EXAMINING THE EFFICACY OF COMPUTER-BASED VISUAL TRAINING TO IMPROVE THE SPEED AND ACCURACY OF WEAPON ACQUISITION IN A DYNAMIC USE OF FORCE SCENARIO

by

M. Hunter Martaindale, B.G.S, M.S.C.J.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Criminal Justice
August 2016

Committee Members:

J. Pete Blair, Chair

Scott Wm. Bowman

Marcus Felson

Beth A. Sanders

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

DEDICATION

This work is dedicated to Harrison; my canine, my compadre, and my non-human better half.

ACKNOWLEDGEMENTS

Many people provided assistance and/or support throughout both the process of this project and the educational journey to reach this point. My parents have always supported and challenged me academically…for this I am grateful. While my committee didn't raise me from an infant, I am also grateful for everything they assisted with along the way. Pete Blair, Beth Sanders, Scott Bowman, and Marcus Felson provided encouragement, advice, and honest feedback throughout the process. Thank you all for everything. Lastly, I want to single out the people that provided direct assistance with creating the images and video for this project's intervention. Jesse from Milo Range, Billy, Kyle, David, Joseph, and Megan with an H, thank you for all your help.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

Law enforcement officers are frequently required to make use of force decisions. When a weapon is present, a key consideration to a successful use of force decision is how quickly an officer can recognize the presence of a weapon. The current dissertation tests the efficacy of a newly developed vision-training program to improve officers' ability to detect the presence of a gun faster and more accurately by employing the key principles of deliberate practice. A 1x2 independent groups design with random assignment to conditions was utilized to assess the effectiveness of the training program. Bayesian data analysis techniques were applied to the collected vision-based data. It was found that deliberately practicing improved participants' ability to visually locate the item. Furthermore, participants receiving the test intervention made fewer errors than participants that received the control intervention. The future direction of the research is also explored.

I. INTRODUCTION

On September 4, 2014, Levar Edward Jones pulled his SUV into a Columbia, South Carolina gas station after being signaled by a state trooper. Trooper Sean M. Groubert initiated the traffic stop due to a seat belt violation. Dash-cam footage showed Jones stepping out of his vehicle as Trooper Groubert stopped his patrol car. Trooper Groubert asked Jones to produce his driver's license. At this point, Jones abruptly turned and reached into his vehicle to retrieve his wallet. Trooper Groubert repeatedly yelled at Jones to "Get out of the car". The video showed Jones quickly turn back toward the officer. Trooper Groubert fired four shots. Fortunately Jones survived being struck in the hip by one bullet (BBC, 2014).

Trooper Groubert was arrested and lost his job. Groubert was charged with assault and battery of a high and aggravated nature. If convicted, Groubert could face up to 20 years in prison. Groubert stated that he saw Jones lunge into the vehicle and then make a quick movement back toward the trooper. It appeared Groubert did not make an effort to visually confirm if Jones possessed a weapon; rather, Groubert took the physical movement toward the vehicle as an indication that he had to be grabbing a weapon (BBC, 2014).

This case illustrates the complex use of force decision-making process law enforcement officers must undergo on a daily basis. Recent use of force cases (e.g., Levar Jones; Michael Brown in Ferguson, MO; or Walter Scott in North Charleston, SC) have highlighted the issue of police use of force, training, and decision-making. Any event resulting in unarmed civilians being shot receives much media attention. This attention can skew the general public's perception regarding the frequency with which police use force—specifically the frequency with which officers fire at suspects (see for example, availability heuristic, Tversky & Kahneman, 1974). However, it is well known among researchers and practitioners that police officers rarely use force when interacting with the public.

The Bureau of Justice Statistics conducted the 2008 Police-Public Contact Survey as a supplement to the National Crime Victimization Survey. An estimated 1.4% of those surveyed had force used or threatened during their most recent contact with law enforcement (Eith & Durose, 2011). In a related study, Hickman, Piquero, and Garner (2008) found that 1.5% of police-citizen contacts resulted in force or the threat of force being administered. No matter how rarely force is administered, law enforcement officers have a responsibility to correctly apply force. For this reason, it is imperative that officers receive adequate training. The concept of providing officers with additional training beyond the training academy is not novel. However, it is the position of the author multiple disciplines should be utilized to develop an effective training program that can improve use of force training program. As such, the intention of this study is to improve law enforcement vision training as it relates to law enforcement use of force through the development of an innovative training program.

Present Study

Scenario-based use of force training is a valuable tool to prepare officers for real-life encounters, such as the shooting of Levar Jones. However scenario-based use of force training requires a large time and resource commitment for law enforcement agencies (Murray, 2004). This is in part because the agency must maintain adequate staffing levels to continue to serve the community. Law enforcement agencies' commitment to protect the community can result in training officers on their days off or reducing the number of officers on duty. Both of these options can prove costly to the department or the community. Additionally, the department must have role players and a location to perform the scenario-based training. The logistical complexity of performing scenario-based training for all officers limits the amount of training received. This means officers may not receive scenario-based training very often.

A key aspect of scenario-based training is one's ability to visually identify actionable threats (e.g., the presence of a weapon). When an officer only receives intermittent training, this valuable visual skill diminishes. Ericsson, Krampe, and Tesch-Romer (1993) posit that expertise at a task is only attained through continued, deliberate practice of the task. For this reason, the present study proposes a method to continually train officers' visual system to recognize lethal threats without requiring departments to expend valuable resources (e.g., officer time, funding). It is proposed that continued visual training will result in faster, more accurate use of force decisions.

Specifically, this dissertation studies a supplemental form of computer-based use of force training that targets training officers to visually acquire the presence of a gun faster. The training would not replace live, scenario-based training; rather, it would give officers continued visual training to improve their ability to detect a gun in a dynamic situation. As officers have limited time for training, the program is designed to be performed in only a few minutes. This will allow for officers to complete a few minutes of training before starting his or her shift. As the program is new, the proposed dissertation is a proof of concept and initial test of the efficacy of visual search training on lethal use of force encounters. It requires years of research and development to produce a validated program. The current dissertation is merely the first step in the process. As such, this dissertation will seek to answer three research questions.

Research Questions

The overarching goal of the training program is to improve the accuracy and speed of decisions in a use of force situation by training participants how to visually search an environment. For this reason, there are three primary research questions.

1. Will a computer-based, vision-training program improve the accuracy of weapon identification in dynamic situation?

2. Will a computer-based, vision-training program improve the speed of weapon identification in a dynamic situation?

3. Will a computer-based, vision-training program improve the visual search strategy of participants in a dynamic situation?

All three of these broad research questions require an in-depth examination of literature outside of the traditional criminal justice literature. The next chapter will present the relevant literature to lay a foundation for the study.

II. RELEVANT LITERATURE

An important component of routine visual behavior is the ability to find one item in a world filled with other, distracting items—e.g., locating a friend in a crowd or your favorite cereal at the grocery store. This is especially true for law enforcement officers when a weapon is present in the environment. The presence of a weapon escalates the manner in which an officer can, and should, respond. However, while lethal law enforcement response to a weapon is justified in some cases, mistaken weapon identification can lead to the application of lethal force when it is unjustified (for example, 1999 shooting of Amadou Diallo by NYPD officers; Fritsch, 2000). The present study attempts to provide visual training to improve law enforcement officers' ability to quickly and correctly detect the presence of a weapon.

This study topic draws on various disciplines to develop a theoretical foundation. As such, the current chapter is broken into four major sections. These sections include:

I. Relevant use of force literature. This includes definitions of force, case law, correlates of force, and the force continuum. Each component of use of force will be discussed in detail.

II. Decision-making models. The chapter's second section will discuss the manner in which humans make a decision in a dynamic situation.

III. Visual system. A key component to identifying the presence of a weapon is the visual system. As such, the third section will explain how the visual system physiologically works as well as its role in image recognition.

IV. Learning models. The last section of this chapter details adult learning models as they apply to use of force training and training retention.

Use Of Force

Before case law or force correlates can be discussed, it is helpful to give a base definition to understand what is meant by *force*, *excessive force*, and *deadly force*. These terms are utilized throughout the use of force literature. Force is the most general term utilized, and there is no single, agreed upon definition of force (IACP, 2001). This paper will utilize the IACP (International Association of Chiefs of Police) definition. IACP defines force as “that amount of effort required by police to compel compliance from an unwilling subject”. The amount of effort required can include the use of physical takedowns, compliance techniques (either verbal or physical), impact weapons, chemical agents, electronic restraint devices, canine encounters, or firearm discharges. When one discusses any means utilized by a law enforcement officer to restrain a citizen or compel a specific behavior from a citizen, this is force (Fyfe, 1988; IACP, 2001).

Excessive force. Force is deemed excessive when it exceeds the degree of force permitted by law or the policies and guidelines of the particular law enforcement agency. IACP defines excessive force as the application of an amount and/or frequency of force greater than that required to compel compliance from a willing or unwilling subject (Fyfe, 1988; IACP, 2001).

Deadly force. Deadly force refers to any force that may be reasonably expected to cause death or critical bodily injury to a person. This level of force generally requires the officer to utilize weapons, although an officer can apply deadly force with his or her bare hands. The most common tool utilized to apply deadly force to a police-citizen encounter is a firearm. Any time an officer fires his/her firearm deadly force is being applied. However, this does not mean the civilian will be killed (e.g., the civilian may live through the deadly force if the projectile does not strike a vital organ). Rather, the force is considered deadly if it can result in the death or critical injury of the individual.

Use of Force Legal Precedent

A discussion in to use of force requires an understanding of the legal precedent regarding use of force. There are three specific cases that summarize use of force case history. These include *Tennessee v. Garner* (1985), *Graham v. Connor* (1989), and *Johnson v. Glick* (1973). These three cases provide law enforcement organizations with legal precedent and boundaries that control use of force encounters.

Johnson v. Glick, 1973. A jail detainee, Johnson, claimed a correctional officer, Officer Fuller, beat Johnson while he was in custody. Johnson testified that Officer Fuller reprimanded Johnson and other detainees for failing to follow instructions. It was claimed that Officer Fuller went into Johnson's holding cell, grabbed him by the collar, and struck him in the head twice. Even though the complaint was dismissed in court, *Johnson v. Glick* plays a pivotal role in use of force cases. In their discussion, the court developed a four-pronged test. The prongs are derived from the following case brief detailing what law enforcement should look for:

> "look to such factors as the need for the application of force, the relationship between the need and the amount of force that was used, the extent of injury inflicted, and whether force was applied in a good faith effort to maintain or restore discipline or maliciously and sadistically for the very purpose of causing harm." (*Johnson v. Glick*, 1973, p. 1033)

This four-pronged test is applied to cases when force is in question. If the use of force does not meet the four factors of the Glick test, the use of force may be deemed excessive.

Graham v. Connor, 1989. In 1989, Delthorn Graham asked a friend to drive him to a convenience store. Graham was diabetic and was in need of orange juice to counteract his recent insulin reaction. Graham entered the store, but he found the line to be too long. Therefore, Graham exited the store in a hurry and asked his friend to drive him to another friend's house instead. A police officer, Officer Connor, thought the manner in which Graham entered and exited the store was suspicious. Connor performed an investigative stop while he attempted to find out what happened in the store. Graham was handcuffed by additional police officers. The police officers did not listen to Graham's attempts to tell his side of the story; therefore, Graham did not receive treatment for his insulin reaction. Ultimately, Graham was released after Connor learned nothing had happened in the store. In his suit, Graham listed several injuries sustained from the detention. Furthermore, it was alleged that Connor and the additional officers used excessive force in making the stop. In deciding *Graham v. Connor,* the court applied the aforementioned four-pronged Glick Test. The court determined that a reasonable jury applying the Glick test to the evidence presented could not find that the force applied by Officer Connor to be constitutionally excessive (*Graham v. Connor,* 1989).

In reviewing the case, the Supreme Court developed the standard of "objective reasonableness" as a new standard for use of force cases. This standard is applied to all use of force cases. The court's decision specifically states that the force used must be "... objectively reasonable in view of all the facts and circumstances of each particular case…”, and that officers must act in accordance with what a reasonable officer would do in a similar situation (*Graham v. Connor*, 1989, p. 396). Furthermore, this reasonableness calculus “must embody an allowance for the fact that police officers are often forced to make split-second decisions about the amount of force necessary in a particular situation (*Graham v. Connor*, 1989, p. 490). While the standard is defined by the Supreme Court as being "objectively reasonable", there is a substantial amount of subjective interpretation in deciding what a "reasonable officer" would do. This process of interpretation is subjective because there is no clear definition of what is considered reasonable. For example, IACP attempts to clarify what is required in the reasonable officer test. Rather than focusing on the intent of the officer, the reasonable officer test is concerned with the objective circumstances of the specific event and what other reasonable officers would conclude as acceptable (IACP National Policy Center, 2005).

Tennessee v. Garner, 1985. *Graham v. Connor* established the objective reasonableness standard as it applies to police use of force situations; however, *Tennessee v. Garner* challenged the application of the reasonableness standard to established statutes. The state of Tennessee had a statute dealing with civilians that fled or forcibly resisted arrest after being given notice of the arrest by law enforcement officers. This statute allowed officers to use all necessary means to effect the intended arrest of felons. On October 3, 1974, this statute was put to the test in Memphis, Tennessee.

Officers Hymon and Wright responded to a breaking and entering call around 10:45 pm. Officer Hymon walked to the back of the house. Once there, he heard a door slam and witnessed the suspect (Edward Garner) fleeing. Garner reached a 6-foot tall fence. Officer Hymon used his flashlight and saw no sign of a weapon. After telling

Garner to *halt*, Garner began to climb the fence. Hymon was sure if Garner climbed over the fence he would escape capture; therefore, Hymon shot him. Garner died on the operating table after being struck in the head with the bullet. Garner had $10 and a stolen purse from the residence.

The court held that the Tennessee statute was in violation of the Fourth Amendment. Namely, the use of deadly force was to be considered a seizure under the context of the Fourth Amendment. Since deadly force is considered a seizure, the act is subject to the reasonableness requirement of the Fourth Amendment and *Graham v. Connor*. Accordingly, law enforcement may only use deadly force when pursuing an unarmed suspect only if the officer has probable cause to believe that the suspect poses a significant threat of death or serious physical injury to the officer or others.

Implications for Presence of a Weapon. The three prior cases lay the foundation for departmental use of force policies. Namely, a law enforcement officer's use of force must be reasonable in order to be deemed justified. As was shown in *Tennessee v. Garner*, deadly force utilized for an unarmed, fleeing felon was considered a seizure and unjustified. The courts in *Tennessee v. Garner*, and other use of force related cases, utilize two methods to assess if an individual application of force is justified—the Glick test and the standard of objective reasonableness. The four-prongs of the Glick test include 1) a need for the application of force; 2) a relationship between the need and the amount of force that was used; 3) the extent of injury inflicted; and 4) whether force was applied in a good faith effort to maintain or restore discipline. Twelve years later, the Supreme Court of the United States utilized the Glick test to develop a standard of objective reasonableness. Under this standard, law enforcement officers must apply force according to what a reasonable officer would do in a similar scenario. When determining the objective reasonableness of a use of force situation, the decision must allow for the dynamic, split-second decisions regarding the amount of force necessary.

Use of force case history plays an important role in determining the legality of deadly force situations when a weapon is present. If a law enforcement officer utilizes deadly force when a civilian is using a firearm, he or she would most likely be found to be justified in their application of force under both the Glick test and standard of objective reasonableness. In this situation, there is a proportional need for deadly to restore discipline. Additionally, if a civilian merely possesses a weapon in the open, the officer would most likely be found to be justified if their response to the weapon and situation fits within the Glick test and standard of objective reasonableness. The mere presence of a weapon would not pass muster for the application of deadly force if the weapon were not a threat to the officer or public. For example, if an individual is by himself in a park with a firearm in his waistband, it would not be objectively reasonable for law enforcement to apply deadly force, nor would the last three parts of the Glick test be satisfied if the officer decided to apply deadly force (e.g., only the first part detailing a need for some level of force would be satisfied). Regardless, the identification of a weapon on a suspect can justify the application of some level of force as applicable to the specific situation. The inverse is also true, the incorrect "identification" of a weapon makes it much more likely that the application of force will not meet the standards and the use of force will be found to be excessive. In turn, the officer is opened up to a variety of criminal and civil penalties.

Organizational and Contextual Factors in Law Enforcement Decision Making

The previous section outlined pertinent case law describing what constitutes acceptable standards for a law enforcement officer's use of force. While important, law enforcement officers may also be influenced by several different factors while making a use of force decision. As such, the upcoming section will give a brief description of the force continuum and correlates of police use of force. While the force continuum will vary by department, the correlates of force were discovered through multiple research endeavors. These correlates will be broken down to micro, macro, and organizational levels of analysis.

Use of Force Continuum. Terrill & Paoline (2012) reported 97% of departments have a written less-lethal policy in place to guide law enforcement officers regarding how much force to apply in given situations. In general the amount of force exerted should be proportional to the suspect's resistance. These policies act along a continuum (i.e., sliding scale) of force where the officer's force increases according to increases in suspect resistance. Additionally, as a suspect deescalates so does the officer's application of force. An example of a force continuum would be as follows.

