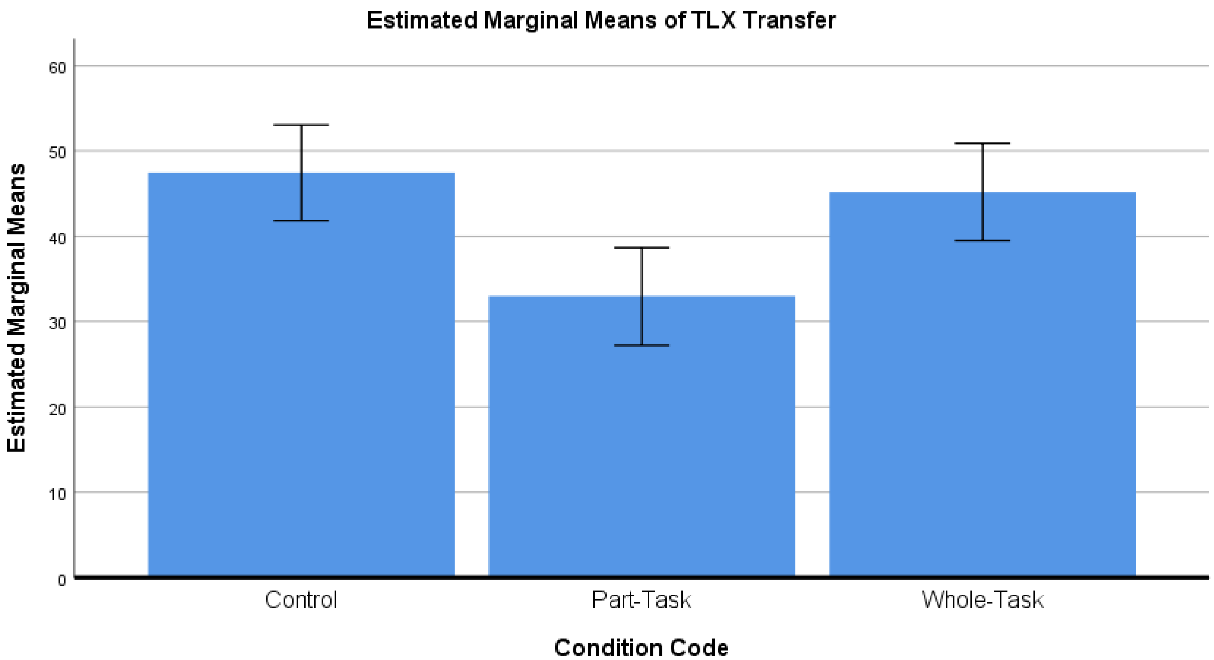


Figure 14. Estimated marginal means of TLX workload scores by condition with standard error bars

Participants in the part-task condition reported significantly lower overall workload scores than the other two conditions (control and whole-task). As such, Hypothesis 1, that there will be a significant difference between the control condition and both experimental conditions, is partially supported. Partial, in that only the part-task condition was significantly different than the control condition (the whole-task condition, was not). Hypothesis 2, that there will be a significant difference in the performance (TLX scores) between the part-task training and whole-task training conditions was supported, as participants in the part-task condition reported significantly less workload than their whole-task counterparts. Finally, Hypothesis 3 predicted that participants in the part-task condition will outperform participants in the whole-task

condition and the control condition. This was also supported, as the part-task condition participants reported significantly lower overall workload than both the control, and whole-task conditions.

The above analysis looks only at the final TLX scores, where each participant was given the exact same transfer task puzzle. It is also informative to explore the perceived workload scores of the participants in the different conditions during the three training tasks. Table 17 reports the breakdown of workload, by condition, on each of the three initial tasks.

Table 17. NASA TLX Scores of All Tasks, by Condition
NASA TLX Scores of each Task, by Condition

		Task 1	*Task 2*	*Task 3*	*Transfer Task*
Condition	*N*	*M*	*M*	*M*	*M*
Control	27	30.1	29.2	28.4	47.4
Part-Task	26	33.1	33.3	29.5	32.9
Whole-Task	26	39.2	36.8	36.7	45.2

Note. The TLX scores range from 0 – 100, with 100 being highest workload.

Participants in the control condition reported a mean workload score of 30.1 (*SE* = 3.1) on task one, 29.2 (*SE* = 2.8) on task two, 28.4 (*SE* = 2.1) on task three, and 47.4 (*SE* = 2.8) on the transfer task. Participants in the part-task condition reported a mean workload score of 33.1 (*SE* = 2.9) on task one, 33.3 (*SE* = 2.4) on task two, 29.5 (*SE* = 3.1) on task three, and 32.9 (*SE* = 3.1) on the transfer task. Participants in the whole-task condition reported a mean workload score of 39.2 (*SE* = 1.8) on task one, 36.8 (*SE* = 2.2) on task two, 36.7 (*SE* = 2.5) on task three, and 45.2

(*SE* = 2.5) on the transfer task. Figure 15 visualizes the scores from Table 15. All of the TLX scores are out of 100, with the lower scores indicating less workload.

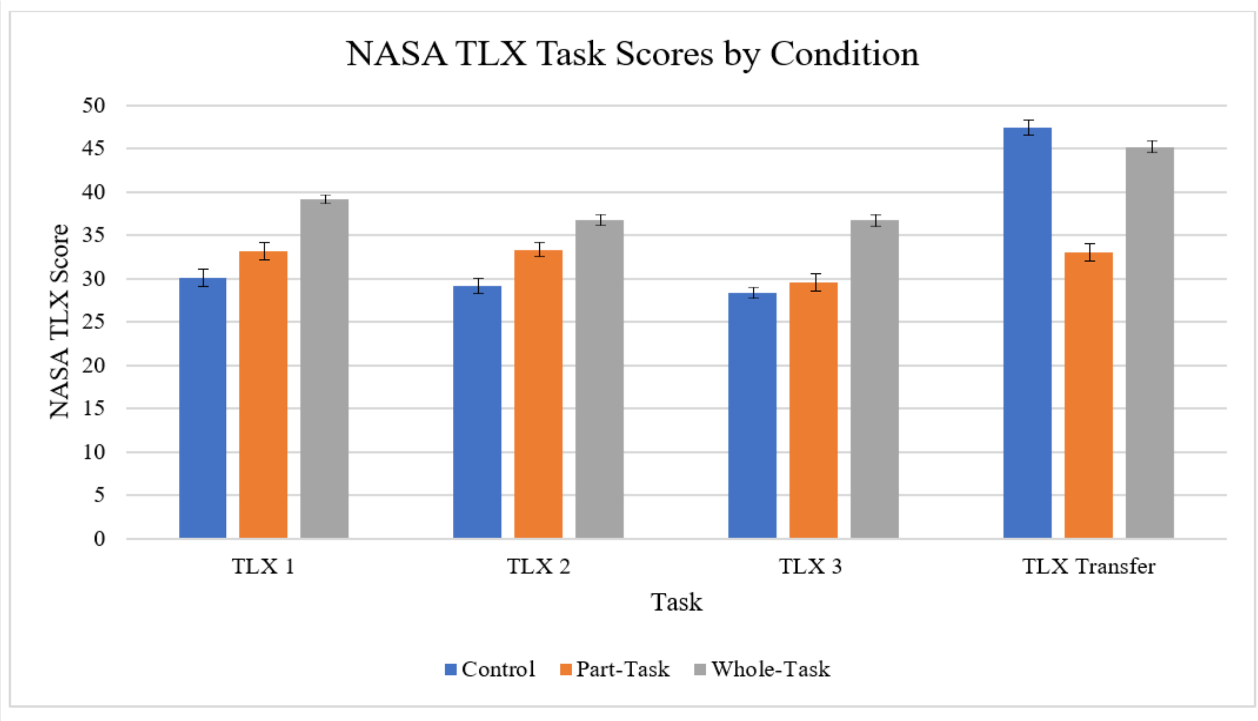


Figure 15. NASA TLX Scores by Condition for All Tasks

As indicated by Table 17 and Figure 15, the control condition reported the lowest perceived workload for the initial three tasks (no manipulation at all) and had the largest, predictable, increase when reporting on the transfer task (four uncertainty variables). The part-task condition was largely falling between the control and whole-task condition scores for the initial three tasks. Recall that these participants were only introduced to one uncertainty variable per puzzle. However, the part-task condition had the least change between the training tasks and the transfer task, as the scores were relatively uniform across all of the tasks. The whole-task condition participants reported the highest perceived workload for each of the initial three tasks, as

expected. These participants were given puzzles with three uncertainty variables in each, so it was expected that they would find it more difficult. However, for the transfer task, which only included one additional uncertainty variable, there was a discernable increase in the whole-task condition.