First, the mere presence of law enforcement personnel acts as a sign of force. When an officer arrives to a call for service, his or her presence acts as a indicator that legal action may be applied if necessary. The second level of force entails officer verbalization. Officers may command civilians to stop performing whatever act of disorder is present before moving to the next level. Empty-hand control is generally the third item on a force continuum. Hand control can be both soft (e.g., holding a citizen's arm) and hard (e.g., fisticuffs to counter an attack). Fourth on the continuum are generally less-lethal methods of force. This level generally includes oleoresin capsicum (OC) spray, batons, and CEDs. Researchers have found CEDs placed with hand control methods, after hand control methods yet before other less-lethal method, or after other less-lethal methods of control (see, for example, Thomas et al., 2010; Alpert & Dunham, 2010; Smith et al., 2010). The last level of force on the continuum is lethal force. In order for lethal force to be reasonable the suspect must pose a serious, imminent, and life-threatening hazard to the officer or another individual in the immediate area (Terrill & Paoline, 2012).

Terrill & Paoline (2012) further explicated the force continuum by studying how departments designed their continuum. The authors differentiated between four distinct designs—linear, modified linear, matrix, and wheel. A linear design is the traditional continuum design. An officer starts with mere presence and moves through the continuum proportionate to the suspect's actions. The modified-linear design follows the same basic idea, but each level is defined with examples. The officers learn what response is expected for each given situation. This model allows for an officer to quickly move to an appropriate force level. The matrix model is simply another method of presenting linear relationships between suspect behavior and officer use of force. Each is placed on a horizontal matrix for easy memorization (Terrill & Paoline, 2012). The last, wheel design, labels suspect actions in a small circle with appropriate officer responses in a larger enveloping circle. This model instructs officers not to assume a linear progression (Terrill & Paoline, 2012, p. 40). Terrill and Paoline (2012) found that 80% of departments in their sample (n=7,306) utilized some form of a force continuum. Of the four options, 73.4% of departments utilized a linear design (a combination of linear and modified linear).

Terrill and Paoline's (2012) results suggest that the majority of officers are trained to follow a linear continuum to ensure proportionate force is applied. Figure 1 illustrates a typical force continuum. In the example, the police response at the bottom of the figure corresponds with the suspect's behavior. In this example, the suspect cooperates at the lowest level, and no force is required. At the next level, the suspect can exhibit psychological intimidation (e.g., yelling during a protest), which warrants police presence at the scene to passively control the individual. If a suspect verbally will not comply, an officer will issue compliance commands in an effort to have the suspect comply. Next, if the suspect passively resists (e.g., no response to verbal commands) the officer may use hand control techniques. If the suspect physically resists or acts aggressively, the officer may utilize chemical agents or other less lethal options. These two activities vary in their order by department. Lastly, if the suspect displays deadly aggression (e.g., firing a weapon at another individual) an officer can use deadly force.

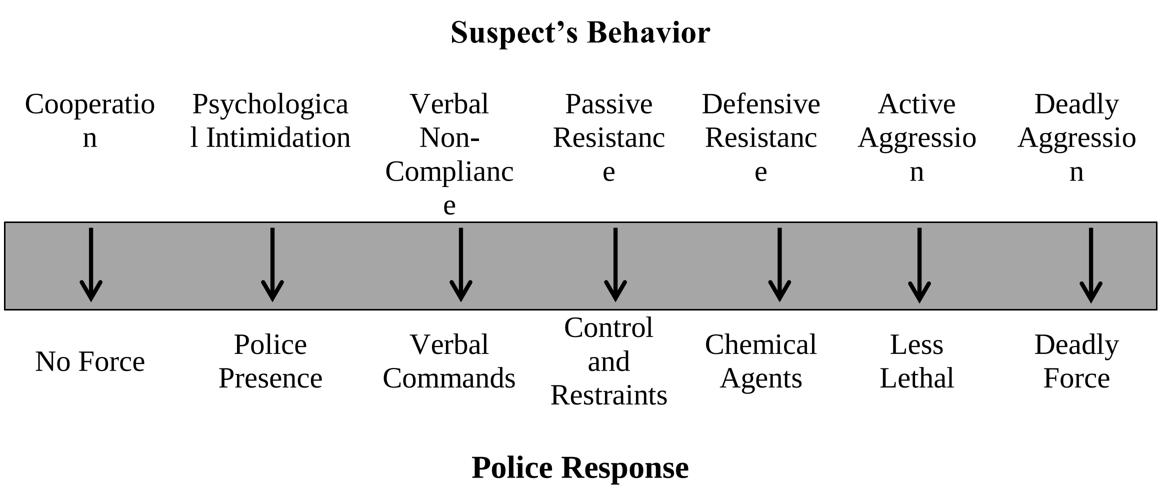

Figure 1. Example of Force Continuum

Beyond departmental use of force policies, the extant literature on use of force has identified three sets of factors that relate to the use of force. These factors can be classified as micro (i.e., individual factors), macro (i.e., community factors), and organizational (i.e., departmental factors). Each of these is reviewed in more detail next to exhibit the complexity of use of force decisions in policing.

Micro-level Correlates. The correlates of force can be viewed from a micro level. Micro-level correlates focus on contextual and civilian factors as they relate to force application. Much of the police use of force research involves suspect demeanor and how it affects the application of force (see, for example, Friedrich, 1980; Meehan, Strange, & McClary, 2015; Worden, 1995; Terrill & Mastrofski, 2002; Garner, Schade, Hepburn, & Buchanan, 1995; Garner, Maxwell, & Heraux, 2002). Researchers have consistently found that the behaviors of the individuals interacting with police officers are the strongest predictors of the use of force (Friedrich, 1980; Worden, 1995; Terrill & Mastrofski, 2002; Garner et al., 1995; Garner, Maxwell, & Heraux, 2002; Hickman et al., 2008; Lee, Jang, Yun, Lim, & Tushaus, 2010; Meehan, Strange, & McClary, 2015). For example, police officers are more likely to use force when a suspect exhibits a negative demeanor, resists arrest, or is carrying a weapon.

Research into the effect of demographic characteristics on the use of force has suggested that these factors are either weakly or inconsistently related to the use of force. For example, some researchers report that force is disproportionately utilized against African Americans (Holmes, 2000; Garner et al., 1995; Payne, 2001; Worden, 1995). Other scholars have argued that once various other factors are accounted for (i.e. actions by suspects, type of offense, presence of weapon, and available law enforcement options) the effect of race on the use of force is diminished (Lee et al., 2010; McElvain & Kposowa, 2008).

The relationship between age and officers' use of force has also been inconsistently supported in the literature. Some scholars (Terrill & Mastrofski, 2002; Worden, 1995) posit that younger adults are more likely to receive a higher level of force from police. Specifically, Terrill and Mastrofski (2002) found a weak, negative relationship with suspect age after controlling for sex, race, socioeconomic status, and situation specific variables (e.g., presence of weapon, demeanor). However, Lawton (2007) and Terrill, Leinfelt, and Kwak (2008) found no significant correlation between age and police use of force once situational variables were controlled for. Intoxicated persons have also been found to receive greater levels of force. This appears to occur primarily because intoxicated suspects are more likely to resist police (Friedrich, 1980; Garner et al., 1995, 2002; Terrill et al., 2008; Worden, 1995).

The mental capacity of offenders has been studied as a correlate to the use of force (see, for example, Engel and Silver, 2001; Johnson, 2011). Research suggests that mentally handicapped individuals are more likely to commit serious offenses, exhibit a hostile demeanor, physically resist police, and be under the influence of alcohol, which as discussed above are related to increased use of force level (Terrill et al., 2008; Worden, 1995; Terrill & Mastrofski, 2002; Garner et al., 1995, 2002; Hickman et al., 2008; Lee et al., 2010) The relationship between mental handicaps and the use of force appears to be weak once these factors are controlled for. Johnson (2011), for example, found that mental instability was positively correlated with substance abuse, abusive demeanor, presence of a weapon, and physically resisting police; however, when level of force was regressed on the contextual variables, mental instability was not a significant predictor of force. Rather, the mentally unstable were more likely to possess weapons and resist arrest than the mentally stable suspects.

Officer and incident specific characteristics may also play a role in the level of force. Some authors have found that male officers are more likely to utilize force or the threat of force (Garner et al., 1995, 2002; McElvain & Kposowa, 2008). Others contend there is no difference in the use of force based upon officer gender (Lawton, 2007; Terrill & Mastrofski, 2002). The effects of a police officer's age, level of experience, and education on the use of force have also been shown to be inconsistent. Several authors contend that younger, inexperienced officers are more likely to use force (Friedrich, 1980; Terrill & Mastrofski, 2002), while Niederhoffer (1967) determined more experienced officers within a patrol position utilized higher levels of force due to an authoritarian mentality. McElvain and Kposowa (2008) posit older, more experienced officers promote to assignments that reduce the likelihood of force being applied. Worden (1995) found that educated officers (i.e., Bachelor's recipients) were more likely to engage in force. However, several other studies found officers with increased education are less likely to use force (Aamodt, 2004; McElvain & Kposowa, 2008; Paoline & Terrill, 2007; Terrill, 2009) and receive fewer citizen complaints. Officer race does not appear to be related to the use of force (Garner et al., 1995, 2002; Lawton, 2007; McElvain & Kposowa, 2008; Terrill & Mastrofski, 2002; Worden, 1995).

The incident specific factors also appear to affect the use of force. The presence of bystanders has been found to be positively correlated to increased use of force (Friedrich, 1980, Garner et al., 1995, 2002; Lawton, 2007; Terrill & Mastrofski, 2002). Additionally, proactive policing practices (i.e., officer initiated police-citizen contacts) are more likely to result in police use of force (Garner et al., 2002; Terrill et al., 2008;

Terrill & Mastrofski, 2002).

Macro-level Correlates. Several studies have examined the relationship between neighborhood characteristics and the use of force. Crime rates, both overall crime rates and homicide rates, show inconclusive and conflicting results in regards to predicting increased use of force. Some authors (Bayley & Mandelsohn, 1967; Fyfe, 1980) posit that higher crime rate neighborhoods receive a higher frequency and severity of force from police including lethal force; however, others (Klinger, 1997; Lawton, 2007) found neighborhood crime rates to be an insignificant predictor of force. Terrill and Reisig (2003) found a positive relationship between homicide rates and increased police use of force. Additionally, economically disadvantaged areas are prone to higher levels of force. Terrill and Reisig (2003) attributed this finding to officers' encountering higher numbers of disrespectful citizens in disadvantaged areas. As previously discussed at the micro level, a hostile demeanor is a strong correlate to police force. Lee et al., (2010) and Smith (1986) found that unemployment rates exhibit a positive relationship with police use of force.

Several authors (Klinger, 1997; Mastrofski et al., 2002; Smith, 1986; Terrill & Reisig, 2003) suggest that the perceived risk is intensified through neighborhood contextual issues and, in turn, positively affects police use of force. While it was previously discussed that race, at the individual level of measurement, does not exhibit a strong relationship with use of force, Smith (1986) found that officers exercised greater levels of force in neighborhoods with high densities of minority residents or racial heterogeneity. However, a more recent study by Lawton (2007) found racial heterogeneity was not a correlate of the use of force. Furthermore, Lee et al., (2010) also found race does not play an important role in predicting levels of police force.

Organizational Correlates. The organizational correlates regarding use of force include departmental policy, informal subculture, and training. Several authors (Alpert & MacDonald, 2001; Fyfe, 1979; Geller & Scott, 1992) discuss how departmental policies influence police use of force. As an example, Alpert and MacDonald (2001) found that departments with policies in place requiring supervisors to fill out use of force paperwork exhibit lower rates of force. However, higher rates of force are found in departments in which the patrol officer involved in the incident completes the use of force paperwork.

Use of force is not only shaped by formal departmental policies. The informal subculture within each department also plays a role in shaping the application of force (Fyfe, 1988; White, 2001). Informal normative behaviors and internal climate affect the decision to use force and how severe of force to use. If the subcultural norm is to apply force, the department is more likely to apply force at greater occurrence. However, the inverse is also true. If the informal norm is to refrain from the application of force, the department will exhibit less force (Fyfe, 1988; White, 2001).

Departmental training procedures also play an important role in officer use of force. Police departments vary in regards to the quantity and quality of both pre-service and in-service training administered. This subject has not been extensively examined in regards to police use of force. Alpert, Dunham, and Stroshine (2006) discusses successful pre-service training as helpful for officers to peacefully resolve various encounters with citizens. However, in-service training must be continually completed to keep officers up-to-date on changing policies and tactics (Alpert et al., 2006). Marion (1998) and Morrison (2006) echo this sentiment by claiming pre-service training does not fully prepare probationary officers to adequately complete their job duties.

Use of Force and Injury

When force is used, the officer and suspect are likely to be injured (i.e., require medical attention at some level). Durose, Schmitt, and Langan (2005) found that 14% of suspects that had force used or threatened against them reported receiving an injury from the encounter. Meyer (1992) analyzed suspect injuries in proportion to the level of force applied. He found that injuries increased as the level of force increased. Specifically, hand control/striking resulted in injury 46% of the time, punching resulting in injury 64% of the time, and the use of a flashlight resulted in moderate or major injuries 80% of the time.

Various authors (Kaminski et al., 1999; Smith & Petrocelli, 2002) discuss the advent of less lethal weapons (e.g., OC spray and CEDs) and the impact on reducing officer and suspect injuries. In relation to hands-on tactics and hard weapons (e.g., batons and flashlights), less lethal methods have been shown to significantly reduce the number of officer and suspect injuries (see, for examples, Kaminski et al., 1999; Smith & Petrocelli, 2002; Strote & Hutson, 2006; Williams, 2008). It is the position of this paper that improving an officer's ability to detect the presence of a weapon can reduce the number of injuries sustained in the long term. Namely, improving an officer's ability to quickly identify an object as a weapon or non-weapon may reduce the level of force required; thus, there may be a reduction in the frequency of officer and civilian injuries.

Decision-Making

While the prior section provided details related to law enforcement use of force, the current section will detail the decision-making process. Decision-making philosophies seek to shed light on how individuals process information to be utilized in making a decision. The following section outlines several different decision-making philosophies from various fields. Each philosophy is examined for its usefulness in the proposed study. These philosophies can be roughly categorized as being either static or dynamic in nature. This classification stems from the environment in which the theories are applicable. Static decision-making theories are those most useful for studying the underpinnings of the decision making process in a laboratory. For example, prospect theory describes economic driven decisions regarding the manner in which individuals delineate between probabilistic risk alternatives in financial decision-making when the outcome is a known value (Kahneman & Tversky, 1979). The core principal for prospect theory is that individuals determine the utility of financial gains or losses from individual reference points rather than final monetary value. While static decision theories, such as prospect theory, are important in many fields, these theories do not shed light on the decision making process in a highly volatile, dynamic situation such as a police-citizen contact where force may be applied. While a police-citizen encounter can be studied in a laboratory setting, studying these types of events in a natural setting can be beneficial. As such, the difference between laboratory and natural decision applications is important.

The distinction between laboratory and natural decision applications was clarified by Tetlock (1991), who said subjects in laboratory studies “rarely feel accountable to others for the positions they take. They function in a social vacuum…in which they do not need to worry about the interpersonal consequences of their conduct” (p. 453). Law enforcement officers must consider the consequences of their conduct at all times. It should also be noted that static decision theories generally rely on simple problems requiring a decision. These theories are best applied if one has time to perform a cost/benefit analysis of their decision's outcome. However, the dynamic decision theories are aptly applied to more complex situations where large amounts of information must be processed in a very short time frame. Complex situations require an efficient method of decision making to quickly determine an appropriate action. Dynamic theories are most applicable to real-life decisions made in the field. For this reason, the discussion will focus on dynamic theories of decision-making.

Dynamic Decision Theories

Dynamic decision theories are applicable to natural applications of choice. There are several dynamic decision-making models, such as bounded rationality and elimination-by-aspects theory. Bounded rationality attempts to explain how individuals utilize shortcuts (i.e., heuristics or rules of thumb) to produce reasonable and workable solutions to complex problems when the individual is constrained by limited information, cognitive constraints, and the finite time (Gigerenzer & Goldstein, 1996). Elimination-by-aspects theory refers to the process of decision-makers eliminating potential options that do not fit the problem (Tversky, 1972). The individual continues to eliminate options until the most viable option is left (e.g., stepping up or down the force continuum proportional to a suspect's exhibited behavior). While these two dynamic decision-making models are useful for complex decisions, they are not ideal models for identifying weapons to aid in dynamic use of force situations. However, there is one decision-making model that effectively describes the process of visually identifying an object, gathering information about the object, making a decision, and acting on the decision—the OODA loop.