Additional t-tests were run between the three conditions to determine if the workload means were significantly different for each of the training scenarios. For training scenario one, these are the results of the t-tests. An independent-samples t-test was conducted to compare reported workload in training scenario one in the control and part-task conditions. There was not a significant difference in the scores for the control condition (M = 30.1, SE = 3.1) and the part-task condition (M = 33.1, SE = 2.9); $t(51) = -.54$, $p = 0.59$. Running the same test for the control condition (M = 30.1, SE = 3.1), and the whole-task condition (M = 39.2, SE = 1.8) yielded a significant result at $t(51) = -2.3$, $p = 0.023$. The last comparison between the part-task condition (M = 33.1, SE = 2.9) and the whole-task condition (M = 39.2, SE = 1.8) did not yield a significant result at $t(50) = -1.8$, $p = 0.083$.

For training scenario two, these are the results of the t-tests. An independent-samples t-test was conducted to compare reported workload in training scenario two in the control and part-task conditions. There was not a significant difference in the scores for the control condition (M = 29.2, SE = 2.8) and the part-task condition (M = 33.3, SE = 2.4); $t(51) = -1.1$, $p = 0.28$. Running the same test for the control condition (M = 29.2, SE = 2.8), and the whole-task condition (M = 36.8, SE = 2.2) yielded a significant result at $t(51) = -2.1$, $p = 0.038$. The last comparison

between the part-task condition (*M* = 33.3, *SE* = 2.4) and the whole-task condition (*M* = 36.8, *SE* = 2.2) did not yield a significant result at *t*(50) = -1.1, *p* = 0.28.

These are the results of the t-tests for the third and final training scenario. An independent-samples t-test was conducted to compare reported workload in training scenario three in the control and part-task conditions. There was not a significant difference in the scores for the control condition (*M* = 28.4, *SE* = 2.1) and the part-task condition (*M* = 29.5, *SE* = 3.1); *t*(51) = -.31, *p* = 0.76. Running the same test for the control condition (*M* = 28.4, *SE* = 2.1), and the whole-task condition (*M* = 36.7, *SE* = 2.5) yielded a significant result at *t*(51) = -2.6, *p* = 0.014. The last comparison between the part-task condition (*M* = 29.5, *SE* = 3.1) and the whole-task condition (*M* = 36.7, *SE* = 2.5) did not yield a significant result at *t*(50) = -1.8, *p* = 0.078.

CHAPTER FIVE: DISCUSSION

This experiment sought to address if part-task or whole-task training methods would foster mindful abstraction from training tasks to a transfer task. The goal being to be able to recommend one of the training schedules over another when training trainees to operate in contested, degraded, or operationally-limited environments. A core issue is that when training for uncertain environments, one cannot train for every scenario that could occur, rather the instructors must pick and choose which uncertainty tasks to train given the allotted time. For example, in the past decade, throughput during training for military trainees has been quick, with limited training time and resources. When met with these conditions, should instructors provide many varied scenarios, each focusing on one aspect of the uncertain environment, or should they provide scenarios that contain many aspects of the uncertain environment, but at lower exposure of each individually? This experiment attempted to provide insight to this problem through a low-tech training task.

The three conditions replicated different training schedules that might be appropriate to recommend. The control condition exposed participants to zero uncertainty variables throughout the three training tasks, and then exposed them to four simultaneously for the transfer task. The part-task condition exposed trainees to only one uncertainty variable per training task. Each training task in this condition comprised a different uncertainty variable than the one before with no overlap. Then, the transfer task (which was the same to all groups), exposed them to “doses” of the original three variables they were exposed to, while also introducing a novel, fourth variable. Finally, the whole-task condition exposed trainees to small doses of three uncertainty

variables during each training task. For all three training tasks, they saw the same three variables represented. For the transfer task, they saw those same three variables, along with the novel fourth variable. With those conditions set, the question became which, if any, would produce better performance results.

The results were viewed through four distinct performance variables: time to complete, number correct (or accuracy), uncertainty variables identified, and overall TLX workload scores. After ensuring equality across conditions for a variety of the demographic and prior experience related measures, the results indicated that there was a condition that produced more favorable results than the others on some of the performance variables. The part-task condition significantly outperformed the whole-task and control conditions on the variables identified measure, as well as the overall workload. For the number correct or accuracy measure, there was a notion of a glass ceiling effect with all conditions eventually completing most of the final transfer puzzle. As such, there was no significant main effect for training condition for this measure. Lastly, for the time to complete measure, there was a significant difference between the control group and the whole-task group, in that the whole-task group completed the transfer puzzle significantly faster than the control condition. However, there was no significant difference between the part-task and the whole-task condition, or the part-task and the control condition.

The results of each of these performance measures are examined in further detail in the following sections, while also noting relevance to the theoretical backing and literature. Support or rejection of each of the research hypotheses will also be discussed.

Completion Time

The performance variable "Completion Time" measured how long, in seconds, it took each participant to complete the final puzzle task as much as he or she could. It was anticipated that the control group would take the most amount of time to complete it, as it was their first time being exposed to all of the uncertainty variables (four in total). After completing three normal puzzles, they were given a chaotic puzzle and spent much of their completion time trying to make sense of it. This is shown in the results, and was expected. The approximate mean time to complete for the control condition was 20 minutes. However, it was less clear how the two experimental conditions (part-task and whole-task) would perform on the time component of the task. As the results indicate, the whole-task condition participants finished the task in the shortest amount of time with an approximate mean completion time of 15 minutes and 45 seconds. This was significantly different than the control condition's 20 minutes, but was not significantly different from the part-task condition, which had a mean time to complete of approximately 17 minutes and 30 seconds. There was also no significant difference between the part-task condition and the control condition, although the mean score was about 2 minutes and 30 seconds faster for the part-task condition.

What can be stated is that exposing trainees to three of the four uncertainty variables, multiple times (whole-task), prior to the transfer task resulted in significantly lowered time, than not exposing them to any at all (control). Even though the part-task condition was not significantly different from the control condition, it does appear that exposing trainees to uncertainty variables prior to the transfer task is favorable, and will enable the trainees to complete the task in a

quicker fashion. One of the takeaways from these results is that prior training and exposure, in some manner, will reduce overall time to complete. However, this is just one of the performance variables, and should be examined in the larger scope of the other three.

In terms of cognitive load, the control condition objectively had the most new factors to deal with, hence the longest time. The whole-task condition participants on the other hand, had the most prior and similar exposure to the transfer task, as each of their training puzzles had all three uncertainty variables. For the final transfer puzzle, they were already familiar with three out of the four variables, and had to adjust to the novel variable. Yet, in the following sections, further results indicate that adjusting to the one novel variable was more difficult for them, possibly due to their whole-task training schedule. Finally, the part-task condition falls somewhat in between the control (no prior experience) and the whole-task condition (most similar experience), and at least on the time variable, fell in between the other two.

One of the reasons that time to complete was one of, and not “the” performance variable, was due to the possible impact of the speed-accuracy trade-off. As the participants were given control over how long it took them to finish the task, some could move very quickly while sacrificing some accuracy, while others could move at a slower, more calculated pace, focusing on accuracy over speed. Additionally, as the puzzles, by and large, were incomplete in the traditional sense, it was a concern that some participants would be unsure when they were actually complete and “finished” with the task. Although the impact of the trade-off was a concern, correlations to confirm were examined, but no significant results were found. Time to complete is important but

should also be viewed along with the other performance variables, as an overall indication of task performance.