OODA Loop. In terms of law enforcement encounters with the citizenry, the law enforcement officer must first gather pertinent information (e.g., visual or auditory data) before making a decision restricting an individual's 4th Amendment rights. One model that can be used to explain this process is the OODA loop. The OODA loop refers to the decision cycle of Observe, Orient, Decide, and Act. The OODA loop is sometimes referred to as Boyd's Cycle, as it was developed by United State Air Force Colonel John Boyd (Boyd, 1996). Each phase of the OODA loop will be discussed next.

Observe. The first stage, *Observe*, refers to the moment an individual is made aware of an object. This awareness manifests through the various senses. While one could observe an object through taste or touch, the two senses most commonly utilized are visual and auditory. This phase is not responsible for gathering any information. Rather, the observation of an object merely initiates the decision making process. A policing example would be that of a patrol officer responding to a call for a suspicious person walking around a car dealership late at night. Once the officer arrives on scene, the officer would direct his or her gaze to the individual once the individual comes across the officer's field of view (Boyd, 1996).

Orient. The second stage, *Orient*, refers to the process of gathering information from the observed object. This stage results in the individual having information to serve as a basis for the decision making process. When the observation is visual, the individual will orient to the object via the visual system. In other words, the individual will utilize the visual system to gather as much pertinent data as possible (Boyd, 1996).

This phase of the OODA loop has been researched independently. Sokolov, Sprinks, Näätäen, and Lyytinen (2002) present orienting responses (ORs) as the building blocks of all investigatory processes. According to the authors, orienting responses can have physiological effects. During an orienting response, the body readies itself to gather the most amount of information possible about the potential threat/stimulus. A well-known physiological example of an orienting response is dilation of the pupil when light levels are reduced. The visual system adapts to the lighting conditions in order to continually gather data.

Utilizing the example of a suspicious person at the car dealership, the officer will orient toward the individual once he is in the officer's field of view. While performing a visual search of the individual, the officer may observe an object in the individual's hand.

The example illustrates the process of orienting in a complex situation. While orienting in a complex situation an inhibitory effect may be experienced and radically slow down the decision-making process for the individual by forcing him or her to gather additional information (Sokolov et al., 2002). This inhibitory effect can be overcome through familiarization (Sokolov et al., 2002). Familiarization is simply the process of repeated exposure to a stimulus or object to the point of acquaintance with the stimulus or object. Returning to the example, once the officer notices the object in the suspect's hand, he or she will further orient on the object. It is at this point that the officer can experience, or overcome, an inhibitory effect. In this instance, the object is a handgun that the suspect is raising to shoot at the officer. For the responding officer, the positive identification of a firearm is paramount to a successful outcome. Familiarization in this scenario may be attained by training the visual system to differentiate weapons from non-weapons. All of these data gathered during the orienting process are utilized to make a decision regarding the best course of action.

Decide. The third stage, *Decide*, refers to the process of making the actual decision. The individual will process the data gathered during the *Orient* stage of the OODA loop. The officer in the armed person at a car dealership example will utilize all the information gathered regarding the presence of the weapon, the movement of the weapon, the situational factors (e.g., time of night, location, suspect demeanor), and the safety of the public to make the decision to apply lethal force. The decision will be influenced by an individual's past experiences, training, and department policies to name a few. All of these factors, including the data gathered during the orienting phase, are combined. In a confrontation that requires lethal force the law enforcement officer must decide on the proper, legal action to take in a split-second (Boyd, 1996).

Act. The fourth stage of the OODA loop, *Act*, refers to the action deemed appropriate during the prior decision stage. It should be noted that action and inaction both fall within the scope of a decision. If one decides not to do anything that is an action in and of itself. The officer in the example could apply lethal force or possible retreat and issue compliance commands. After the decision is made, the actual action of drawing his or her weapon and firing on the suspect or retreating and issuing compliance commands completes the loop (Boyd, 1996).

The prior example of applying lethal force does not end the process of the OODA loop. In fact, the process is repeated indefinitely. The action taken will then prompt the individual to evaluate the effectiveness of the action. At this point, the OODA loop begins again. Once the officer draws and fires his or her weapon, the officer will need to evaluate the effectiveness of the action and decide if further action is needed. The manner in which the officer proceeds is dictated by the OODA loop (e.g., the officer will perform the loop to determine if the suspect is still a threat or not). The officer will perform the loop throughout the entire process of securing the scene.

Furthermore, the OODA loop entails a linear movement through processes. One must observe before one can orient. One must orient to an object before a decision is made. And, one must make a decision before an action is taken. This process is competitive when two individuals are interacting with each other. Regarding law enforcement, this process is most evident when there is confrontation requiring the application of force. When two individuals are engaged in confrontation, the individual with the faster OODA loop has a distinct advantage. The faster OODA loop results in an action being taken before the other individual. In the above example, the suspect is also processing an OODA loop. The suspect observes and orients on the officer approaching. The suspect then makes the decision to raise and fire his weapon at the officer. The suspect may initially have a time advantage due to a shortened orienting period and predetermined decision that is simply awaiting a cue (e.g., police officer). Law enforcement officers are at an even greater disadvantage when one considers the need to process all available information prior to making a use of force decision due to the legal ramifications inherent with an incorrect decision. Due to the competitive nature of a suspect confrontation, the speed in which a law enforcement officer can progress through the OODA loop determines their ability to have an advantage during the encounter. Decision speed has been assessed via reaction time research. In fact, there are several authors that have tested reaction time, both with and without law enforcement officers as participants.

Reaction Time Research. One could debate what models of decision-making are applicable to reaction time research. While both bounded rationality and elimination-by-aspects may be studied for their influence on reaction time, the OODA loop is indirectly tested when assessing reaction time—i.e., the process of reacting to a stimulus requires the participant to observe and orient to the stimulus. The amount of time required to observe and orient to a stimulus affects the speed of the reaction. In other words, increased complexity increases the amount of time needed to react (Brebner & Welford, 1980; Luce, 1986). Scholars have found that reaction time to noncomplex visual tasks (i.e., pushing a button when a card is turned over) ranges from .20 to .30 seconds (Eckner, Kutcher, & Richardson, 2010). However, more complex reaction tasks (i.e., correctly indicating whether the card presented is red) were found to be slower (Eckner et al., 2010). These results have been replicated in studies of reaction time as it relates to law enforcement officer use of force (Lewinski & Hudson, 2003a; Lewinski & Hudson, 2003b). When performing simple reaction tasks (i.e., firing a pistol when cued by a light) officers had an average reaction time of .31 seconds (Lewinski & Hudson, 2003a). However, officer reaction time nearly doubled to .56 seconds when performing a more complex task (i.e., firing pistol when a specific series of lights in different rows light up).

In a more recent study, Blair, Pollock, Montague, Nichols, Curnutt, and Burns (2011) studied reaction times of officers when forced to make a deadly force decision in a dynamic scenario rather than a laboratory setting. The suspects in the study were armed with a pistol either pointed to the ground or to their own head. The officer would turn and face the suspect with his or her gun drawn and issued compliance commands. The suspect either complied or attempted to shoot the officer. The officer would attempt to fire on the suspect once they recognized the suspects' intent to fire at the officer. The authors utilized a repeated measures design with four counterbalanced test conditions (i.e., gun high or low and shoot at the officer or surrender). Suspects fired in .38 seconds on average while officers fired in .39 seconds on average. Thus, the process of observing a suspect's hostile movement, interpreting the movement, deciding to take lethal action, and acting on the decision resulted in the officer firing after the suspect even though the officer already had his or her gun aimed at the suspect before the hostile movement. The authors closed by suggesting improvements in training to improve reaction times of law enforcement officers and highlighted the importance of avoiding a similar situation if possible (Blair et al., 2011).

Acute Stress Response. While the orienting response will naturally tend to draw an officer's attention to threats, there is another process that is also important. This process is acute stress response. Acute stress response is a set of physiological changes to a perceived threat. This process is also sometimes referred to as the fight, flight, or freeze response. Some of these physiological changes include increasing the heart rate to improve blood flow, tensing muscles to prepare for quick movement, and dilating the eyes to increase imagery light. The response is dependent on perception and training; specifically, how serious you perceive the threat to be and how quickly your training is recalled determines how you respond to a threat. The level of stress also affects how well you are able to perform.

Grossman and Christensen (2008) spend considerable time discussing the harmful aspects of stress on an officer. The authors discuss stress and performance along a color-coded, sliding scale. White and yellow are attributed to a state of rest and partial arousal.

These are generally seen with a heart rate below 115 BPM. A level red reflects optimal physical performance with a BPM up to 145. At this level, the heart's output fuels the brain, muscles, and necessary organs for optimal performance. After this level senses and fine motor skills begin to lose their ability to function properly. These levels are known as gray and black. At this level of physical stress, fine motor skills diminish. The body also undergoes a widespread vasoconstriction. Blood stops flowing to the hands and eventually the major muscle groups are deprived of precious oxygen. These physiological processes also produce a number of psychological and physiological distortions.

Substantial research into the rare occasions when police utilize deadly force has revealed that the acute stress response often produces distortions in the perceptions of officers involved in shootings. The three most researched distortions are tunnel vision, audio exclusion, and time distortion. Tunnel vision occurs when the peripheral vision is no longer perceived. Officers experiencing tunnel vision only register visual input for images that are in the fovea. Audio exclusion (i.e., auditory blunting) results in an officer lacking the ability to hear what is happening during the situation. Time distortions involve the officer perceiving the event as occurring either faster or slower than normal (Solomon & Horn, 1986; Campbell, 1992, Klinger & Brunson, 2009).

In his study of 167 officers involved in shootings, Campbell (1992) found that 44% experienced tunnel vision, 42% experienced auditory blunting, and 34% experienced slow-motion time. A study conducted by Artwohl and Christensen (1997) was largely consistent with Campbell's study, but Artwohl and Christensen also identified new forms of perceptual distortion, including intensified sound and heightened visual detail. These occurred in 16 and 72 percent of the cases respectively.

The most recent research into perceptual distortion was completed by Klinger and Brunson (2009). The authors established a snowball sample of 80 officers (113 total incidents) that had shot citizens in the line of duty. They found varying levels of auditory blunting in 82% of cases. Fifty-one percent reported tunnel vision, while 56% reported heightened visual acuity. Slow-motion time was experienced by 56% of the officers and fast-motion time occurred in 23% of the cases. The authors also studied the temporal variability of perceptual distortion to determine if the distortion occurred prior to shooting or during the shooting. Tunnel vision occurred prior to firing in 31% of cases and during firing in 27% of cases. Auditory blunting occurred prior to firing in 42% of cases and during firing in 70%. In all, 87% of officers reported some form of perceptual distortion before shooting and 92% during a shooting (Klinger & Brunson, 2009).

Interaction of Orienting Response, Reaction Time, Stress Response and Decision-Making. The prior discussion highlighted several key factors that can influence a law enforcement officer's ability to quickly identify a weapon and make an actionable decision. There are three specific takeaways present: 1) orientation time impacts decision-making; 2) acute stress response impacts one's ability to properly orient and make a decision; and 3) the overall process of decision-making influences one's reaction time.

First, the amount of time spent orienting to an object/environment in a complex situation is positively correlated with one's ability to make an accurate decision. Researchers have found as visual complexity increases so does the amount of time required to gather information (Just & Carpenter, 1976). This process is known as an inhibitory effect. The individual is required to slow down and gather additional information. However, it also appears that training can reduce one's inhibitory effect in a complex situation through habituation to the stimulus. In terms of weapon identification, visual training to recognize the presence of a weapon can habituate the trainee (i.e., a law enforcement officer habitually recognizing distinct properties of a weapon without having to spend additional time orienting to the object). In effect, visual habituation can speed up the orienting stage by requiring less visual data be gathered to make a correct decision.

Second, the physiological and psychological aspects of acute stress response can act as an inhibitor to one's ability to gather needed data for accurate decision-making. Recall, when an individual experiences acute stress he or she can undergo perceptual distortions. The most influential distortion in terms of weapon identification is tunnel vision. Both Campbell (1992) and Klinger and Brunson (2009) found approximately 50% of their sample experiences tunnel vision. This process can prevent an officer from ever seeing a weapon. For example, if a law enforcement officer experiences tunnel vision and only focuses on a suspect's face, he or she may not notice the presence of a firearm in the hand of the suspect. In this example, the law enforcement officer is unable to gather pertinent data to aid in making an accurate and potentially life-saving decision.

Third, it is evident that the overall process of making a decision impacts a law enforcement officer's reaction time. As previously discussed, Lewinski and Hudson (2003) found that officers took almost twice as long to fire their weapon when the cue to fire was more complex. Blair et al., (2011) assessed law enforcement officers' ability to react to a hostile movement. They found that suspect action was faster than law enforcement reaction, on average. Their study design had the suspect on a much shorter OODA loop (i.e., the suspect already knew if he or she was going to fire the weapon or not while the officer had to process the scenario). This highlights the importance of visual training to speed up the early phases of the OODA loop for law enforcement officers. If the action of the suspect is 0.01 seconds faster in a laboratory setting where the officer did not have to identify the presence or absence of a weapon, it stands within reason that the action of the suspect will be even faster in a real-world dynamic situation.

Vision

The prior discussion of the OODA loop illustrates the important role the visual system contributes to the first two phases—Observation and Orientation.

If a potential threat is detected in the periphery of the officer's vision during a room entry, the orienting response will automatically reorient the eye to place the image of the potential threat in the optimal visual position in order to gather better information about the possible threat. In short, the natural tendency of the entering officer will be to look at the threat. In some cases, the orienting response may be strong enough that it halts the other behaviors (e.g. moving) in order to maximize the information acquired. If this occurs, the officer may momentarily stop (stall) in his current position to gather more information.

Function of the Eye

The visual system is a complicated mechanism that converts light reflected from objects in the environment into images. Figure 2 below is a graphical representation of the eye's physiology. Not shown in Figure 2 are the six extraocular muscles. These muscles control the fine movements of the eye. The importance of these muscles will be discussed in a later section.

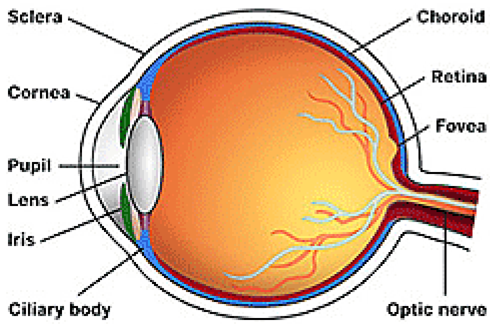

Figure 2. The Eye

Vision begins when the light waves enter the eye via the cornea. The cornea is essentially the window of the eye. However, unlike a flat window, the curvature of the cornea bends the light rays so they may pass through the pupil. Figure 3 illustrates the process of corneal refraction. In the image, the light rays are shown as light and dark gray stripes. As the rays contact the curvature of the cornea they are angled in a manner that condenses the rays. This angling of the light rays allows a larger “image” to pass through the pupil (Ahmad, Klug, Herr, Sterline, & Schein, 2003; Ungerleider & Mishkin, 1982).

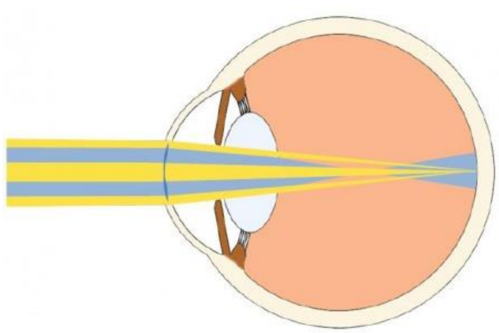

Figure 3. Corneal Refraction

The pupil is simply the opening of the eye. The diameter of the pupil is determined by the iris. The iris is similar to the shutter on a camera. It regulates the amount of light entering the eye by enlarging or shrinking in an attempt to keep the amount of light consistent. Once the light enters the eye it is passed through the eye's lens. The lens is clear and flexible. A series of tiny muscles flex the lens in order to correctly focus the light rays to optimize vision. This focused light is directed to the eye's retina (Ahmad et al., 2003).

The retina is made up of millions of light-sensitive sensors, called rod and cone receptors. Cone receptors are responsible for high-acuity color vision. Rod receptors lack detail and color capabilities, but do respond to motion and light (Ahmad et al., 2003). The retina utilizes the cone and rod receptors to transform light into electrical signals sent via the optical nerve.

While any light contacting the retina will result in an image being processed, the quality of vision is dependent on where the imagery is concentrated on the retina. For optimal acuity, the lens system directs light onto a small position with the densest concentration of cone and rod receptors. This location is called the fovea centralis (fovea). Light that contacts the fovea accounts for the clearest vision (Ahmad et al., 2003). The entire process is shown in Figure 4. In this example, light is reflected from the tree. This reflected light then passes through the cornea and is properly refracted through the pupil. The lens focuses the light so it strikes the fovea centralis. The optic nerve then transmits the impulses to the occipital lobe of the brain where the electrical impulses are processed as a tree. Specifically, the impulses are processed by the visual cortex located in the occipital lobe at the back of the brain (Ungerleider & Mishkin, 1982). The functionality of the visual cortex will be discussed in a subsequent section dealing with visual perception.