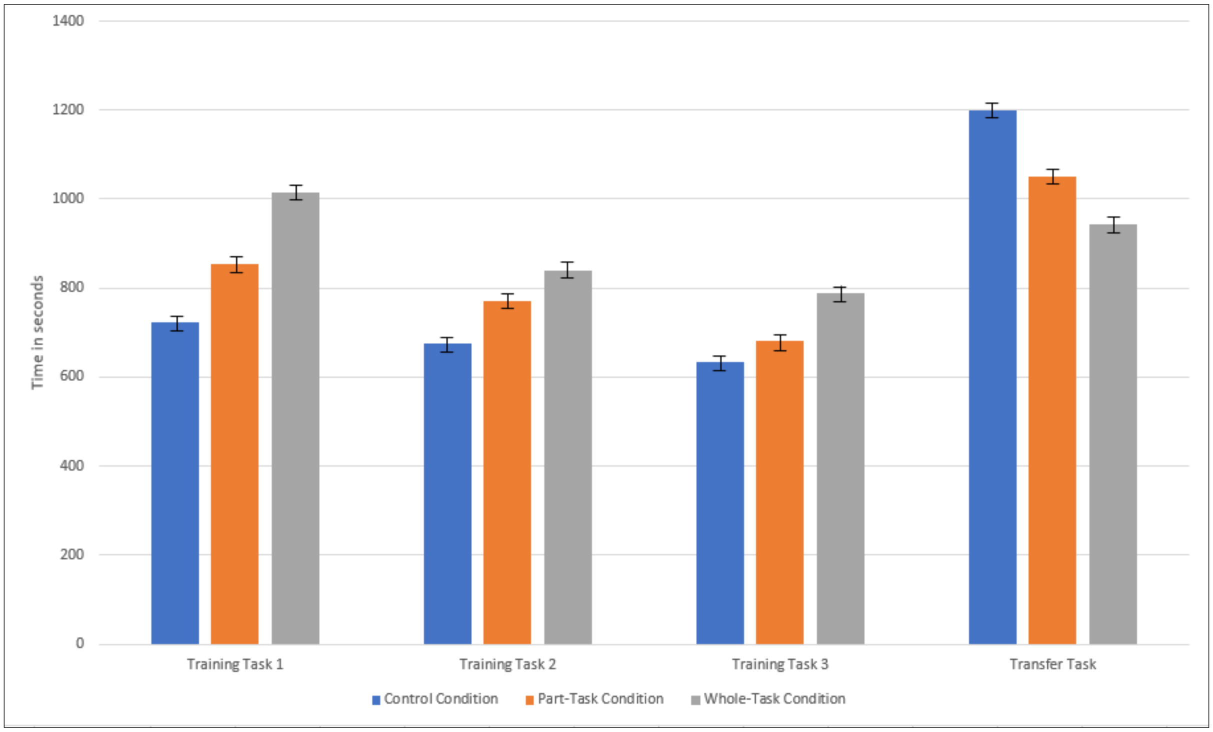


Figure 16. Time to complete, by condition, by task in seconds. Standard error bars shown

Additional analyses looked how the conditions performed on the time to complete performance variable on the proceeding three training tasks, rather than just on the final transfer task (figure 16). This closely mirrors the TLX workload scores, by condition by task show in figure 15, at least for the three training tasks. The control condition participants were able to finish the puzzles in the least amount of time, while encountering no CDO variables. The part-task participants only had to deal with one CDO variable at a time, albeit in high doses, and generally took longer than the control condition, but shorter than the whole-task condition. Lastly, the

whole-task condition participants took the longest on every training task, which could be expected, as they had to deal with three CDO variables in each puzzle (albeit in small doses).

In summary, and based on the transfer task to complete, Hypothesis 1, that there will be a significant difference between the control condition (no intervention), and both experimental conditions (part-task training and whole-task training), was partially supported as only the whole-task condition was significantly different from the control group. Hypothesis 2, that there will be a significant difference in the performance (time) between the part-task training and whole-task training conditions, was not supported by the data, and the two conditions had relatively similar mean completion times. Lastly, Hypothesis 3, that participants in the part-task training condition will outperform (time) participants in the whole-task training condition and the control condition, was unsupported by the results.

Number Correct

The performance variable "Number Correct" was an indicator of accuracy, and recorded how many of the participant's puzzle pieces on the transfer task were in the correct spot. Due to the uncertainty variables, of the thirty-six-piece transfer puzzle, there would be thirty-two correct pieces to place, with four "open" spots. A perfect score was 32 out of 32. As the results indicate, there seems to be a glass ceiling effect for this performance variable, as the mean scores of all three conditions fell between 28 and 29 correct pieces placed. Although seemingly a speed-accuracy trade-off, as mentioned before, no significant correlation was found to support that. It is more likely that participants simply put in enough time and effort into the task to get it so that it was close enough for them to say that they were complete. A typical answer was "I think I'm

done” or “Ok, I don’t think I can complete anymore.” It appears that the part-task and whole-task conditions were exposed to enough uncertainty variables in the training sessions that they were okay with a close, but not fully complete task, knowing that a complete puzzle might not even be possible. Likewise, the control condition, being exposed to all four of these uncertainty variables for the first time ever, was overwhelmed (as is indicated by the TLX scores later), and simply did their best with what they were given.

After reviewing the mean “Number Correct” results across the conditions, there were no significant differences between any of the groups. This was expected to be the case for this performance variable, as participants were asked to complete as much as they could, without a true time limit. Essentially, when measuring time and accuracy, it was expected that the time would be the more varied measure, than the accuracy, at least for this experiment. The part-task condition did record a higher mean score than the other two conditions, although it was not significant, and a future investigation point. There would likely be more of an apparent difference in the number correct measure if the time to complete the task was set, rather than flexible.

In summary, neither Hypothesis 1, 2, or 3 were supported by the number correct data. This is likely caused by the glass ceiling effect, as mentioned in the previous sections. Future investigations should aim to minimize this effect as much as possible in order to identify any intervention main effects.

Variables Identified

The performance variable, Variables Identified (qualitative measure), was a free-response count of how many of the uncertainty variables a participant was able to describe and note after the final transfer puzzle. As stated earlier, there are four uncertainty variables total: missing pieces, extra pieces, false pieces, and cropped images. The cropped image variable was only included in the transfer task. The participants were never told or made aware of what these variables were, rather they were asked to complete the puzzle in front of them to the best of their ability, and to let the experimenter know when they were done, or could not finish anymore. Coding and counting the responses participants gave at the end of the transfer task is one of the better ways to note if they were paying attention, and mentally abstracting the abnormalities they were encountering. Their responses inform questions such as: were they looking for something outside of what they already saw, were they consciously looking and using mental resources to focus on abnormalities, and were they able to identify and write out what it was? This section was scored out of four, and was liberally scored, where credit was given if the participant came close to describing the uncertainty variable. Examples of that include when participants wrote "cut off image" or "can't see the edges of the puzzle," and were essentially noticing that the box / solution image was cropped. A score of four means they noted all variables without ever being explicitly informed about them.

As the results section indicate, the part-task condition was a clear outperformer. Controlling for spatial ability, the control condition's mean score was 2.4, the part-task condition's mean score was 3.4, while the whole-task condition's mean score was 2.6. The part-task condition noted significantly more of the uncertainty variables, than the control group, as well as the whole-task

group. There was no significant difference between the whole-task group and the control group, as their means were very similar.

The proposed reason for this disparity in performance between the part-task condition and the other two is due to having a lower cognitive load while performing the final task. Workload scores will be discussed in the next section in more detail; however, it is important to examine it some with respect to this performance variable. Consider that the control condition was just exposed to any of these four uncertainty variables for the first time during the transfer task. The participants could be best described as confused while performing this task. They started to piece it together, and solve what they could, but describing their thought process as overwhelmed is not inaccurate (as will be shown in the workload section).

However, the whole-task condition had already seen three of these variables before, multiple times. Recall that in each of the three training task puzzles the whole-task condition received, the same three uncertainty variables were in each one (missing, extra, and fake pieces). After seeing those three repetitively, once they saw the pattern again in the transfer puzzle, they had already met the previous parameters to which they were trained. This could even be a case of a cognitive bias called functional fixedness, where one only sees patterns or objects a certain way after a while. Once the whole-task participants saw the three previous variables, they largely "stopped" at those, and did not seek out, further uncertainty variables. This is a quintessential aspect of Salomon and Perkins' (1989) Low- and High-Road Theory on Transfer. As discussed in Chapter 2, the low-road process of transfer involves repetitive, but varied practice to automaticity. There is some variance in the repetitive training to slowly expand the trainee's awareness of the use

case for the skill, and to add flexibility. For example, in this experiment, training the three uncertainty variables repetitively, but on different puzzles. This is essentially an incremental approach to adding flexibility and building broader applicability. It is time consuming, as repetition is required, but it is less useful for higher-order cognitive skills required in contested, degraded, or operationally limited environments. There typically are not any leaps or bounds in applicability, as the trainee focuses on learning the narrow use case of the training task.

In contrast, the high-road process focuses on mindful abstraction, a deliberate and effortful process of grasping core elements of the situation and abstracting them to re-represent them. Once abstracted, the concepts can be re-applied in various training environments outside of the original training environment. For this experiment, by design, the part-task group was “trained” (or exposed) to each uncertainty variable, in high doses. The next puzzle contained another high dose of a completely different uncertainty variable, and so on for the three training tasks. This was theoretically fostering mental abstraction, and allowing the trainee to consider that pieces are not just missing in the first puzzle, and there are not just extra pieces in the second, but rather that these puzzles are all abnormal and it is important to be on the lookout for anything out of the ordinary. This way, without ever being exposed to the fourth variable, participants in the part-task group would already be looking for it, and more keen to recognize it, which the results showed they significantly were better at. This is the crux of the matter at hand for applying this research to broader CDO environments. Instead of trying to train someone on every single thing that could go wrong or be encountered in a CDO environment, could an instructor instead train them on a few highly varied issues, and foster the abstraction that they need to keep up their awareness of abnormalities. The part-task training schedule appears to foster abstraction, as the

participants have extra germane load to apply to abstracting core concepts. It appears that the results from this experiment, support part-task training for the CDO environment, at least partially, within the constraints of this effort.

In summary, Hypothesis 1, that there will be a significant difference between the control condition (no intervention), and both experimental conditions (part-task training and whole-task training), was partially supported as only the part-task condition was significantly different from the control group (the whole-task condition was not). Hypothesis 2, that there will be a significant difference in the performance (variables noted) between the part-task training and whole-task training conditions, was supported by the data, with the part-task condition outperforming the whole-task condition. Lastly, Hypothesis 3, that participants in the part-task training condition will outperform (variables noted) participants in the whole-task training condition and the control condition, was supported by the data. In terms of variables identified, not only did the part-task condition outperform the control and whole-task condition, they also noted more instances of the fourth novel variable (cropped image) than the other two conditions.

NASA TLX Workload Scores

The last, yet important, performance variable is the participant's NASA TLX overall workload scores. These self-reports were completed at the end of each puzzle, and provide insight and measurement in to each participant's perceived workload while completing the task. Recall, that TLX scores are scored out of 100, with a lower number indicating lower workload. Workload is important as mindful abstraction, the key to high-road transfer, is only possible if the participant

has the cognitive resources (germane load) available to allocate to it, as it is an effortful and conscious process.

Participants were encouraged to complete the puzzles as quickly and as accurately as possible, which likely placed additional pressure on them while completing a somewhat complex and increasingly disorganized transfer puzzle. The results section presented the perceived overall workload scores by condition for the final transfer task, as well as the training task workload scores. The workload scores by condition across the training and transfer task are more informative than the raw scores themselves, and allow for context and comparison.

The workload scores on the transfer task alone revealed that the part-task training condition participants indicated a significantly lower workload score than both of the other training conditions. The part-task condition's mean workload score was a 32.9 out of 100, while the whole-task condition's mean workload score was a 45.2, and the control condition was a 47.4. These scores indicate that for the complex, and somewhat disorganized transfer task puzzle, that included four uncertainty variables, the part-task condition encountered the least cognitive load while completing the task. The results suggest that this was due to the training schedule, the part-task nature of their exposure to the variables, in the training tasks leading up to the transfer task. This coincides with the fractionation research discussed in Chapter 2 regarding dividing a task into sub-tasks that are typically done simultaneously. Each sub-task is trained in isolation, and then combined at the end so that the trainee is doing them all at once (Wightman and Sistrunk, 1987). By exposing trainees to one variable at a time during training, when they were all put together in the transfer task, there was less overall cognitive load imposed on them, even when

presented with a novel uncertainty variable. The workload results on the transfer task indicate that compared to the whole-task and control conditions, the part-task condition participants had extra cognitive load to "spend" on the novel variables.

Another interesting finding is what the TLX scores look like across the training tasks and the transfer task. From the above discussion, the part-task condition had significantly lower workload on the final task, but how did those participants fare during the fractionated training tasks? These reported scores are shown in Table 16 in Chapter 4 and re-represented in Figure 17 below, and indicate fascinating patterns.

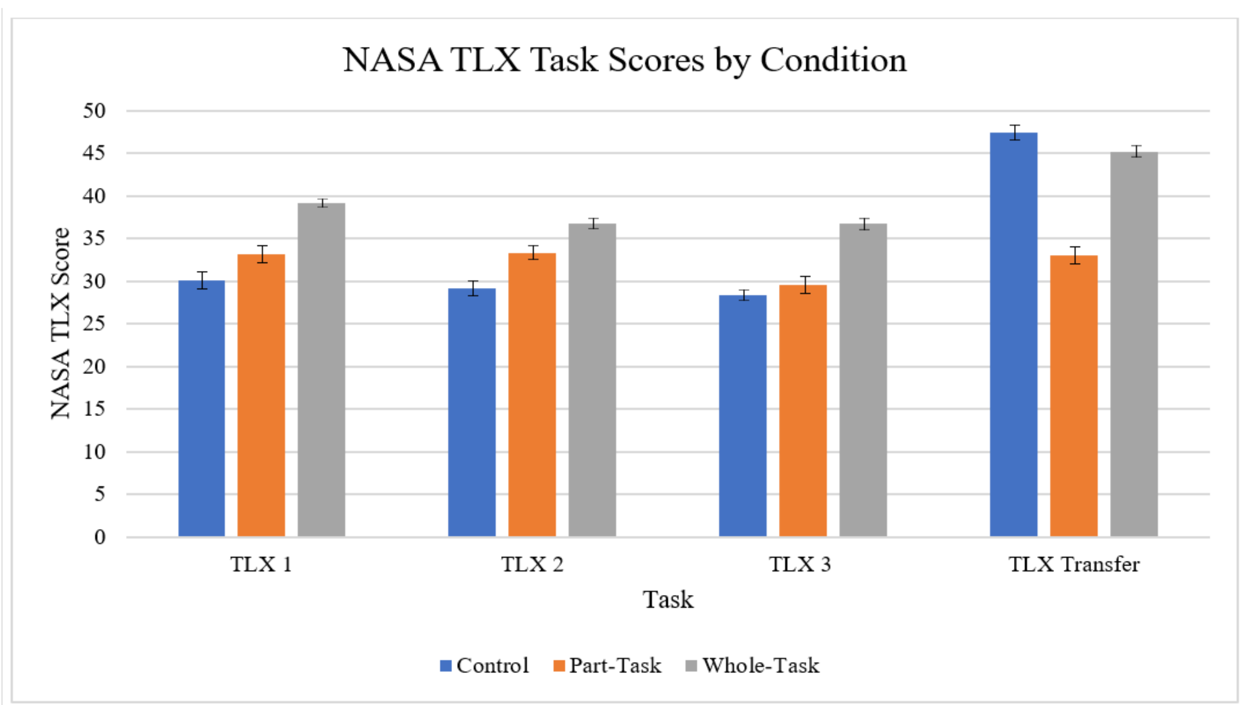


Figure 17. NASA TLX Scores by Condition for All Tasks - Patterns

As shown, there is a predictable pattern for the first three training tasks. The control condition participants reported the lowest overall workload scores, and is somewhat predictable, as they were solving puzzles that were not manipulated in any way. This trend, of the control condition reporting the lowest workload scores, continued across all three training tasks. The next predictable pattern, is the part-task condition scores. Recall that these participants were given a puzzle with only one (and unique) uncertainty variable integrated with it at a time. For example, task one would be a puzzle with six missing pieces only, task two would be a puzzle with six extra pieces only, and task three would be a puzzle with six extra or spoofed pieces replacing the real pieces (in practice, the order of the uncertainty variables and puzzles was randomized however). So although more challenging and workload inducing than a control, unmanipulated puzzle, the participants did have a more difficult task to solve these, as indicated by the reported workload scores. This remained true across all three transfer tasks. Finally, the whole-task condition reported the highest workload scores across all three training conditions. This also is somewhat predictable, as this is the largest difficulty change from the unmanipulated control task puzzle. For each training task, the whole-task participants were exposed to three uncertainty variables at once. So for each puzzle, there were two missing pieces, two duplicate pieces, and two fake pieces. Although the puzzle image itself changed for each task, the uncertainty variables remained the same for all three tasks.