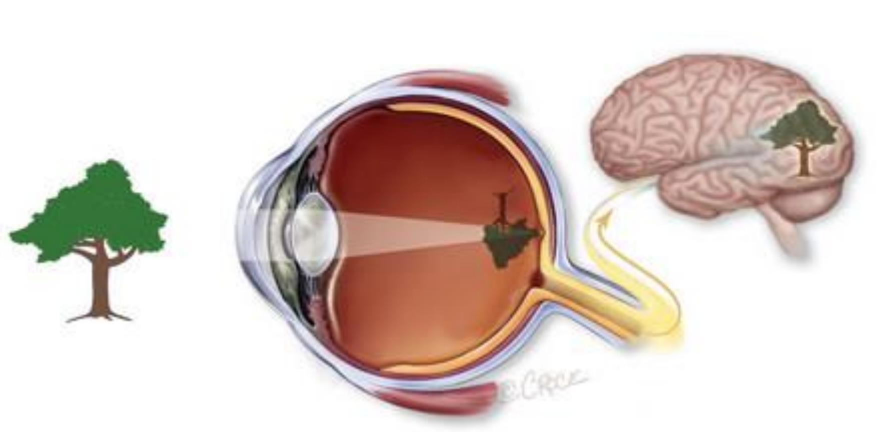

Figure 4. Vision Example

Visual Acuity

Visual acuity is known by a couple of names—e.g., arc of vision, visual angle. Visual acuity is paramount in fully understanding the properties of the eye and how the eye operates in the orientation process.

As previously discussed, optimal vision is obtained when light reflected from an object hits the fovea directly. However, visual acuity drops rapidly as imagery moves away from the optimal angular distance from the fovea—1 to 2 degrees of arc. As an example, the width of the thumb with the arm fully extended covers approximately 2 degrees of arc. Ruch (1965) found that acuity falls to 50% at 2.5 degrees, 25% at 7.5 degrees, and 4% in the furthest periphery.

Peripheral Vision

Peripheral vision refers to the vision occurring outside of the center of foveal gaze. Recall that optimal vision is achieved when light is focused on the fovea. If one imagines the visual field of view (FOV) as if one was positioned in the middle of a circle (360°), the 2 degrees of arc where one is focused for optimal visual acuity would account for approximately 5° of the circle directly where one's gaze is focused (see Figure 5). If one was to continue gazing straight ahead and hold her arms straight out to her side she could see both hands when she wiggles her fingers. The arms would account for slightly more than half of the circle (i.e., slightly over 180° FOV). Therefore, peripheral vision allows people to see slightly over 180°. Furthermore, peripheral vision lacks static sensitivity at the edge of the visual field while maintaining motion sensitivity; this is why the person must wiggle her fingers to acquire the object in the periphery (Ware, 2013).

Peripheral vision is broken into three categories—i.e., near, mid, and far peripheral vision. The near periphery accounts for approximately 30° FOV per eye. The mid periphery accounts for approximately 60° FOV per eye. Lastly, the far periphery can account for as much as 110° FOV for each eye (see Figure 5). Peripheral vision plays a vital role in the visual orientation process. As an object is detected in the periphery, the eye reorients itself to provide more detailed information about the object. This means that the position of the eye changes to place the image of the object of interest on the fovea.

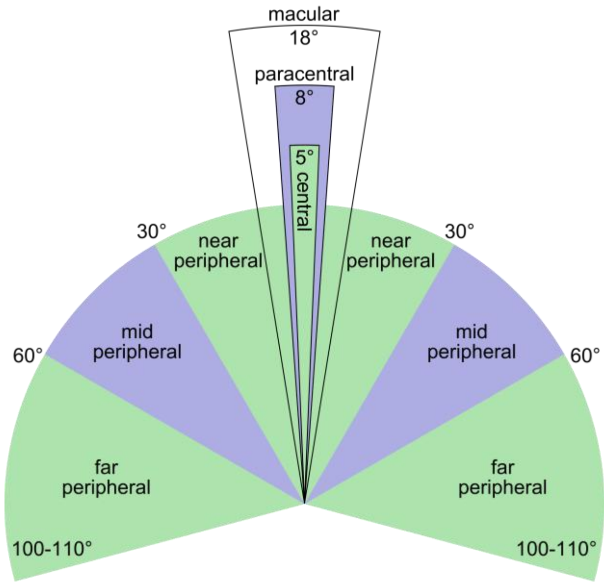

Figure 5. Visual Acuity Example

Visual Perception

Visual perception is the process of processing the light rays received by the eye. Specifically, once the light reflecting from an object has impacted the retina, the optic nerve sends the electrical impulses to the visual cortex. Both hemispheres of the brain contain a visual cortex responsible for processing impulses from the opposing eye's visual field (i.e., the left hemisphere processes images from the right visual field and the right hemisphere processes images from the left visual field).

The impulses pass through the various sections of the visual cortex, each of which is responsible for a separate function. These sections include visual area one (V1; the primary visual cortex), visual area two (V2), visual area three (V3), visual area four (V4), and visual area five (V5). The initial impulse is sent to V1. Each hemisphere's V1 then transmits impulses to two primary pathways known as the ventral and dorsal stream. The ventral stream utilizes V1, V2, V4, and a separate area known as the inferior temporal cortex. The ventral stream is concerned with object perception and is known as the "What

Pathway" and is associated with form recognition and object representation. The dorsal system utilizes V1, V2, V3, V5, and the posterior parietal cortex. The dorsal system concerns itself with spatial perception and is also known as the "Where Pathway" or "How Pathway". This system is associated with motion (i.e., an ability to recognize objects that are both stationary and dynamic), representation of object locations, and control of eye movements. As such, the dorsal system plays a vital role in hand-eye coordination (Chinellato, Grzyb, Marzocchi, Bosco, Fattori, and Pobil, 2009).

Image Recognition and Search Strategies

Researchers have found the amount of time spent fixating on an object is positively correlated to the amount of information gathered and/or the complexity of the image (Just & Carpenter, 1976; Ahmad et al., 2003). Therefore, the more complex an image, object, or situation, the slower an individual is in making an informed decision. There are two predominant theories regarding the complexities of image recognition regarding search strategies—feature integration theory and guided search theory.

Feature Integration. Treisman and Gelade (1980) developed feature integration theory to describe the process of combining individual attributes of an image into a known object. The process of becoming aware of the object is known visual attention. According to the authors, the process of feature integration is split into two phases—i.e., preattentive and focused attention. The first phase in feature integration is known as the preattentive stage. It is during this stage that an object is analyzed for as much data as possible (e.g., color, shape, size, movement). These data are processed independent of each other in the brain. Since this stage asserts that one collects large amounts of data at once, it is considered a parallel search mechanism (i.e., multiple features are assessed independent of each other). It is during the second phase, focused attention, that individual features are integrated with each other. This phase is named focused attention because it requires effort to combine the individual features that were automatically gathered during the preattentive stage into a single object. Since the features are integrated into a single object, this phase is considered a serial search mechanism (i.e., the focus is on a singular object rather than multiple objects). This process of integrating multiple features into a single object through the serial process is known as binding (Treisman, 1999; Von Der Malsburg, 1995). Early studies into feature integration found that more complex images required more time to process. Treisman and Gelade (1980) validated this by studying differences between single-feature and conjunctive targets. In their seminal work, participants searched for a target within a display containing up to 30 items. When assigned to the single-feature condition, participants were required to find an item with a single attribute such as a blue letter. Participants assigned to the conjunctive condition were required to find an object with multiple features such as a green letter T. All the other letters in the conjunctive condition shared a single feature (i.e., all brown letter T or a green letter X). Treisman and Gelade found support for focused attention with the conjunctive condition taking twice as long to identify as the single feature condition (Treisman & Gelade, 1980).

Other scholars posit that feature integration theory is more complex than originally described by Treisman and Gelade (Duncan & Humphreys, 1989, 1992; Humphreys, Riddoch, & Quinlan, 1985). Namely, there are two factors excluded that impact the visual search times. Duncan and Humphreys (1989, 1992) explained that the similarity between the target image and the non-target images, as well as the similarity between the non-targets explained the larger effects seen in the preliminary work by Treisman and Gelade. According to Duncan and Humphreys, if the target and non-target images are very similar then the visual search times are longer as it is more difficult to pick out the target image regardless if it is a single feature or conjunctive image. However, if the non-target images are very similar to one another the visual search times are much shorter as the target image will be easier to locate. This position was based on feature integration research by Humphreys, Riddoch, and Quinlan (1985). The authors had participants located inverted T targets mixed into different quantities of non-inverted T non-targets. According to feature integration theory, the visual search time should have significantly increased as the number of non-targets was increased (these would be considered conjunctive targets as one has a vertical and horizontal line to identify). However, the similarity of the non-targets leant to essentially equal visual search times regardless of the number of non-targets.

Guided Search. Wolfe (1989) developed guided search theory to advance the visual search literature beyond the classic serial and parallel search mechanisms such as feature integration theory. Specifically, Wolfe suggested that the preattentive, parallel phase could guide the focused attention serial phase in conjuncted visual searches. This position differed from feature integration theory's position that the initial process is inherently parallel followed by a serial attention phase; rather, the process of image recognition is much more efficient with conjunctive searches. For example, if one was to search for a red letter T in a field of red and black letters, guided search would suggest that a preattentive process would screen out only the red letters (essentially a serial search for color). In doing so, another serial attentive phase would then perform a single factor search of the red letters for the letter T (Cave & Wolfe, 1990).

Eye Movement. There are two primary types of eye movement—i.e., saccadic and pursuit. Saccade movements are rapid reallocations of foveal fixation from one fixation point to another (Carpenter, 1988; Rosenbaum, 1991). It is during saccadic movement that an individual bounces his or her foveal fixation from one object to another. An example of saccadic movement is searching a crowd of people for a familiar face. The searcher focuses foveal gaze on several faces without searching the space between or around the faces. Scholars have found that saccadic movement reduces the amount of information gathered via the visual system (Ditchburn, 1973; Festinger, 1971; Massaro, 1975). These findings follow Just and Carpenter (1976) and Ahmad et al.'s (2003) finding of a positive correlation between amount of visual fixation time and the amount of information gathered. Rapid saccadic movement does not allow the visual system time to gather information between fixation points.

Unlike the rapid saccadic movement, smooth pursuit movements allow an individual to track slow-moving targets. This process keeps the target object in foveal vision to maintain clarity. An example of pursuit tracking would be following an individual as he or she walks past. Studies have shown if an object is moving quickly, the effectiveness of pursuit movement diminishes (e.g., attempting to follow the puck in ice hockey; Haywood, 1984). In scenarios with rapid movement (e.g., sports with fast moving balls), it has been found that experienced performers utilize saccadic eye movements to predict where the ball will be and then pursuit movements once it is slow enough to track (Bahill & LaRitz, 1984; Ripoll, 1991).

Examples of Visual Training Research

Criminal justice decision making policy and training research have yet to test various aspects of the visual process. However, many relevant examples are found in the extant sports science literature. For example, the sports science vision literature focuses on both static and dynamic visual search. Static visual search research entails showing participants static images and assessing the speed and accuracy of their decisions. Tyldesley, Bootsma, and Bomhoff (1982) examined differences between novice and experienced soccer players' ability to anticipate penalty kick direction based on pictures of players in the process of kicking a ball. Experienced players responded significantly faster than novice players. Additionally, the experienced players gaze fixated on the shooting leg while novice players performed saccadic searches of the image. The inexperienced players' fixation time was 26.6 ms longer, on average, than the experienced players before a decision was made. Tyldesley et al., (1982) also tested the effects of increasing the complexity of the decision. To accomplish this, the authors required participants to choose not only if the player was going to shoot left or right, but also if the player was going to shoot high or low based on their body position. With the additional task, the experienced players utilized twice as many fixation points before responding. Instead of focusing solely on the leg, the experienced players performed a saccade and looked at the shooting hip first 60% of the time. Their second foveal fixation was at the shooting side shoulder (55%). The experienced goal keepers utilized their prior knowledge to determine shot direction by the hip/leg and the height of the shot by the shoulder placement (Tyldesley et al., 1982).

Bard and Fleury (1981) took visual search research out of the laboratory and into the field. Novice and expert hockey goalkeepers responded to a model performing a slap or sweep shot on the ice. The speed of visual initiation and the percentage of time spend concentrating on certain aspects were captured. Expert goalkeepers performed their response faster than novice goalkeepers regardless of the shot condition. Much like prior laboratory research on soccer players, the expert goalkeepers' visual fixation points were different than the novice goalkeepers. For examples, expert goalkeepers fixated their gaze on the stick 71% of the time during a slap shot. However, the inverse occurred for novice goalkeepers. The novice goalkeepers fixated on the puck 70% of the time during a slap shot. Similar findings regarding differences between expert and novice visual fixations have been reported in other sports (Bard et al., 1980; Petrakis, 1986; Petrakis, 1987; Salmela & Fiorito, 1979; Vickers, 1992).

More recent research has attempted to include complementary object recognition models to traditional vision tracking, namely the use of verbalization concurrent with visual fixation. Scholars have agreed that the inclusion of verbal reports may provide insight into higher mental processes present (Brinkman, 1993; Ericsson & Simon, 1993; Green, 1995). Ericsson and Oliver (1989) suggested that verbalization enables participants to identify information as it is being processed by short-term memory. Their position is that the verbalized information may not be available for recall in a retroactive report of the experiment. Others have also suggested that valuable information regarding the visual system can be lost if not verbalized during data collection (Brinkman, 1993; Russo, Johnson, & Stephens, 1989). Furthermore, verbalization can identify recognition of objects not seen in visual tracking playback if the individual utilized peripheral vision (Vickers, 1988).

Video-based Visual Training. While the prior discussion focused on static images and models for visual training, this section will describe the importance of video-based visual training. In fact, several scholars have found that video-based training improves participant performance along multiple dimensions (see, for example, Burroughs, 1984; Christina, Barresi, & Shaffner, 1990; Singer, Cauraugh, Chen, Steinberg, Frehlich, & Wang, 1994; Williams & Burwitz, 1993). Much like the prior discussion, the scholarly work is predominately located in the sports science field of study. Regardless, there are two key takeaways from research utilizing video-based visual training.

1. *Visual Training Can Improve Speed and Accuracy of Object Recognition/Decision-Making.* This dissertation proposes that visual training can improve the speed and accuracy of weapon identification. While there is a paucity of research relating to law enforcement and visual training, this is not the case in other fields. For instance, Christina et al., (1990) studied the efficacy of video-based training on football linebackers’ ability to make accurate decisions. The authors utilized video to give linebackers practice deciding on the correct direction to move based on recognizing offensive play. The video was taped so it was representative of what the linebacker would see in a game situation. As the play unfolded, the linebacker would move a joystick as soon as he knew which direction he would move based on recognizing typical offensive moves. Christina et al. found that accuracy improved dramatically over the four week training period. Furthermore, while accuracy increased, there was no sacrifice in the speed of decisions.

While Christina et al. did not find marked increases in the speed of decisions after video-based training, Haskins (1965) was able to. Haskins developed a video in hopes of shortening the time needed to determine the direction of a tennis ball return shot. Participants viewed a series of return shots from where their point of view would be on a tennis court. Haskins utilized actual match film from before and after the video-based training intervention. He found significant improvements in response time in the post-training matches. This finding suggests that Haskins was able to improve the speed of tennis players' decision-making.

2. *Visual Training Can Teach People Where to Look for Important Cues.* An important aspect of a fast visual search strategy is knowing where to look. Salmela and Fiorito (1979) utilized visual training video for hockey goaltenders. The goaltenders were shown point of view video of a hockey player shooting a puck in their direction. The goaltenders were asked to predict where the shot would go. The authors found that goaltenders accurately predicted the location of the shots based on the visual cues highlighted in the video. Additionally, Burroughs (1984) studied the effectiveness of visual training to enhance baseball players' ability to recognize pitch locations based on visual cues from the pitchers' natural movements. The author also filmed so the participant would have the perspective of actually playing the game. The participants were shown training video in both slow-motion and real time. Burroughs found no difference between the test and control group's ability to distinguish between a fast ball and breaking ball. However, he did discover that players participating in the video-based training were more accurate at perceiving the correct location of the pitch than players that did not participate in video-based training. These two sports-related studies show that video-based visual training can teach people where to look for visual cues. This method of training should cross over to other fields beyond sports, such as law enforcement officer's ability to learn where to look for a weapon.

These two key points (i.e., visual training can improve speed and accuracy and vision training can teach people where to look) are vital for the development of the vision-training program proposed in this study. The vision-training program will utilize the concepts present in the aforementioned studies, such as presenting training video from the point of view of the law enforcement officer.