The workload trend that held throughout all three training tasks was that the control group reported the least workload, then the part-task group, and then the whole-task group reported the most workload. Yet, that pattern disappeared on the final transfer task. Both the control condition and the whole-task condition's reported workload scores markedly rose, while the part-task

condition scores stayed pretty much the same. The control group jump in workload scores is predictable, as they went from solving a normal puzzle repetitively, to a chaotic and difficult one. The whole-task condition's scores rose too, although not as drastic as the control groups, but certainly higher than any of the training task scores. It was a more difficult task, and two of the three conditions demonstrated that in their workload scores, except for the part-task condition group. The part-task condition's scores stayed almost the same as they were for all the previous training tasks. There was no marked rise when the difficulty rose, and there was a significant difference in their workload scores, versus the other two conditions. The caveat is that the TLX scores are from a subjective measurement, so conclusions are hard to draw only from it, but it can be stated that there is more opportunity for mindful abstraction if there are more cognitive resources available. Additionally, it can be noted that the part-task condition found the most difficult puzzle, to be the "easiest" in terms of workload. This certainly merits further investigation.

Theoretical and Practical Implications

The literature is all but settled on when part-task training excels for closed-skill or procedural training, although a noted gap exists when referring to open-skill or cognitive tasks, such as mindfully abstracting concepts. This work addresses that question, in part. The results from the study indicate that the part-task training (or exposure) schedule did have statistically significant benefits, in at least two of the performance variables. Additionally, when speaking of the benefits, the only real comparisons should be between the two experimental training schedules: part-task and whole-task. The control condition was more or less expected to fail, or do poorly on most measures, although it served as a useful baseline from which to explore the other

conditions. So in terms of variables noted, if the participants understood all the uncertainty variables they were encountering and listed them in their minds, the part-task condition participants performed better. Likewise, in terms of workload scores, the part-task condition participants held impressively steady throughout all three training tasks and the transfer task. There was no marked increase in workload in their reports for the transfer task, as there were for the whole-task and control condition. It is likely that this is due to the way the uncertainty variables were presented to them, and the resultant cognitive load, or lack thereof, imposed on them during the tasks.

Recall, the Naylor/Briggs continuum suggested that a task consisting of low organization (such as a puzzle with various uncertainty variables integrated in it) may fare better being trained in a part-task method. This is due to the inability to easily “link” the disparate sections together, as one would in a normal environment. When the participant is able to see the entire task for what it is, and able to follow logical steps to complete it, whole-task training is suggested, even if the task is difficult. The “flow” from one task step to the next builds on each other and lessens the cognitive load as each new piece of information logical fits in the proper location. Yet, the less the task flows and is linked together, the more mental processing it will require, and cognitive load will increase, unless a new training schedule is chosen to reduce this. That is where the benefit of part-task training appears. The fractionation of the different uncertainty variables in this experiment seems to have enabled the participants to focus on just that one variable, figure it out, notice what is happening, and verify that that was all that was going on in that task. When there is only one uncertainty variable to consider, even if it is exposed in higher doses, it frees up more cognitive capacity to do other things, such as wonder why this puzzle is missing pieces,

and consider abstracting that concept and saving it for later. After the second training task is presented to them, the gears should start turning and integrating what they learned from the first task, “there are pieces missing” into a more abstract concept of “there are going to be things wrong in each puzzle.” So, when the transfer puzzle appears, they are already expecting something else, and probably even looking for it. The fact that they also find previous uncertainty variables did not seem to impact workload at all. As discussed throughout this dissertation, mindful abstraction is an effortful process, the trainee needs to focus on it and be aware of it. That is more difficult, the more overloaded cognitively someone is.

The previous section, and the results indicate that there is something to high-road transfer, and using part-task training methods to optimize it. Unfortunately, it is hard to say theoretically why the time to complete scores between whole-task and part-task conditions were similar. It is possible that this was also due to the ceiling effect of the number correct (or accuracy) scores, and that the tradeoff relationship between time and accuracy accounts for it. However, statistically there was no difference between the part-task and whole-task times, or their number correct scores.

Although the experiment was a very basic, laboratory version of a much more complex real-world task (i.e., monitoring a sophisticated sensor display) there are promising practical implications that can be forwarded with additional future research. The lab results indicate that part-task training can be better in this narrow circumstance, and the suggestion for how to best train for CDO-type environments, judging by this data set, is to use a part-task training schedule. There will be no significant accuracy or time gains or losses, but there will be anticipated

abstraction gains, as well as lessened workload. Although the ceiling effect limited some gained possible knowledge from this task, the workload and variables noted indicated gains could be made if similar tasks were trained via part-task exposure.

Limitations and Future Research

As with most studies, this experiment was constrained and limited by various factors. A key aspect of furthering research in a particular field is presenting methods and results clearly, while also being transparent about limitations and shortcomings. Only after that, can others place the work in the appropriate context, and use it for future research.

One of the first limitations worth noting is with the selection of the puzzle solving task itself. Although initially assumed to be a skill that the general population is familiar with and would require no pre-training with, it was found that there was more variation in puzzle solving skill than was optimal. Although individual performance varying by ability was expected, a future iteration of this study would benefit from a task with a stronger baseline. For example, a similar future investigation into this type of problem would benefit from a task that is either trained during the study, or one that is even more commonly known than puzzle solving. Anecdotally, it may be a generational difference as quite a few thirty-year-old pilot participants thought the thirty-six-piece puzzle was significantly too easy. Yet, 18-to-20-year old pilot participants seemed to struggle more than the older participants.

Additionally, the constraint that the training tasks and transfer tasks were within the same sessions could be a limitation. Additional work can investigate similar training schedules that separate training tasks from transfer tasks by more time.

Another limitation and future opportunity would be dealing with the possible ceiling effect with respect to the number correct performance variable. The nature of a puzzle is likely the cause of this, and them having the solution or box image available. A solution would be to impose a strict time limit on each of the task and then assess accuracy and completeness at the cutoff point.

Although fairly complex, the puzzle size limitations may have limited some variability as well. Initial puzzle sizes were pared down from 64 pieces, to 49, and then finally to 36 pieces. Unfortunately, during pilot testing, larger puzzle sizes were taking participants thirty minutes or more per task and were also visibly frustrating some participants. Increasing the puzzle size in the future would be beneficial if time and resource constraints allow for it.

A final limitation was that this was a very specific task looking to substitute for more applied tasks in the future, essentially an experiment at the basic level, replicating only the higher-level concepts of the applied task, not the task itself. Manipulating puzzles looking for effects is a cost-efficient and low-tech solution that can provide insights and answer specific research questions, but puzzle solving may not be close enough to the applied tasks to draw real conclusions extending to personnel in CDO environments (such as Naval personnel training on sensors and other situational awareness technologies). This experiment was a stepping stone in the direction to making research-based training recommendations for future training programs

and schedules. Future work can refine and increase the fidelity of the tasks to better suit the applied task that is in question.

The results, both significant and otherwise, from this dissertation should assist and encourage future research into how to best facilitate mindful abstraction during training for uncertain training environments. Transfer of training is a huge domain, so adding another piece of insight into a subsection of it, such as determining the optimal training schedules (part-task or whole-task) for this type of training is beneficial and warranted.