Learning and Performance

Beyond an understanding of police use of force, decision-making, and the visual process, the manner in which adults learn and acquire expertise (i.e., high level of skill, proficiency, or competence in a task) in performance is important for the proposed dissertation. There are a multitude of adult learning theories found within the psychology and adult education disciplines (see, for example, behaviorism, cognitive constructivism, and humanistic psychology). While theories of learning are important, the proposed research is concerned with how adults acquire expertise of skills through training (i.e., the speed and accuracy of weapon acquisition). Furthermore, the proposed research is intended to illustrate how training certain skills impact overall performance (i.e., training proper visual search strategies). As such, a discussion of relevant skill acquisition research is warranted.

Lifespan Development

One traditional view of expert-level skill acquisition is lifespan development. Lifespan development is based on the assumption that skills and abilities begin development during childhood, and the progression of this development is linear in nature. These skills and abilities are limited by innate biological factors that will constrain the highest level attainable on an individual basis. According to these models, skill acquisition is asymptotic in that individuals reach a plateau of performance (Denney, 1984). Galton (1869) is often credited with the formation of this model. Galton presented findings that height and body size were, generally, genetically established. Therefore, according to Galton, the same genetic predisposition applied to mental aptitude/capabilities. As an example, an individual born to parents with a proclivity for medicine is genetically suited to become an expert in medicine. Galton did acknowledge that training was needed to reach expertise in any domain. However, he posited that fast improvements in performance would only be seen at the beginning of training. The upper bounds of performance were determined by hereditary factors and these hereditary factors established an individual's proficiency at any given task. Utilizing the medical field example, an individual whose parents were farmers would plateau and not reach the same level of expertise as the individual whose parents were physicians.

Galton's (1869) principles are still found in contemporary theories of skill acquisition and expert performance (see for example Anderson, 1982; Simon & Chase, 1973). The contemporary framework discussed by Fitts and Posner (1967) breaks skill acquisition and performance into three phases. The process of learning to play golf will be utilized as a working example to illustrate these phases. According to Fitts and Posner

(1967), the first phase of skill acquisition starts with beginners attempting to understand the requirements and reduce major mistakes while performing the actions. In terms of golfing, a novice golfer would learn the basic body movements necessary to make contact with the golf ball. In addition, the novice golfer exercises much effort to avoid mishitting the ball. The second phase begins once the individual has gained more experience. During the second phase, performance appears smooth, noticeable mistakes become rare, and the learner is capable of maintain an acceptable level of performance without intense focus. According to Fitts and Posner (1967), the second phase can frequently be obtained with less than 50 hours of training. Therefore, a golfer performing in the second phase is capable of striking the golf ball with ease after training for approximately 50 hours. As the individual performs within the second phase, the performance skills become automated. At this point the individual is in the autonomous third phase. During this phase, the individual performs the skills with minimal effort. However, while in the third phase, the individual's automated performance begins to plateau. This plateau is gradual but certain to result in a performance limit in accordance to Galton's (1896) position. Under this framework, no matter how long the individual plays golf she will be limited by hereditary factors.

Continued Experience

Somewhat related is the theory of continued experience. Simon and Chase (1973) propose that learners gradually acquire knowledge about how to react to situations by storing memories of their past actions in similar situations. Simon and Chase's position is similar to the prior discussion of long-term memory as it related to object recognition. The authors propose that approximately ten years of performance are required to be considered an expert. Simon and Chase (1973) studied chess masters and determined that expertise (measured as winning international chess tournaments) was not possible without a decade of playing chess. Thus, a shift occurred with some scientists and expertise was given a longitudinal component instead of a hereditary one. According to these scientists, expertise was a linear progression from novice to expert. Furthermore, the amount of time was mediated primarily by training and prior experience (Chi, Glaser, & Farr, 1988; Hoffman, 1992).

The theory of continued experience leading to performance expertise has been met with criticism (see for example, Ericsson et al., 1993; Ericsson & Smith, 1991; Vicente & Wang, 1998). The unconvinced authors do not believe that length of time performing in a given domain equates to expertise. For example, Doane, Pellegrino, and Klatzky (1990) found that experienced computer programmers' performance is not always superior to computer science students when it comes to programming tasks. The same findings of the less experienced performing at the same or higher level of societal experts are present in wine tasting (Gawal, 1997), investment banking (Camerer & Johnson, 1991), diagnosis of heart sounds and x-rays by general physicians (Ericsson, 2004) and auditor evaluations (Bédard & Chi, 1993).

While many researchers do not agree that longevity automatically equals expertise, they do concede that an extended period of time is required to acquire expertise. In fact, many researchers have supported the classic study of Simon and Chase (1973) through additional research into the years required for expertise in many domains. For instance, examples of domains that have found support for the Simon and Chase (1973) 10-year rule include, music (Sosniak, 1985), mathematics (Gustin, 1985), swimming (Kalinowski, 1985), evaluation of livestock (Phelps & Shanteau, 1978), medical diagnosis (Patel & Groen, 1991), and diagnosis of X-rays (Lesgold, 1984). However, the time one devotes to a performance cannot be spent unproductively if the goal is expertise. Rather, the current body of research points to a more active theory of skill acquisition and expertise, deliberate practice.

Deliberate Practice

The theoretical framework for deliberate practice was first put forth by Ericsson, Krampe, and Tesch-Römer (1993). Ericsson et al. (1993) outline the framework of deliberate practice as neither short-lived nor simple. The authors follow Simon and Chase (1973) and assert that a minimum of 10,000 hours (10 years of practice) is required for one to become an expert at a task. The methods in which deliberate practice leads to improvements and, ultimately, expertise are three-fold. First, Ericsson et al. (1993) propose the "monotonic benefits assumption". That is, the amount of time invested in domain-specific deliberate practice activities is positively correlated with the level of performance expertise attained. Next, the individual pursuing expertise must have adequate resources to optimize practice. These resources include teachers and suitable facilities. Lastly, individuals that actively pursue deliberate practice are to find it challenging and less enjoyable than other activities. The primary focus of deliberate practice is to improve performance; thus, individuals engaging in deliberate practice will not find it motivating or enjoyable.

In their seminal work, Ericsson et al. (1993) first showed promise with deliberate practice through a study of violinists in Germany. In testing the theoretical framework, Ericsson et al. (1993) created four test groups differentiated by level of skill attainment.

The groups included the best violinists in the university, good violinists, music teachers, and middle-aged professional violinists. The “best” and “good” violinists were chosen from a sampling frame established by the university instructors. The instructors selected violinists capable of soloist careers as the “best,” and students that are not quite to that level as “good”. Ericsson et al. (1993) had all participants complete recall interviews where they were asked about their practice habits growing up. Furthermore, the participants completed diaries of their practice habits. Lastly, the violinists completed surveys of their attitudes and valuation of practices.

Ericsson et al., (1993) found support for the monotonic benefits assumption. The best and professional violinists had accumulated 7,410 and 7,336 hours of deliberate practice by the time they were 18 years-old, respectively. Alternatively, the “good” violinist accumulated 5,301 hours and the music teachers had 3,420 hours. Support was found for deliberate practice requiring much effort while being less enjoyable than everyday activities. Again, this was an assumption the authors made as deliberate practice improved future performance rather than brings immediate enjoyment. Charness, Krampe, and Mayr, (1996) also found support for deliberate practice. In short, Charness et al. found the sheer amount of chess playing was unable to predict chess skill once the amount of deliberate practice was controlled for statistically. The authors found that the best chess players were those that engaged in the most deliberate practice.

Ericsson and Lehmann (1996) later developed the theoretical framework into a general theory of expertise. According to the revised theory of deliberate practice, in order to attain an expert level of performance one must practice in accordance the following four facets:

1) The task is practiced at an appropriate level of difficulty. A task that is too easy will become mundane while a task that is too difficult may not be attainable (e.g., organic chemistry is not understandable without a foundation of courses).
2) The learner must be provided with informative and immediate feedback regarding his/her performance. A significant part of feedback is the opportunity to correct actions.
3) The learner must be provided with sufficient repetition of the performance tasks.
4) The learner must be motivated to exert effort in hopes of improving future performance. If the learner is not motivated to improve, the performance will not be improved.

While the prior discussion has dealt with obtaining expertise in the performance of a skill, one need not view performance only through the lens of expertise. Ericsson et al., (1993) classify the activity performance levels as work, play, and deliberate practice. Individuals undertake each type of activity for a variety of reasons. Work is generally undertaken for the sole purpose of external rewards. The most basic external reward of work is pay for services rendered; however, competition and public performances can serve as a secondary reward system. While competitions and performances may not reap monetary gains, recognition and acceptance may serve as adequate rewards for some. Play includes activities that are enjoyable to the individual and do not have overt goals other than enjoyment (e.g., no monetary gains or trophies). The implicit goal of play is the activity in itself and the enjoyment said activity brings. An example of play can be seen at a park with friends throwing a Frisbee. The primary goal of deliberate practice is to improve performance. Ericsson et al., (1993) purport that deliberate practice “requires effort and is not inherently enjoyable” (p. 368). Rather, the motivation to practice is due to practice improving performance; thus, deliberate practice is not driven by financial or pleasure incentives. Also recall that true expertise is estimated to take approximately 10 years, or 10,000 hours, of deliberate practice. While this may be true for expert-level performance, it is the position of this paper that deliberate practice can improve one’s ability to perform a task even if expert level performance is never attained. Furthermore, task competency can be attained with much less time deliberately practicing. For this reason, the proposed study does not intend on producing expert-level performance; rather, the proposed study seeks to improve the level of competency by providing law enforcement officers with shorter amounts of deliberate practice.

Rapid Skill Acquisition

While most skill acquisition research focuses on expert-level skill attainment, some authors propose methods for learning skills at a lower level of competence. For instance, Kaufman (2014) posits that an individual can attain marked improvement in any skill utilizing the principles of deliberate practice. He calls the process “rapid skill acquisition.” According to Kaufman there are five major steps to rapid skill acquisition. First, one must decide on what skill to learn. Second, it is important to deconstruct the skill into as many subskills as possible. Third, one must learn enough about each subskill to be able to intelligently practice and self-correct if needed. Fourth, one must remove all barriers that may get in the way of practicing. Lastly, it’s imperative to practice the most important subskills for a minimum of 20 hours. Kaufman’s prescribed steps loosely follow the deliberate practice model. However, his steps are designed to work according to the “power law of practice.” The power law of practice is utilized to describe why cognitive and motor skill acquisition study participants exhibit dramatic improvements in their performance after a very short period of practice time.

Figure 6 illustrates the power law of practice as studied by Kolers (1975). Kolers had participants read pages of inverted text for comprehension. Figure 6 shows a dramatic decrease in the amount of time required to read a page of inverted text. However, once the individual has read several pages, the variation in reading time decreases dramatically and the data begin to clump (Kolers, 1975; Newell and Rosenbloom; 1981). A common example of a power law is the area of a square in relation to the length of the square’s sides. If one doubles the length of a side, the area is multiplied by a factor of four (i.e., the area functions as a power of the square’s side length). In the below example, the two variables are reading time and pages read. Reading time acts as a power to the number of pages read. The power function is -.44 resulting in a dramatic decrease in reading time as the number of pages read increases.

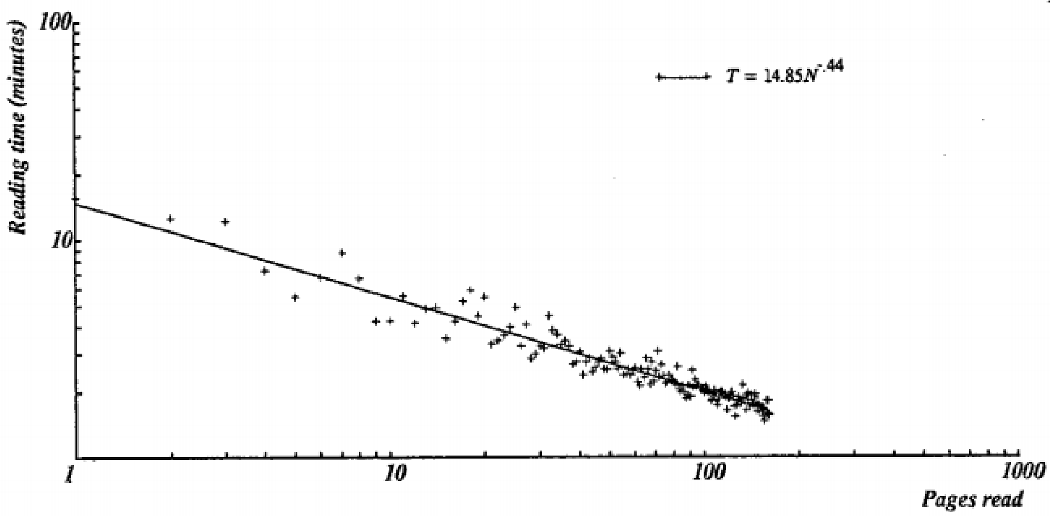

Figure 3: Learning to Read Inverted Text (Log-Log Coordinates).
Plotted from the original data for Subject HA (Kolers, 1975).

Figure 6. Example of Power Law of Practice

Methods to Improve Retention

There are three main methods people use to practice memory improvement—i.e., massed practice, spaced practice, and interleaved practice.

Massed and Spaced Practice. Many people have performed the "massed practice" method of memory retention to prepare for an exam. Students refer to massed practice as cramming for the exam. The students will go over tested information repeatedly until it is etched into their memory. However, evidence has shown that massed practice is not as effective as other methods for longer term retention (Brown, Roediger, & McDaniel, 2014). Spaced practice simply means to practice over a period of time. A comparison of massed and spaced practice was recently undertaken.

Moulton, Dubrowski, MacRae, Graham, Grober, & Reznick (2006) provided surgical residents with four lessons in microsurgery (i.e., repairing tiny vessels surgically). Half of the 38 physicians completed all four lessons in a single day. One month after completing the lessons, the physicians were given a test to measure their skill retention. The physicians that completed the lessons in a single day scored lower and 16% failed the practice surgery. In the same study, the other half of the participants practiced their microsurgery skills utilizing a spaced practice model instead of massed practice. These participants completed the four lessons one week at a time (i.e., one week per lesson). As previously discussed, the participants utilizing spaced practice significantly outperformed the massed practice participants. In another study, Litman and Davachi (2008) performed two experiments assessing the effectiveness of massed and spaced practice on the retention of paired words. The authors found that spaced practice did not enhance immediate memory. However, participants that performed spaced practice exhibited a slower rate of forgetting relative to massed practice over a 24-hour period (Brown, Roediger, & McDaniel, 2014).

Interleaved Practice. A key component of deliberate practice is to practice a task with an appropriate level of increasing difficulty. One method of increasing the difficulty of practice is to interleave the material practiced. Interleaving simply refers to mixing the sequence of material instead of practicing the material in a linear, massed fashion. Rohrer and Taylor (2007) performed an experiment of interleaved practice utilizing a 1x2 experimental design (massed practice condition and interleaved practice condition). Both groups of students were shown how to calculate the volume for four different geometric solids (i.e., wedge, spheroid, spherical cone, and half-cone). The massed practice group would practice calculating the volume of each solid one at a time. For example, the massed group would practice calculating the volume of a half-cone several times before moving to the next shape. However, the interleaved practice group would practice calculating the volume of each solid in a varied sequence (e.g., calculate one wedge, calculate one half-cone, and repeat in a varied sequence). During the practice period the massed practice group outperformed the interleaved practice group (89% correct versus 60% correct respectively). After a week, both groups were tested to measure their memory retention. The massed practice group only got approximately 20% correct. However, the interleaved practice group averaged 63% correct. So, while the practice itself was more difficult, as shown in the practice scores, the final scores were significantly better for the interleaved practice group than the massed practice group.

Memory Consolidation. Beyond interleaving practice, research has shown that the process of memory consolidation improves one's ability to recall learnt information. The process of memory consolidation involves the stabilization of a memory trace after initial acquisition—i.e., the process of transforming a new memory into a longer term memory for recall. There are two main types of memory consolidation…synaptic and systems consolidation. Synaptic consolidation occurs within a few hours of learning a new task while systems consolidation occurs over a period of weeks to years. When an individual learns a new task it first undergoes synaptic consolidation before being transplanted to a longer term memory region of the brain through systems consolidation. The first phase is known as synaptic due to the process of synaptic plasticity. Kavanau (1997) utilizes the concept of synaptic plasticity and proposes that sleep after learning behavior reinforces the memory by refreshing synapses in the brain. Research suggests that spacing learning over a 24-hour period (e.g., sleeping) decreases the rate of forgetting as compared to learning material on a single day (Litman & Davachi, 2008).

Other scholars have examined the impact of rapid eye movement (REM) sleep and memory consolidation (Vertes, 2004; Walker, Stickgold, Alsop, Gaab, Schlaug, 2005). Some have theorized REM sleep increases synaptic plasticity; therefore, memory consolidation is improved during REM sleep (Ribeiro, 1999). However, others have found that participants denied REM sleep do not show deficits in task learning. This suggests REM sleep may not improve memory consolidation beyond regular synaptic consolidation (Vertes, 2004). While synaptic memories are first stored in the hippocampal region of the brain, during systems consolidation, the memory is moved from the hippocampus to the neo-cortex for long-term memory and recollection. This process starts after the memory has been in the hippocampus for approximately a week (Frankland & Bontempi, 2005).

Based on the material discussed here, this dissertation seeks to evaluate the utility of a training method utilizing the principles of deliberate practice, interleaved practice, and memory consolidation. Specifically, the training method will provide participants with repetitive, interleaved training of increased complexity. The training will also give participants immediate feedback so they may perform corrective action. Lastly, the training method will allow for memory consolidation by requiring participants to come back a second day and execute more training before the final testing phase of the proposed dissertation. The next chapter will discuss the methodological process in detail.

III. METHODOLOGY

The current research was concerned with testing the effectiveness of a vision-training program to improve participants' ability to recognize the presence of a gun. As previously discussed at the end of the prior chapter, the proposed research is based on Ericsson's (1996) learning model of deliberate practice, Brown, Roediger, and McDaniel's (2014) discussion on information retention, memory consolidation, and the OODA loop. The current chapter provides detail on the research's methodology. This discussion includes four sections, starting with the proposed research design. The design section includes detail on the sample, intervention, and pre-/post-test. The second section will include definitions of all variables followed by the research hypotheses. Third, potential threats to validity and ways to account for the threats will be discussed. Fourth, the analytic technique utilized for the study will be briefly discussed.

Design

The proposed design of the dissertation was a 1x2 independent groups design with random assignment to test conditions (i.e., gun acquisition threat training program or non gun acquisition threat training program). Both the control and treatment group performed a pre-test to determine if the groups were comparable (O_1). Both groups received the appropriate, randomly assigned intervention. Lastly, both groups underwent a post-test to determine the effectiveness of the intervention (O_2). A graphical representation of the design is as follows.

Random	Test	O_1	$\rightarrow$	UoF* Visual Training	$\rightarrow$	O_2
Assignment	Control	O_1	$\rightarrow$	Non UoF* Visual Training	$\rightarrow$	O_2

*UoF = use of force

Figure 7. Study Design

A pre-post design with random assignment was chosen to best evaluate the effectiveness of the training program. This design was chosen for a couple reasons. First, by offering a pre- and post-test, the design allowed for a stronger examination of causality. Second, the design controlled threats to internal validity through randomly assigning participants to test conditions, which will be discussed in a subsequent section (Shadish, Cook, & Campbell, 2002).

Sample

A convenience sampling strategy was employed. As a non-probabilistic sampling strategy, the generalizability of findings would be weak. As this study was a proof of concept, generalizability was not a concern. However, the random assignment to test conditions should have removed the threats to internal validity and provided a solid test of efficacy regarding the training program. The researcher visited several undergraduate criminal justice courses at Texas State University to solicit participation. Many of the instructors of record provided extra credit opportunities for students. Additionally, the researcher received approval from the IRB to provide a small monetary compensation as an incentive for participation ($10; IRB Approval # 2016O7324). Table 1 presents the descriptive data for the sample.

Table 1. Sample Descriptive Data

	Control Group	Test Group
Sample Size	N = 44	N = 43
Age	22.70 (5.83)	22.00 (3.63)
Sex		
Male	26	24
Female	18	19
Race/Ethnicity		
African American	N = 8	N = 9
Asian	N = 2	N = 2
Caucasian	N = 18	N = 18
Hispanic	N = 14	N = 15
Other	N = 2	--
Prior LE	N = 1	N = 1
Prior Military	N = 1	N = 1

*Standard Deviation in Parentheses

There were originally 92 participants; however, only 87 were able to be included in the analysis. Two participants did not wear contacts on the second day. They were unable to complete the post-test, as they could not see well enough to perform. Furthermore, these two individuals squinted their eyes trying to see rendering the eye tracker useless…for these reasons they were excluded. There were equipment malfunctions with the other three participants that resulted in no audio or video being recorded. Table 1 illustrates the descriptive data for the sample. As seen here, the random assignment to conditions resulted in both the control and test group being comparable. The average age was 22 years old. The majority of participants were Caucasian.

After solicitation, participants signed up for specific time blocks via an online sign-up program "Jooners". Jooners allowed volunteer participants to reserve available time blocks that fit their schedule.

Videos

The videos used for both the pre-test and the post-test were provided by MILO Training Systems (previously IES Systems). MILO Training produces advanced use of force simulators to train law enforcement officers and military personnel on the proper application of force. MILO utilizes several different systems; however, the premise is the same. The system projects a professionally made video depicting a use of force encounter on a screen. The trainee is armed with a variety of infrared weapons (e.g., pistol, rifle, pepper spray, conducted energy device, flashlight) that can interact with the MILO video environment. The trainee then works through the scenario presented in the video. The scenarios are dynamic—i.e., the camera point of view is that of a responding officer. There are approximately 800 MILO videos that challenge trainees to detect viable threats and perform the appropriate level of force.

MILO provided the researcher with a company laptop pre-loaded with all the available training videos. Four videos were chosen for the pre-test. There were two videos where the perpetrator was armed with a gun and two videos where the perpetrator was "armed" with something other than a gun (e.g., cell phone). The study participant was required to virtually "move" through the scenario and make a lethal force decision based on the presence of a weapon. As a large portion of the sample was college students without formal force continuum training (two participant had prior law enforcement training and two participants were prior military), the participants were instructed to verbalize the presence of a weapon if present. For a detailed description of all four pre-test videos, please see Appendix A.

Vision Tracker

Both the pre- and post-test utilized vision tracking technology to measure the study participants' performance during testing. The participant wore a backpack containing the computer. The computer was connected to the head-mounted vision tracking hardware. The head-mounted hardware consisted of a microphone, an infrared LED light bulb, and two cameras. One of the cameras and the LED bulb faced the user's eye. The camera was focused on the user's pupil and was capable of capturing infrared light. This camera recorded the pupil's directional movements and was known as the "Eye Camera". The LED light put a small reflection on the face of the cornea to aid in the tracking process. The other camera faced away from the user. This camera was referred to as the "Scene Camera". The scene camera was static and captured everything in front of the user—i.e., wherever the user's head was oriented. Both camera feeds were processed by the laptop computer worn on the participants' back.

The software utilized a series of regression analyses to interpolate where the user's gaze was fixated based on pre-established calibration points. The LED light's reflection was used to determine the breadth of eye movement in conjunction with the centroid of the eye's pupil. This information was also utilized by the software to give the most accurate visual track. The result was a superimposed moving dot corresponding to the position where the individual's gaze was focused.

Pre-Test

A pre-test was utilized to ensure both groups were equivalent in terms of their visual performance. Furthermore, the pre-test established a baseline performance level. The established baseline was compared to the post-test to assess the effectiveness of the intervention. The pre-test consisted of four use of force videos. The videos and related technology used during the post-test will be discussed next.

Post-Test

The post-test utilized the same type of videos used during the pre-test—i.e., videos provided by MILO Training. While the pre-test only used four videos, 10 videos were chosen for the post-test. There were 4 videos where the perpetrator was armed with a gun, 4 videos where the perpetrator was armed with something other than a gun, and 2 videos where no item was present. Just like the pre-test, the study participant was required to virtually “move” through the scenario and verbalize the presence of a weapon if applicable. For a detailed description of all 10 post-test videos, please see Appendix B.

Intervention

Both the control and test group received an intervention specific to their condition. The intervention in this experiment was a computer-based training program. The test group received a program that trained the participant to look at a suspect’s hands for a gun. This trained the participant to visually scan the suspect for a gun. The control group received a program unrelated to use of force training. This was to ensure the use of force training program had an actual effect beyond the presence of any training. The computer-based training program will now be discussed in detail.

Programming Information. The program was built in E-Prime 2.0. The E-Prime Suite utilized a custom programming language designed around the established visual basic language. The E-Prime Suite was composed of several independent programs that work in conjunction with each other. The two main programs used during the project were E-Studio and E- Run. Each will be briefly described below.

E-Studio was where the raw program was built. The parameters for each training level were established in this program. That is, the images and video to be used were selected, the correct responses for each image and video were determined, and feedback was written for each image and video based on the user's input. While one could start each training level directly from E-Studio, E-Run streamlines the process.

E-Run allowed each training level to be exported from E-Studio. The participant was not required to navigate through E-Studio to start the training. Rather, the participant simply selected which training level to begin and the appropriate training level was brought up automatically. This process was similar to opening a PowerPoint presentation. In PowerPoint the user has the option of opening the root PowerPoint program and starting the presentation from PowerPoint. Alternatively, the user can save the presentation as a PowerPoint show and start the presentation directly without opening PowerPoint. This process not only streamlines the manner in which the presentation is started, it also keeps the user from altering any aspect of the presentation. E-Run performed this same procedure for the E-Studio training program. The study participant selected the appropriate training level and it automatically started. This kept the study participant from accidently changing any of the pre-established parameters in E-Studio.

E-Studio only worked on Windows operating systems and required a moderate amount of computing power. For this reason, the researcher used an established computer lab on the campus of a large southwestern university.

The training program consisted of three distinct levels. According to Ericsson (1996), the key to becoming an expert in any given task was repetition, increasing difficulty within the task, and ample feedback. Therefore, the program consisted of three levels of increasing difficulty with feedback. Each of the levels will be discussed in further detail.

Description of Levels. There were three distinct training levels. Both the control and test group used identical training programs. However, the control group did not receive any images related to weapons or use of force. Rather, the control group determined if letters of the alphabet were consonants or vowels. The manner in which letters were presented changed in the same manner the treatment group's levels changed. The treatment group was shown images and video relating to the identification of weapons and the application of deadly force. Each training level will be described separately. Furthermore, image examples will be provided.

Level 1. The first level contained the simplest images. The purpose of the first level was two-fold. First, this level acclimated the study participants to the training program. In other words, the participant became familiar with which keyboard buttons were utilized and how to navigate through the program. Second, the simple images prepared the participant for the more complex images according to the deliberate practice model.

The control group participants were shown a single letter of the alphabet. The participants pressed the number 1 if the letter was a consonant and the number 3 if the letter was a vowel. Figure 8 illustrates this level for the control group.

Figure 8. Control Group Level 1 Example

The test group participants were shown an image of a weapon (i.e., gun) or a non-weapon (e.g., hairbrush). The participants pressed the number 1 if a weapon was displayed and the number 3 if a non-weapon was displayed. Figure 8 provides example images for this level. As seen in Figure 9, the images presented are clearly a gun or not a gun. Ambiguous objects were not utilized.

Figure 9. Test Group Level 1 Example

Both groups were immediately informed if they made a correct selection or not. A screenshot of the image was shown to the user as well as a statement at the top of the screen that read *Correct* or Incorrect. This process allowed the participants an opportunity to see how a mistake was made and how to correct for the mistake in the future. The user then pressed the space bar to move to the next image. There were 30 images presented in the first level for both the control and the test group.

Level 2. The second level added complexity to the first level. In the second level, the control group was shown an image with sixteen symbols (i.e., a 4x4 table). Only one of the symbols was a letter of the alphabet. This style of image was chosen as it required the participant to visually scan multiple aspects of the image to locate the vowel—if a vowel was present. The user was required to find the letter of the alphabet and select the number 3 for a consonant and the number 1 for a vowel. Figure 9 shows example media for this level. In Figure 10, the table containing the vowel is on the left, while the table not containing a vowel is on the right.

α	W	¥	⏎
E	≤	đ	β
G	∑	Ø	^
Ψ	❖	Ω	w

✔	W	N	Ŧ
α	₠	§	€
ϖ	Ψ	K	v
#	✪	∑	@

Figure 10. Control Group Level 2 Example

The test group was required to determine if an object was a gun or non-gun while in the hands of an individual. The object was in either the right or left hand of the individual, and the individual was standing at various angles in relation to the camera. The participant selected the number 1 if a gun was present and the number 3 if a non-gun was present. Figure 11 illustrates example media for this level. As seen in Figure 11, the model is a white male, wearing a blue shirt, in front of a nondescript background. All of the models utilized in the creation of the media reflect this model. Namely, all models were white males, wore blue shirts, and stood in front of a blank background. This was to remove additional correlates from the image. This process meant that the only thing participants would focus on is the location and identification of the item. The image on the left of Figure 11 contains the gun. As seen in the left image, the gun is partially obstructed by the model's hand. However, features of the gun are still visible. This requires the participant to utilize accurate foveal vision to pull the necessary data needed to make a correct decision.

Figure 11. Test Group Level 2 Example

There were also 30 images presented in the second level. By showing the images from various non-sequential angles, this level contained interleaved images to improve retention (Brown, Roediger, & McDaniel, 2014). In other words, the images within the level were randomized so there was a non-discernable order. This process was more mentally taxing on the participant.

Level 3. The third level was the most complex. This level increased the complexity of the second level by adding movement. This was accomplished by creating videos of the second level. For the control group, the various sixteen symbol images were printed and placed on individual boards. The board was placed on a table. The video started where the participant could not see the content of the board. The camera passed over the board and participants pressed the number 1 when they determined if there was a vowel present on the board. Alternatively, participants pressed the number 3 if there was not a vowel present. The test group was shown video of the individual turning around with an object in either their right or left hand. The videos were shot in the same location with the same models. They only differed from Level 2 by having movement added. As before, the user pressed the number 1 if a gun was present and the number 3 if a non-gun was present. The videos in Level III were also interleaved according to the procedures discussed by Brown et al. (2014). For both the control and test group, the environment remained neutral. In other words, there was not any auxiliary imagery competing with the training material. The background and surrounded area were cleared of any tertiary objects that could draw the user's attention. There were 30 videos presented in the third level.

Level IV was created to give participants additional training at the beginning of the second day of participation. There was a fourth level that was structurally the same as Level III. Level IV contained the same media as Level III. However, the videos used in

Level IV were not recycled from Level III. A Level IV was created for both the control and the test group. There were 50 videos presented in the fourth level.

Procedure

Recall the discussion on memory retrieval/consolidation presented by Brown, Roediger, and McDaniel (2014). According to the authors, memory retrieval was most effective if sessions are spaced out. This process required an individual to utilize cognitive effort and sped up one's ability to recall information. For this reason, study participants were asked to participate for two days. On the first day, the participant completed the pre-test and Levels I, II, and III of the training program. The next day, the participant completed Level IV and the post-test. This process required participants to practice memory retrieval during the completion of Level IV prior to the post-test. A more detailed look at both days' procedures follows.

Day 1. The first day of a participant's involvement began with reading and signing an IRB approved consent form (see Appendix E for a copy of the consent form). The participant then completed a short form collecting basic demographic information (see Appendix C for an example of the demographic collection tool). The demographic information included age, sex, race, prior LE experience, and prior military experience. The participant was given an instructional sheet detailing the process of signing into the training program (see Appendix D for an example of the instructional sheet). The instruction sheet was numbered. Even numbers denoted assignment to the test condition and odd numbers denoted assignment to the control condition. The participant utilized this number to sign into the program. The participant number also linked the individual to their demographic information for later analysis.

Once the participant read the instruction sheet he/she began. The participant had the vision tracker calibrated to their individual physiology. The participant was positioned in front of the projector to undergo the pre-testing process. At this point the research moved to the MILO Training laptop to begin the pre-test. The participant worked through the four pre-test videos. The participant was required to vocalize the moment he/she is aware of a threat by stating “gun” or “no-gun”. According to Brinkman (1993), Ericsson and Simon (1993), and Green (1995) this vocal latency process involved higher order skills. This process also triangulated the moment the user vocalizes what is seen and what the vision tracker showed.

Once the participant had completed all four pre-test videos, the researcher removed the vision tracker from the participant and showed them to the room where the training program was located. Once in the training room, the participant was shown to a desktop computer equipped with the training program. Each training level had an individual file placed on the computer’s desktop. The participant double clicked the appropriate file based on their assigned condition to begin. The program asked for the participant number. The participant entered their assigned participant number. The program then displayed a description of the specific level and the instructions needed to complete the level. The participant was instructed to make his/her decision as quickly as possible. The participant then completed Level I, Level II, and Level III of the training. It took participants approximately 15 minutes to complete the procedure on Day 1.

Day 2. Once the participant arrived at the test location for Day 2, he/she completed Level IV of the training program. This process was in place to let participants perform information retrieval prior to the final test. Once the participant completed the level, he/she completed the post-test. The vision tracker utilized in the post-test limited the number of participants to five an hour. This equated to two minutes to set up and remove the equipment and 10 minutes for the actual post-test. Once the participant had been calibrated to the vision tracker, the post-test began. First, the researcher reiterated to the participant that they were to verbalize the presence of a gun or non-gun (i.e., “gun” or “no-gun”). Once the participant acknowledged that he/she understood, the researcher moved to the laptop equipped with the MILO use of force videos. Each video began with a short countdown. Once the video began the participant worked through the scenario until each scenario was complete. After the participant had completed all post-test scenarios the vision tracker was again calibrated and removed.