APPENDIX A: DEMOGRAPHICS

Age ______

Gender M / F

Do you have normal or corrected vision? Yes / No

Are you colorblind? Yes / No

Approximate GPA ______

How many puzzles have you completed in the past year? Please circle your response

None

Between one and three

Three or more

How would you rate your level of experience solving puzzles? Please circle your response

Beginner (no real experience with them)

Intermediate (solved puzzles in the past and understand the basic concepts)

Advanced (solved many puzzles, know the basic & expert concepts, consider it a hobby)

How would you rate your confidence in your ability to solve an intermediate level puzzle (approximately 60 pieces)? Please circle your response

Not at all confident

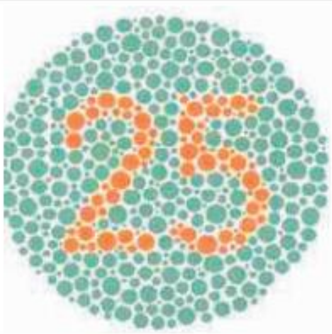
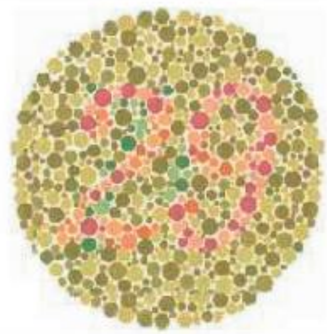
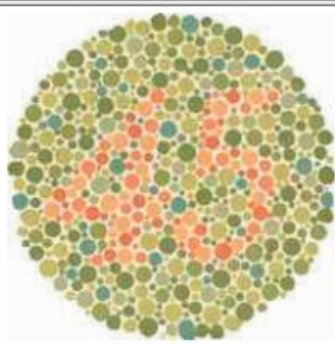
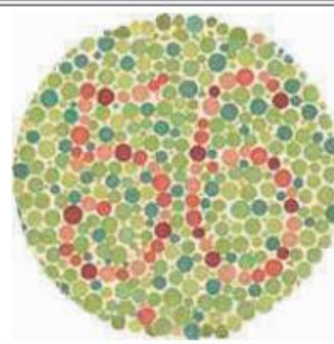
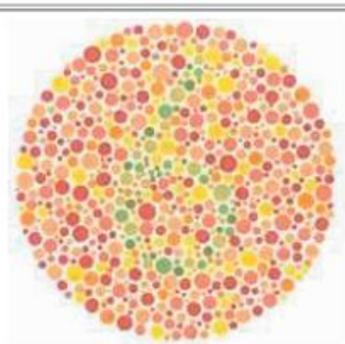
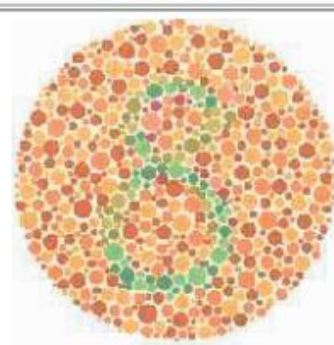
Somewhat confident

Confident

APPENDIX B: ISHIHARA COLORBLIND TEST

Ishihara Test for Color Blindness

What numbers do you see revealed in the patterns of dots below?

APPENDIX C: WONDERLIC PERSONNEL TEST SAMPLE

Wonderlic Personnel Test Sample Questions

1. Which of the following is the earliest date?

   A) Jan. 16, 1898   B) Feb. 21, 1889   C) Feb. 2, 1898   D) Jan. 7, 1898   E) Jan. 30, 1889

2. LOW is to HIGH as EASY is to _?_.

   J) SUCCESSFUL   K) PURE   L) TALL   M) INTERESTING   N) DIFFICULT

3. A featured product from an Internet retailer generated 27, 99, 80, 115 and 213 orders over a 5-hour period. Which graph below best represents this trend?

A   B   C   D   E

4. What is the next number in the series?   29   41   53   65   77   _?_

   J) 75   K) 88   L) 89   M) 98   N) 99

5. *One word below appears in color. What is the OPPOSITE of that word?*
   She gave a complex answer to the question and we all agreed with her.

   A) long   B) better   C) simple   D) wrong   E) kind

6. Jose's monthly parking fee for April was $150; for May it was $10 more than April; and for June $40 more than May. His average monthly parking fee was _?_ for these 3 months.

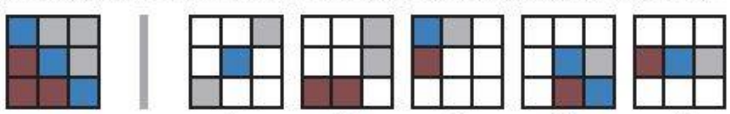
   J) $66   K) $160   L) $166   M) $170   N) $200

7. *If the first two statements are true, is the final statement true?*
   Sandra is responsible for ordering all office supplies.
   Notebooks are office supplies.

   Sandra is responsible for ordering notebooks.

   A) yes   B) no   C) uncertain

8. Which THREE choices are needed to create the figure on the left? Only pieces of the same color may overlap.



J   K   L   M   N

9. Which THREE of the following words have similar meanings?

   A) observable   B) manifest   C) hypothetical   D) indefinite   E) theoretical

10. Last year, 12 out of 600 employees at a service organization were rewarded for their excellence in customer service, which was _?_ of the employees.

    J) 1%   K) 2%   L) 3%   M) 4%   N) 6%

Correct Answers: 1. E, 2. N, 3. D, 4. L, 5. C, 6. M, 7. A, 8. KLM, 9. CDE, 10. K

APPENDIX D: NASA TASK LOAD INDEX

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?

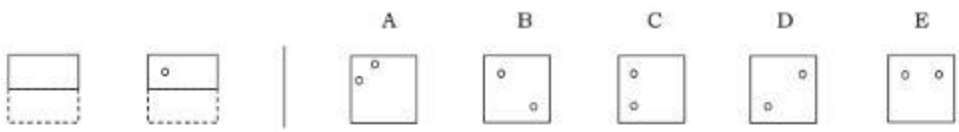
Very Low — Very High

APPENDIX E: PAPER FOLDING TEST

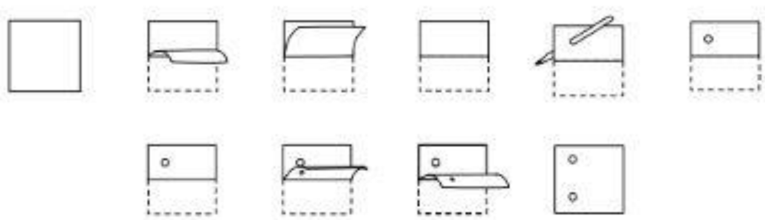
Paper Folding Test—Vz-2-BRACE

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.