Data Collection

The primary data were collected during the pre- and post-test phases of the study. These data were collected via the previously mentioned vision tracking system. The researcher extracted the video from the vision tracking software. The video was loaded into Camtasia, a video editing program. Camtasia provided the researcher the ability to view video frame-by-frame. Both vision tracker cameras record at 30 frames per second (fps). Therefore, all data gathered from the post-test were restricted to a specificity of 1/30th of a second.

Variables

A total of 12 root variables were collected. Six of the defined variables represent a collection of variables for each image or video presented in the training program. The aggregate total number of variables is 48. Each root variable was defined as below.

Age – *Age* identifies the whole number years of age for the participant. *Age* is a ratio level variable.

Sex – *Sex* identifies the biological distinction of the participant as male or female. *Sex* is a dichotomous, nominal level variable (i.e., male/female).

Race/Ethnicity – *Race/Ethnicity* identifies the racial and/or ethnic group the individual participant identifies with—i.e., African American, Asian, Caucasian, Hispanic, Other. *Race/Ethnicity* is a nominal level variable.

Prior_law_enforcement – *Prior_law_enforcement* identifies whether or not the individual participant has any prior law enforcement experience. *Prior_law_enforcement* is a dichotomous, nominal level variable.

Prior_military – *Prior_military* identifies whether or not the individual participant has any prior military experience. *Prior_military* is a dichotomous, nominal level variable.

Condition – *Condition* identifies whether the individual participant was randomly assigned to the control or test group. *Condition* is a dichotomous, nominal level variable.

Pre_dec_time – *Pre_dec_time* identifies the length of time, measured in 30 fps, that passes from the moment the pre-test video begins until the individual participant verbalizes the presence or absence of a weapon. There are 4 *pre_dec_time* variables projected, one for each pre-test video. The variable names are pre_dec_time_1 through post_dec_time_4. *Pre_dec_time* is a ratio level variable. A high degree of interrater reliability was found for this variable (ICC = .96, 95% CI [.92, .98], $p < .001$).

Pre_test_accuracy – *Pre_test_accuracy* identifies whether or not the individual participant makes a correct decision for the presented pre-test video. There are 4 *pre_test_accuracy* variables projected, one for each pre-test video. The variable names are pre_test_accuracy_1 through pre_test_accuracy_4. *Pre_test_accuracy* is a dichotomous, nominal variable. A high degree of interrater reliability was found for this variable (ICC = .98, 95% CI [.96, .99], $p < .001$).

Pre_test_fixate – *Pre_test_fixate* identifies whether or not the individual participant looks directly at the weapon or non-weapon prior to making their decision to apply force or not. There are 4 *pre_test_gaze* variables projected, one for each post-test video. The variable names are pre_test_gaze_1 through pre_test_gaze_4. *Pre_test_gaze* is a dichotomous, nominal variable. A high degree of interrater reliability was found for this variable (ICC = .93, 95% CI [.89, .94], $p < .001$).

Post_dec_time – *Post_dec_time* identifies the length of time, measured in 30 fps that passes from the moment the post-test video begins until the individual participant verbalizes the presence or absence of a weapon. There are 10 *post_dec_time* variables projected, one for each post-test video. The variable names are post_dec_time_1 through post_dec_time_10. *Post_dec_time* is a ratio level variable. A high degree of interrater reliability was found for this variable (ICC = .95, 95% CI [.90, .97], $p < .001$).

Post_test_accuracy – *Post_test_accuracy* identifies whether or not the individual participant makes a correct decision for the presented post-test video. There are 10 *post_test_accuracy* variables projected, one for each post-test video. The variable names are post_test_accuracy_1 through post_test_accuracy_10. *Post_test_accuracy* is a dichotomous, nominal variable. A high degree of interrater reliability was found for this variable (ICC = .98, 95% CI [.97, .99], $p < .001$).

Post_test_fixate – *Post_test_fixate* identifies whether or not the individual participant looks at the weapon or non-weapon prior to making their decision to apply force or not. There are 10 *post_test_gaze* variables projected, one for each post-test video. The variable names are post_test_gaze_1 through post_test_gaze_10. *Post_test_gaze* is a dichotomous, nominal variable. A high degree of interrater reliability was found for this variable (ICC = .94, 95% CI [.90, .96], $p < .001$).

Coding

The author was the primary data coder. He coded 100% of the pre- and post-test video data. A second coder performed an inter-rater reliability check by coding 10% of the recorded video. A coding instrument and instructions was provided to the reliability coder to assist in consistency. Intra-Class Correlation Coefficient was utilized to assess inter-rater reliability. These statistics were presented with the definitions of the variables above.

Possible Validity Issues and Steps to Correct

As with all data collection, the proposed research was not without some potential validity threats. By utilizing randomization in the research design, both the control and test group should have been equally impacted by potential threats to internal validity. As the purpose of the proposed study was to provide a proof of concept test, external validity concerns were not incorporated. It was not the researcher's intention for these initial results to be generalizable to a large population at this phase of the training program development. Future iterations of the training program will take external validity into consideration. The following section will discuss the common threats to internal validity and what measures were taken to pacify the identified threats.

1) History. The participants took part in the research over a two-day period. It was possible that an event may occur between the first and second phase of study. A history threat was difficult to account for. However, the research took place on back-to-back days. The short time between the first and second phase of study, and the random assignment of study participants, should have limited possible history threats to validity.

2) Testing. Both the pre-test and post-test utilized use of force video. While a testing effect may have been present, the impact of the effect was assessed via the control group. Furthermore, the pre-test only consisted of four videos to reduce the exposure of force scenarios for the control group.

3) Selection. Group equivalency was vital was between group differences were examined. The research design mitigated a potential selection threat through randomized assignment to test conditions. Demographic data and the pre-test were utilized to assess group equivalency.

4) Mortality. Another potential threat to the two-day study was participant mortality. The online signup sheet (i.e., Jooners.com) provided participants with reminder emails and/or texts. The researcher also sent a reminder email to participants the day before their allotted time. At the end of their first day, the researcher expressed gratitude for their participation and reminded them to come back the second day. The activities planned for the second day were reiterated.

5) Maturation. Participants may have changed over the course of an experiment that has more than one data collection period. However, maturation was not considered a legitimate threat to the proposed study. Participants performed phase 2 of the study the

day after phase 1. This short period of time should not have allowed for maturation of the participants.

6) Instrumentation. Participants were not exposed to evolving instrumentation. The intervention utilized for both the test and control group was set prior to the beginning of the experiment. Each participant completed the intervention in the same order.

7) Regression. Regression to the mean is a threat when performance measures are utilized. Some participants may have initially performed at an elevated level while others performed at a below than average level. It is possible that some participants may have then performed at an equal level as the mean of the sample. This process can appear to show changes in the participants' performance due to the intervention rather than regression to the mean. The randomized assignment to test conditions should have eliminated potential regression effects.

Research Hypotheses

There are three hypotheses that the study will address. Each hypothesis was developed according to the three research questions.

H_1: Deliberate practice on the visual acquisition of weapon presence will improve the speed of use of force decisions over deliberate practice in unrelated visual searches.

H_2: Deliberate practice on the visual acquisition of weapon presence will improve the accuracy of use of force decisions over deliberate practice in unrelated visual searches.

H_3: Deliberate practice on the visual acquisition of weapon presence will decrease the amount of time needed to fixate on the item over deliberate practice in unrelated visual searches.

Analytic Technique

Bayesian analytic techniques were applied to the data described in this chapter. This section will give an introductory description of the statistical programs utilized to conduct the analysis. The upcoming chapter will present the results.

Reallocation of Credibility. A key concept in Bayesian inference is the reallocation of credibility. Credibility is synonymous with probability in this discussion. As such, credibility ranges from 0 to 1 (Krushcke, 2014). When solving a problem, several outcomes may be credible, and as data are collected some outcomes become non-credible. The credibility of the removed outcomes is then reallocated to other possible outcomes. For example, if one was to step outside in the morning and the sidewalk was wet, there are several credible reasons for the wet sidewalk. Someone may have dropped a drink, the sprinkler system may have gone off, the neighbor's kids may have thrown water balloons at each other, or it may have rained. As the individual walks down the sidewalk, the concrete is still wet without any dry spots. The possibility of a spilled drink or water balloons is no longer credible. Therefore, the credibility assigned to those potential outcomes are reallocated to the remaining credible outcomes. The further the individual walks, he or she notices the adjacent property, street, and cars are also wet. A sprinkler system is no longer credible, so the most credible possibility is that it rained. This process is graphically represented in Figure 12, below.

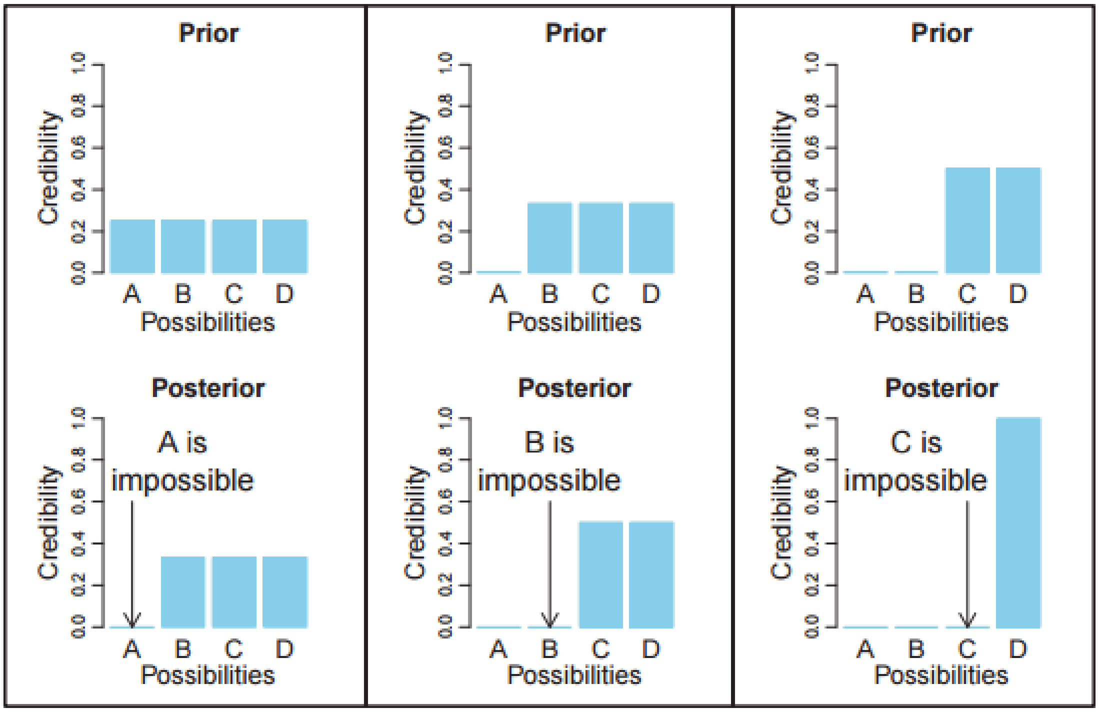

Figure 12. Reallocation of Credibility Example (Kruschke, 2014)

At the top left of the graphic one can see that all four possibilities are credible. Directly below, in the bottom left portion of the graphic, one of the possibilities was found to no longer be credible. Therefore, the credibility was reassigned to the remaining three possibilities. This process is repeated in the middle and the right of the graphic. As possibilities are found to be non-credible the only remaining possibility is assigned all the credibility. In this case, it rained while the individual was inside.

This example was a simple illustration of the credibility reallocation process. As seen in the above graphic, our prior knowledge is influence by observations in order to reallocate credible outcomes. This concept is core in Bayesian analysis. The prior knowledge can be determined by previous research into the topic. Or, if no research has been conducted related to the topic, an uninformed prior can be assigned. Both of these types of priors are illustrated in the previous Figure 8. The top left graphic represents an uninformed prior. In this instance, all things are possible. As information is gathered, future assessments can utilize an informed prior based on what was learned. This is illustrated by the top right graphic. This graphic shows that only C and D are possible outcomes based on two separate instances of reallocating credibility. Once more, information is applied to the informed prior and credibility is reallocated to the outcome D. Figure 13 shows this process with a distribution of values.

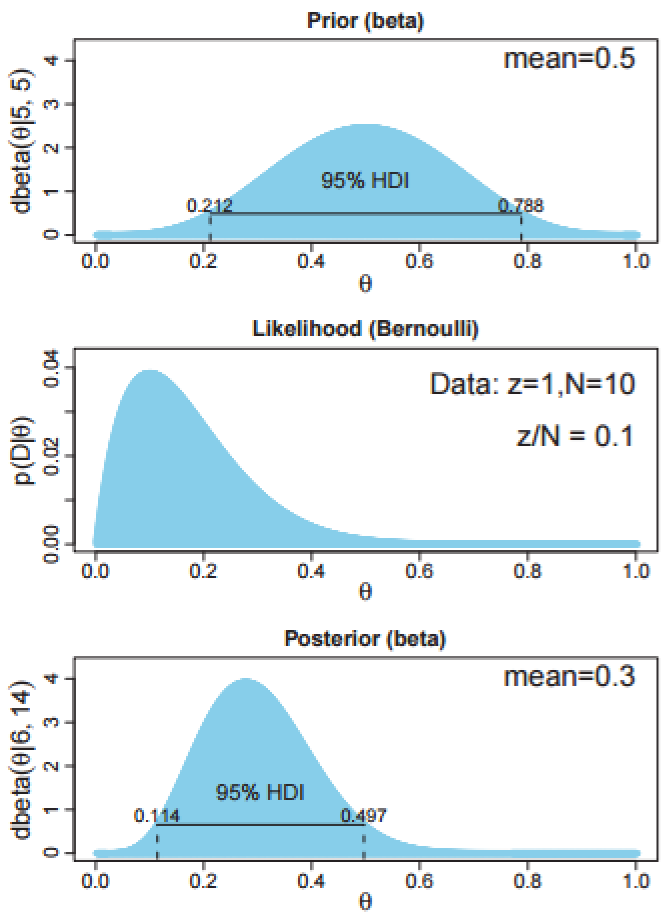

Figure 13. Data Applied to Prior Distribution

The prior knowledge in Figure 13 is a normal distribution. For the purposes of this example, the data will be employee performance scores on a 10-point Likert scale. In the past 10-years, employees were equally above and below a 50% score on the scale. For this reason, the prior knowledge would reflect the past performances. The middle graphic illustrates employee performance scores after a new manager was hired. The bottom graphic illustrates credible outcomes based on the new data being applied to the prior knowledge. Had there been no prior knowledge of employee performance, the posterior distribution would have reflected the data presented in the middle graphic more closely. However, in this example, the prior knowledge of employee performance results in a shift of credibility away from a score of 50%. An explanation of this fictitious data could be the new manager evaluated performance much more harshly. For this reason, the most credible outcomes are somewhere between the prior distribution and the new data. The new posterior distribution will now become the prior distribution for future analysis. If performance scores stay at the same level, the new posterior outcomes would have credibility reallocated to a lower level.

MCMC Simulations. Bayesian analytic techniques have been in use for several decades. However, the process of calculating the posterior outcome requires a great deal of computing power. For this reason, only recently have computers advanced to the level that Bayesian techniques can be utilized to efficiently calculate posterior distributions for analysis beyond the simplest (within the last 30 years; Kruschke, 2014). The Bayesian programs utilized in this research use Markov chain Monte Carlo (MCMC) simulations to produce the posterior parameter distribution. MCMC simulations generate parameter estimations from a target distribution. This process results in thousands of parameter estimates. In a frequentist analysis, the researcher performs the analysis to represent the data. The researcher is presented with a single parameter estimate and a confidence interval. The researcher then concludes that the parameter estimate falls somewhere within the confidence interval. The MCMC process essentially performs this task thousands of times (actual rate determined by the researcher). Each parameter estimate falls within a plausible distribution. These estimates form a distribution of credible parameter estimates. Figure 14 illustrates this process below.