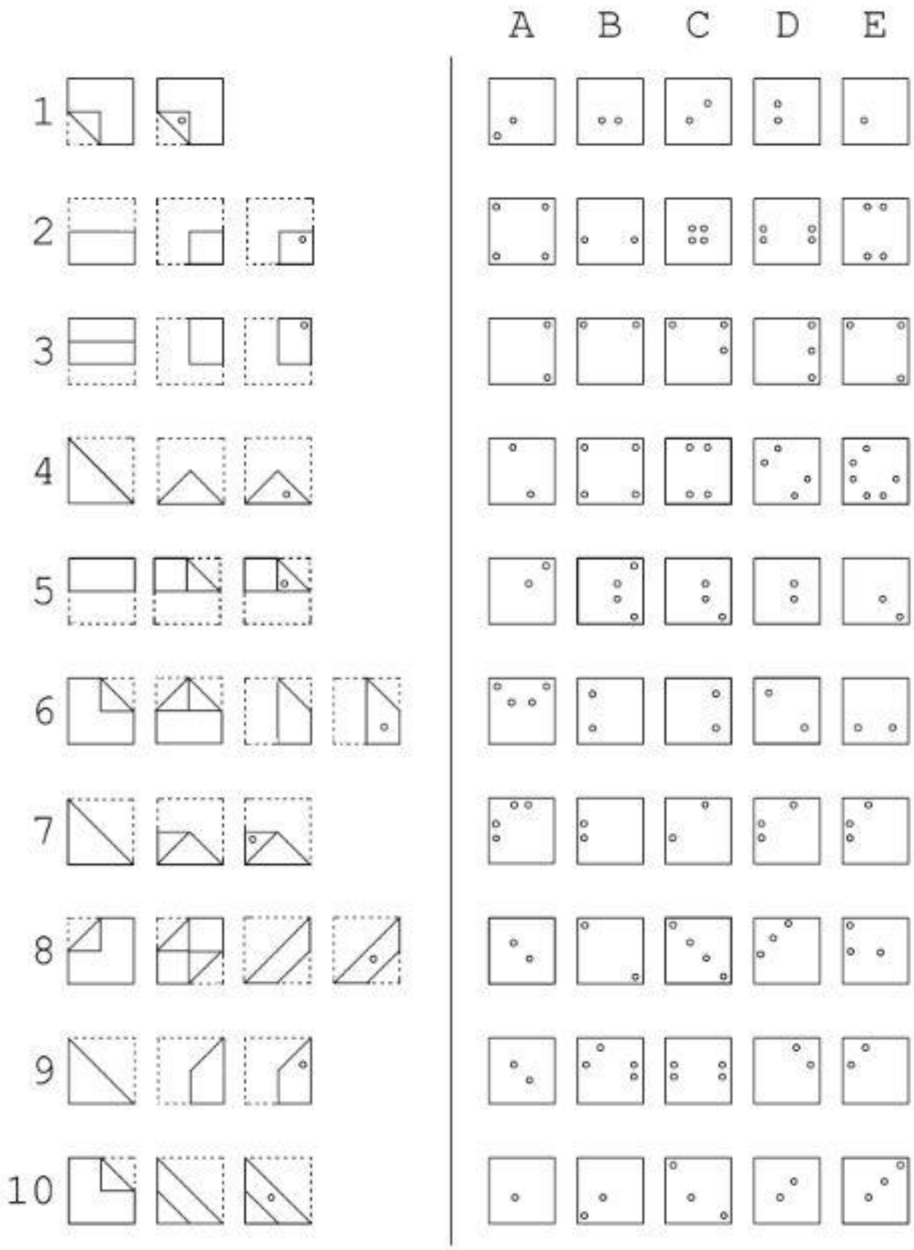
In these problems all of the folds that are made are shown in the figures at the left of the line, and the paper is not turned or moved in any way except to make the folds shown in the figures. Remember, the answer is the figure that shows the positions of the holes when the paper is completely unfolded.

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

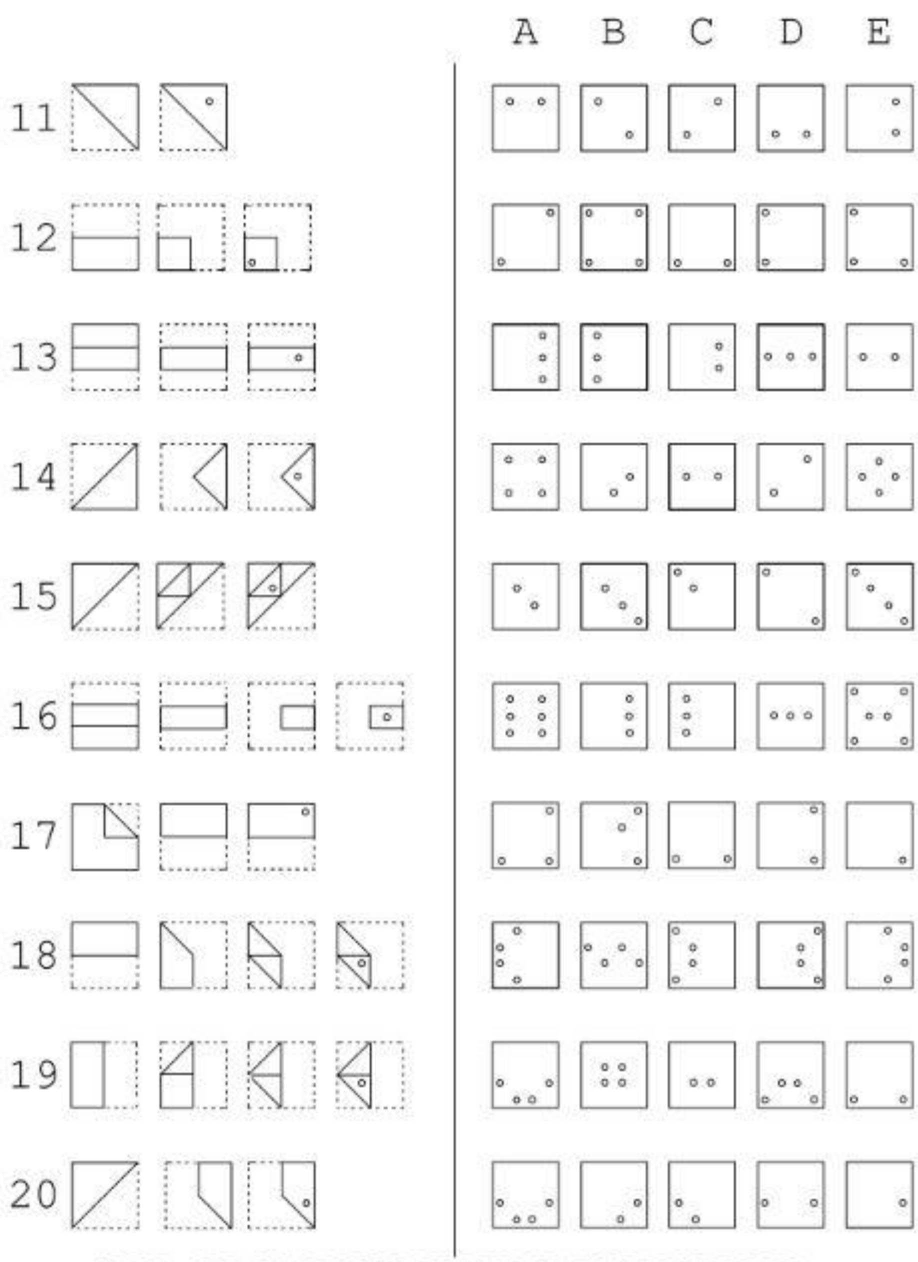
PART ONE (3 MINUTES)



STOP

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

PART TWO (3 MINUTES)



STOP AND WAIT FOR FURTHER INSTRUCTIONS
DO NOT GO BACK TO PART ONE

APPENDIX F: POST-PARTICIPATION SURVEY

Post-Participation Survey

Please check the box next to any puzzle-solving strategies used during these tasks (check as many as apply)

☐ Separate the edges from the internal pieces

☐ Work on the edges and corners first

☐ Place pieces in approximate locations remembered from box cover image

☐ Sort pieces by color and solve "mini-puzzles"

☐ Sort pieces by identifying markings other than color and solve those "mini-puzzles"

☐ Trial and error: attempting to solve through puzzle shape and fit with other pieces

☐ No discernable strategy employed. You just "solved it"

Please list any and all puzzle abnormalities encountered while solving

1. ____________________
2. ____________________
3. ____________________
4. ____________________
5. ____________________

APPENDIX G: UCF IRB APPROVAL 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: **John P. Killilea**

Date: **August 14, 2017**

Dear Researcher:

On 08/14/2017 the IRB approved the following human participant research until 08/13/2018 inclusive:

Type of Review:	UCF Initial Review Submission Form Expedited Review
Project Title:	Investigating the effectiveness of using part-task or whole-task training for uncertain training tasks
Investigator:	John P. Killilea
IRB Number:	SBE-17-13277
Funding Agency:	
Grant Title:	
Research ID:	N/A

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 08/13/2018, 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).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

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:

Page 1 of 2

Signature applied by Gillian Amy Mary Morien on 08/14/2017 09:17:08 AM EDT

IRB Coordinator

Page 2 of 2

LIST OF REFERENCES

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

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baddeley, A. (1986). Oxford psychology series, No. 11. Working memory.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology, 41*(1), 63-105.

Baldwin, T. T., Ford, J. K., and Blume, B. D. (2009). Transfer of training 1988-2008: An updated review and new agenda for future research. In G. P. Hodgkinson & J. K. Ford (Eds), *International review of industrial and organizational psychology*, 24, 41-70. Chichester, UK: Wiley.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, *36*(4), 1065-1105.

Boster, J. S., & Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgements of similarity among fish. *American Anthropologist*, *91*(4), 866-889.

Briggs, G. E., & Naylor, J. C. (1962). The relative efficiency of several training methods as a function of transfer task complexity. *Journal of Experimental Psychology*, *64*(5), 505.

Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human resource development review*, *6*(3), 263-296.

Byrnes, J. P. (1996). *Cognitive development and learning in instructional contexts*. Boston: Allyn and Bacon.

Cain, B. (2007). *A review of the mental workload literature*. Defence Research And Development Toronto (Canada).

Clarke, N. (2002). Job/work environment factors influencing training transfer within a human service agency: Some indicative support for Baldwin and Ford's transfer climate construct. *International journal of training and development*, *6*(3), 146-162.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillside. *NJ: Lawrence Earlbaum Associates*.

Cormier, S., & Hagman, J. (1987) *Transfer of learning: Contemporary research and applications.* San Diego: Academic Press.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*(5), 795.

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, *77*(5), 615.

Dykens, E. M. (2002). Are jigsaw puzzle skills spared in persons with Prader-Willi syndrome? *Journal of Child Psychology and Psychiatry*, *43*(3), 343-352.

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

Ellis, H. C. (1965). The transfer of learning.

Facteau, J. D., Dobbins, G. H., Russell, J. E., Ladd, R. T., & Kudisch, J. D. (1995). The influence of general perceptions of the training environment on pretraining motivation and perceived training transfer. *Journal of management*, *21*(1), 1-25.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected*(No. IDA-D-3061). INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA.

Ford, J. K., Quiñones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. *Personnel psychology*, *45*(3), 511-527.

Ford, J. K., & Weissbein, D. A. (1997). Transfer of training: An updated review and analysis. *Performance Improvement Quarterly, 10*(2), 22-41.

Foxon, M. (1997). The influence of motivation to transfer, action planning, and manager support on the transfer process. *Performance Improvement Quarterly*, *10*(2), 42-63.

Friedenberg, J., & Silverman, G. (2011). *Cognitive science: An introduction to the study of mind*. Sage.

Gagne, R. M., Baker, K. E., & Foster, H. (1950). On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychological Review*, *57*(2), 67.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12(3*), 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.

Green, B. F., McCloskey, M., & Caramazza, A. (1982). The relation of knowledge to problem solving with examples from kinematics. *Thinking and Learning: Research and Open Questions*, *2*, 127-139.

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.

Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *The Quarterly Journal of Experimental Psychology*, *28*(4), 603-621.

Holding, D. H. (1965) *Principles of training.* Oxford: Pergamon.

Huczynski, A. A., & Lewis, J. W. (1980). An empirical study into the learning transfer process in management training. *Journal of management studies*,*17*(2), 227-240.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kirkpatrick, D. L. (1967). Evaluation of training.

Knerr, C., Morrison, J., Mumaw, R., Stein, D., Sticha, P., Hoffman, R., Buede, D., & Holding, D. (1985). *Simulator-based research in part-task training*. Alexandria, VA: Human Resources Research Organization.

Laker, D. R. (1990). Dual dimensionality of training transfer. *Human Resource Development Quarterly*, *1*(3), 209-223.

Langer, E. (1989) *Mindfulness.* Reading, MA: Addison-Wesley.

Larken, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-1345.

Lim, D. H., & Johnson, S. D. (2002). Trainee perceptions of factors that influence learning transfer. *International journal of training and development*,*6*(1), 36-48.

Lim, J., Reiser, R. A., & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of a complex cognitive skill. *Educational Technology Research and Development*, *57*(1), 61-77.

Macaulay, C., & Cree, V. E. (2000). Transfer of learning. *Transfer of Learning in Professional and Vocational Education: Handbook for Social Work Trainers,* 1-26.

Maddox, M. (1987). Environment and supervisory influences on the transfer of training. *Unpublished manuscript, Michigan State University. E. Lansing, MI*.

McGehee, W., & Thayer, P. W. (1961). Training in business and industry.

Moore, S. C., Stockfisch, J. A., Goldberg, M. S., Holroyd, S. M., & Hildebrandt, G. G. (1991). Measuring military readiness and sustainability. *Santa Monica, Calif.: RAND.*

Morrison, J. E., & Fletcher, J. D. (2002). Cognitive readiness (No. IDA-P-3735). INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA.

Navy Warfare Development Command Public Affairs. (2012, June 11). Navy Warfare Development Command Releases Tactical Memorandum. *America's Navy*, p. 1.

Naylor, J. C., & Briggs, G. E. (1963). Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *Journal of Experimental Psychology, 65*(3), 217.

Newell, K., Carlton, M., Fisher, A., & Rutter, B. (1989). Whole-part training strategies for learning the response dynamics of microprocessor driven simulators. Acta Psychologica, 71, 197-216.

Newstrom, J. W. (1984). A role-taker/time-differentiated integration of transfer strategies. *Vol. August*.

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.

Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New ideas in psychology, 2*(2), 137-168.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International encyclopedia of education*, *2*, 6452-6457.

Prudente, T. (2015, November 2). Naval Academy reinstates celestial navigation. *Military Times*.

Salas, E., Milham, L. M., & Bowers, C. A. (2003). Training evaluation in the military: Misconceptions, opportunities, and challenges. *Military Psychology,15*(1), 3.

Salomon, G., & Perkins, D. N. (1987). Transfer of cognitive skills from programming: When and how?. *Journal of educational computing research,3* (2), 149-169.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational psychologist*, *24*(2), 113-142.

Singley, M., & Anderson, J. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Spector, J. M., & Anderson, T. M. (2000). *Holistic and integrated perspectives on learning, technology, and instruction: Understanding complexity*. Mahwah, NJ: Lawrence Erlbaum.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science, 12*(2), 257-285.

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.

Taylor, P. J., Russ-Eft, D. F., & Taylor, H. (2009). Transfer of management training from alternative perspectives. Journal of Applied Psychology, 94(1), 104.

Teague, R. C., Gittelman, S. S., & Park, O. C. (1994). *A review of the literature on part-task and whole-task training and context dependency* (No. ARI-TR-1010). ARMY RESEARCH INST FOR THE BEHAVIORAL AND SOCIAL SCIENCES ALEXANDRIA VA.

Thompson, M. M., & McCreary, D. R. (2006). *Enhancing mental readiness in military personnel*. DEFENCE RESEARCH AND DEVELOPMENT TORONTO (CANADA).

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. II. The estimation of magnitudes. *Psychological Review, 8*(4), 384.

Underwood, B. J. (1951). Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. *Journal of Experimental Psychology*, *42*(1), 44.

van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learners' mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5–13.

van Merriënboer, J. J., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied cognitive psychology*, *20*(3), 343-352.

Verdine, B. N., Troseth, G. L., Hodapp, R. M., & Dykens, E. M. (2008). Strategies and correlates of jigsaw puzzle and visuospatial performance by persons with Prader-Willi syndrome. *American Journal on Mental Retardation*, *113*(5), 343-355.

Wickens, C. D. (1984). Engineering psychology and human performance. New York, NY: Harper Collins.

Wickens, C. D. (2002). *Multiple resources and performance prediction. Theoretical issues in ergonomics science*, *3*(2), 159-177.

Wightman, D. C., & Sistrunk, F. (1987). Part-task training strategies in simulated carrier landing final-approach training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(3), 245-254.

Yelon, S. L., & Ford, J. K. (1999). Pursuing a multidimensional view of transfer. *Performance Improvement Quarterly*, *12*(3), 58-78.

Yelon, S., Sheppard, L., Sleight, D., & Ford, J. K. (2004). Intention to transfer: how do autonomous professionals become motivated to use new ideas? *Performance Improvement Quarterly*, *17*(2), 82-103.