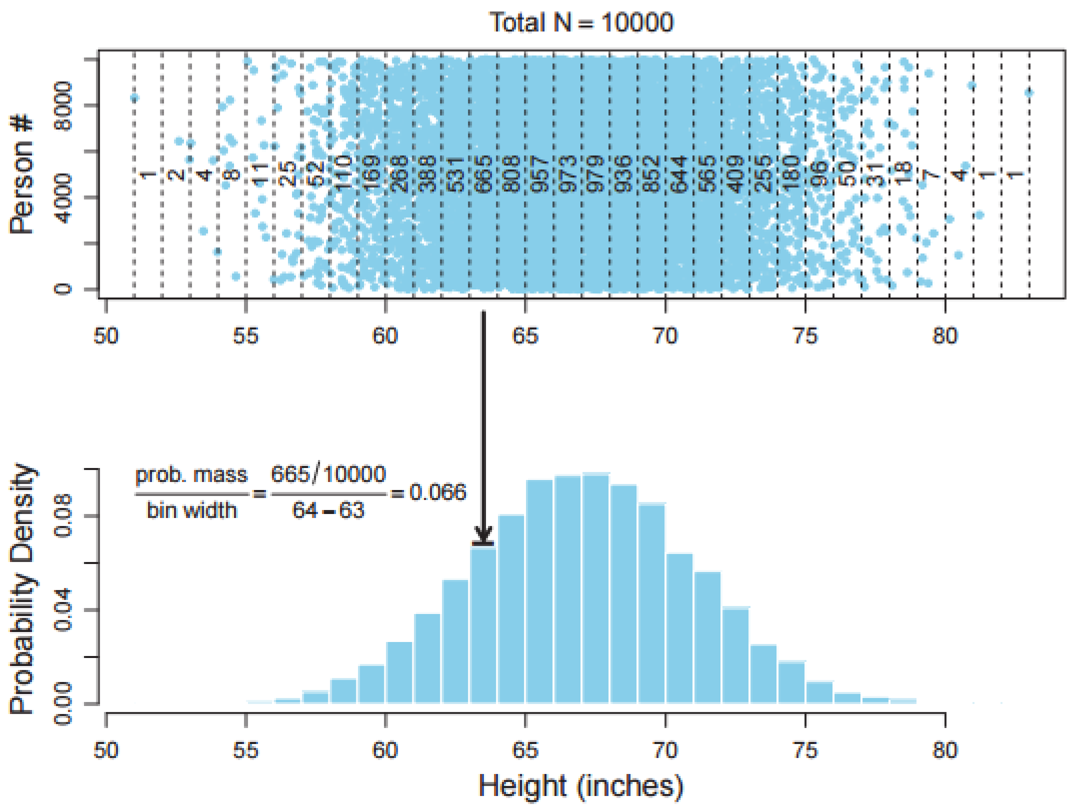

Figure 14. Probability Density

As seen in Figure 14, the top graphic is a scatterplot of credible parameter estimates. In the figure, the majority of the observed values in the scatterplot fall between 60 and 75. The scatterplot is transformed into a histogram to illustrate the distribution within the program. The histogram now resembles a distribution most are familiar with. However, rather than this being a distribution of observed values (i.e., data from a dataset), these values are credible parameter estimations based on the observed values and prior knowledge. A 95% interval of credible estimates can be created for each posterior distribution. This process results in what frequentist statistics put forth as a confidence interval.

Benefits of Bayesian Techniques. While the above discussion highlighted Bayesian techniques ability to improve specificity with parameter estimations by simulating thousands of credible values, there are additional benefits to performing Bayesian analyses. First, the MCMC process normalizes the distribution of data through simulating thousands of parameter estimations. During the MCMC process, the simulation bunches credible estimates around the central tendency of the observed data. This process normalizes the parameter estimates even if the data are not normally distributed. The program treats values closer to the central tendency as more credible. This leads to the second benefit of Bayesian analysis – processing outliers. Outlier data are problematic in frequentist analysis as these values can skew the single parameter estimate. This requires the researcher to decide what to do about extreme cases (e.g., exclude the values). However, in a Bayesian analysis these values are left alone. The analysis views these values as credible. This is because they occurred in real life. While they are a credible value, the program will assign less credibility to the outlier value and more credibility to the data closer to the central tendency. This allows the value to contribute to the overall process, albeit in a reduced capacity.

A third benefit of Bayesian analysis is the analysis does not rely on p-values. There has been a big push in the last few years for researchers to avoid the pitfalls of p-values. First, p-values cannot inform the reader the probability that a result occurred by chance. P-values cannot inform the reader what the size of any observed effect is.

Furthermore, a p-value is heavily dependent on a large enough sample size. For these reasons, many scholars are advocating for the use of effect sizes, confidence intervals, and estimations of parameters (i.e., Bayesian techniques) to replace null hypothesis significance testing (see Anderson, Burnham, & Thompson, 2000; Wagenmakers, 2007).

In Bayesian analysis, data are applied to a prior distribution to test for a treatment effect. If the collected data are different enough from the prior, credible distribution, what is considered credible is reallocated based on the collected data. This process of reallocation results in credible parameter estimates. Normally, a non-committal prior distribution is utilized based on the expected distribution. However, the design of the current study allowed for a different approach to be taken. Specifically, the pre-test data were used to establish a prior distribution. This process allowed for informed priors to be applied to the post-test data analysis. The next chapter will present the research findings.

IV. RESULTS

The current chapter details the study results. Before data are presented, it is imperative that the variables utilized for the analyses be thoroughly described. For this reason, each variable will be defined and descriptive data regarding the variables will be presented. Furthermore, the analysis will follow the Bayesian framework discussed in the prior section. To ensure readers understand how to read the graphical representation of the data, an example graphic will be utilized to explain the nuances of interpretation. Following the example graphic, the results from the study will be presented.

For consistency, the results are presented in the order of the study's research questions and hypotheses. As such, the speed of the decision making process will be presented first. Next, data regarding participants' ability to visually locate the item in question will be presented. Lastly, findings regarding participant decision making accuracy will be presented. Following the research questions and hypotheses, findings that were not anticipated will be presented.

Primary Variables

The data collected from each video were utilized to create a series of variables. Summing the requisite values for each participant created each variable. There is a variable for the number of errors, speed of decision, and speed of item fixation for both the pre-test and post-test. Descriptive data are presented next for each variable.

Error Variable. Each video's accuracy values were summed to create the error variable. Participants were told to verbally identify the presence of a gun in each video. Videos either contained a gun in the individual's hand, an item other than a gun in the individual's hand, or no item in the individual's hand. Each time a participant incorrectly identified an object he/she was given a score of 1. Alternatively, if a participant did not verbally identify a gun if present, he/she was given a score of 1 for the error. Participants could potentially make more than one error. As such, the score is cumulative for each participant. This cumulative score for each participants was combined with the scores from participants within the same test or control condition to create the overall error variable. Descriptive data regarding the pre-test and post-test accuracy variable are presented next.

Pre-test Error Variable. The pre-test consisted of four videos. The civilian in Video #1 was holding a wallet in her hand. Participants could make an error by incorrectly identifying the wallet as a gun. Both civilians in Video #2 and Video #3 were holding a gun in their hand. Participants could make an error by not verbally identifying the gun. The civilian in Video #4 had a gun in his lap, but not in his hand. The only part of the gun in Video #4 that was visible was the end of the handgrip. This was a difficult item to locate; therefore, many participants were unable to correctly identify the gun. There were 81 total participants with error data to present (Control n = 42; Test n = 39). Participants were excluded because there was no audio recorded during their pre-test session (two were excluded from the control group while four were excluded from the test group).

Table 2. Pre-Test Accuracy Variable

	Control Group	Test Group
N	42	39
Mean	.37	.30
SD	.62	.47
Min	.00	.00
Max	2.00	1.00

Post-test Accuracy Variable. The post-test consisted of ten videos. The civilians in Videos #1, #2, #7, and #10 were holding objects other than a gun in their hand. Participants could make an error by incorrectly identifying the object in the civilian's hand or environment as a gun. Videos #5 and #9 did not contain any objects in the civilians' hands. Participants could make an error by incorrectly identifying any object in the environment as a gun. Videos #3, #4, #6, and #8 contained civilians with guns in their hands. Participants could make an error by failing to identify the object as a gun. There were 83 total participants with data available to create this variable (Control n = 43; Test n = 40). Participants were excluded because there was an equipment malfunction with the microphone resulting in no error data being recorded during their post-test session (one was excluded from the control group while three were excluded from the test group).

Table 3. Post-Test Accuracy Variable

	Control Group	Test Group
N	43	40
Mean	1.39	.50
SD	1.17	.59
Min	.00	.00
Max	5.00	2.00

Decision Speed Variable. Each video was coded to capture the speed of the participants' decisions. The vision tracker captured video at 30 frames per second (fps). The number of frames was counted from the moment the item in question was visible until the participant verbalized the presence of the gun. The number of frames was converted to seconds for easier comprehension. This was accomplished by dividing the number of frames needed to verbalize the decision by 30 (i.e., decision in seconds = frames/30). Unless the participants made an error on a non-gun video, the decision was only verbalized when a gun was present. As such, the decision speed variable was composed of videos where a gun was present.

Pre-test Decision Speed Variable. There were three pre-test videos with a gun present. Video #2 and Video #3 presented the participants with a civilian holding a gun. Video #4 had a gun present in the scenario. However, the gun was lying in the civilian's lap and only the bottom of the handle was present. The difficulty of identifying a handle of a pistol pushed the decision time out several seconds. For this reason, video #4 decision times were removed from the variable to keep from skewing the average decision time. The control group contained 42 participants. The test group contained 39 participants. There were microphone malfunctions that resulted in the exclusion of participants (two were excluded from the control group and four were excluded from the test group). This was because the decision could not be heard to capture the needed data.

Table 4. Pre-Test Decision Speed Variable

	Control Group	Test Group
N	42	39
Mean	.64	.68
SD	.24	.22
Min	.02	.23
Max	1.07	1.13

Post-test Decision Speed Variable. The post-test had four videos with a gun present in the civilian's hand. The four videos made up the decision speed variable for the post-test. Descriptive data are presented in Table 5. The control group was made up of 43 participants. The test group was made up of 40 participants. Again, there were microphone malfunctions that resulted in the exclusion of participants (one was excluded from the control group and four were excluded from the test group).

Table 5. Post-Test Decision Speed Variable

	Control Group	Test Group
N	43	40
Mean	.40	.36
SD	.18	.17
Min	.07	.03
Max	.74	.72

Fixation Variable. The fixation variable measures the amount of time, in seconds, it took a participant to visually fixate on an item once it was present. As with the decision variable, these data were calculated by counting the number of frames that passed from the time the weapon is present. However, in the fixation variable, the counting stopped once the participant visually fixated on the item. A visual fixation occurred when the individual's superimposed gaze point stopped on an item for 2 or more frames.

Pre-test Fixation Variable. The pre-test fixation variable consisted of three videos. For Video #1, the time required by participates to fixate on the civilian's wallet was captured. For Video #2 and Video #3 the time required by participants to fixate on the civilian's gun was captured. As previously discussed, the gun in Video #4 was only partially visible (i.e., only the handle was visible in the civilian's lap). For this reason, most participants never visually fixated on the item. Video #4 was excluded from the pre-test fixation variable. Participants were excluded if none of their fixation data were able to be used (e.g., squinted the entire time, knocked the headgear out of alignment). Four participants were excluded from the control group while none were excluded from the test group.

Table 6. Pre-Test Fixation Variable

	Control Group	Test Group
N	40	43
Mean	.55	.54
SD	.12	.10
Min	.28	.26
Max	.77	.78

Post-test Fixation Variable. The post-test fixation variable consists of visual fixation data from eight different videos. The civilians in Videos #1, #2, #7, and #10 were holding objects other than a gun in their hand and contained fixation data. Videos #3, #4, #6, and #8 contained civilians with guns in their hand. These four videos also contained fixation data. Therefore, the fixation variable was developed by combining the fixation time values derived from each of the eight videos. Videos #5 and #9 did not contain an object in the civilian's hand. For this reason, there were no fixation data to present and these two videos were not included in the fixation variable. As with the pre-test, participants were excluded if their video was not usable. One person was excluded from the control group while one person was excluded from the test group.

Table 7. Post-Test Fixation Variable

	Control Group	Test Group
N	43	42
Mean	.46	.38
SD	.07	.10
Min	.29	.07
Max	.60	.57

Missing Data Procedure

As with most data, the current fixation variable had a small number of missing data. Specifically, the fixation variable was missing 34 of the 680 value observations. These missing values occurred when the participant blinked or squinted at the exact moment the item was presented. Generally, a participant would only be missing a single value in the fixation variable. For this reason, efforts were made to correct the missing values.

There are many ways to address missing data; however, some of these methods include removing observations or ignoring the missing data. This was not ideal for the current study as ignoring the data or removing the participant (i.e., removing the entire observation due to a single missing value) would have resulted in lowering the overall sum of the variable. In other words, if one response time value were missing from the variable, the overall fixation time would be inherently shorter for the participant. For this reason, the best method for dealing with these missing data required retention of the participant data. There are multiple missing-data approaches that can retain the data—e.g., mean imputation, last value carried forward, informed by related observations (Sterne, White, Carlin, Sprat, Royston, Kenward, Wood, & Carpenter; 2009). Mean imputation occurs when one utilizes the mean value for the group to replace the missing values. The downside to mean imputation is a narrowing of the variance. This happens due to having less variation from the mean. Last value carried forward refers to the process of imputing data based on prior performance. Last value carried forward does not account for variations in the test measure. In other words, if test variable #4 is inherently different from #3, imputing from #3 may not give an accurate estimation of what variable #4 should be.

Mean imputation was chosen to address the few cases of missing data. All of the videos were very different from one another. Any attempt to predict missing values based on prior performance would be affected by the complexity of the different videos. In other words, some of the videos had slower response times due to the complexity of the scenario while some videos had fast times. Mean imputation allowed for the individual video performance to fill in the missing data. Fixation variable results are presented both with the missing speed values and with imputed means in place of the missing speed values.

The fixation variable was made up of several observations. The mean for the specific observation for the specific group was chosen to fill in any missing value for that observation and group. For example, if the control group has a missing fixation time for a participant on post-test video #3 the mean fixation time for the control group on video #3 was utilized to replace the missing value. Each video is different; thus, the decisions and fixation times are different from the other videos. For this reason, replacement from prior observations would not be an accurate representation of the data.

Reading the Graphics/Results

Before the data are presented, the current section will describe how to read the analysis results. Recall from the previous chapter, the data presented in these analyses are estimations of the parameter via Bayesian data analysis programs developed by John Kruschke. In this case, there are 100,000 parameter estimates based on the data originally obtained in the experiment. There are no p-values present; rather, the data are presented as a credible posterior parameter distribution for each group (i.e., control and test group), credible posterior distribution of mean difference, and credible posterior distribution of effect sizes. This method of analysis allows the user to understand what the data look like, what the difference is between groups, and the level of impact presented by this difference. An example analysis was created (see Figure 15 below). To illustrate the data distribution a series of graphs are created. The graphs are histograms representing the distribution of credible posterior parameter estimates created via the MCMC simulation process.

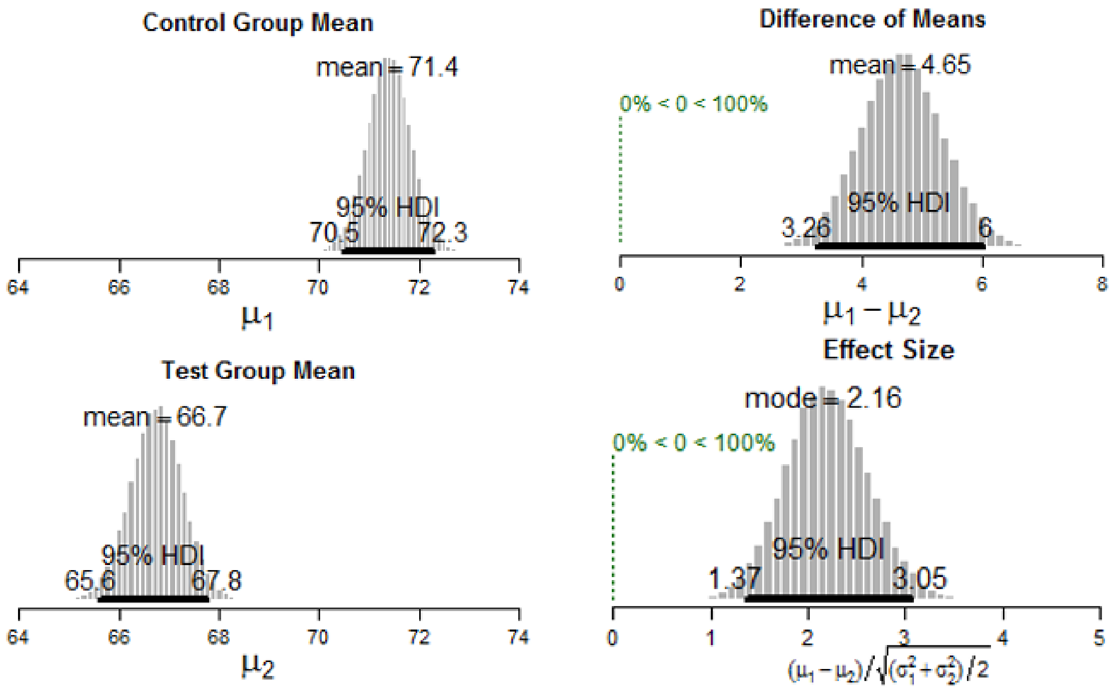

Figure 15. Example Data Output

The data depicted in Figure 15 are fictitious. These data represent height differences between plants undergoing a study to develop a fertilizer to reduce the height of corn stalks so shorter farmers can harvest corn. During this fictitious study, the control group received normal fertilizer while the test group received new fertilizer that the farmer would hope to reduce the height of the corn plant. These data were input in inches. This section will first describe what is shown in the graphics, and then the manner in which the graphic should be interpreted will be discussed.

Description of Graphics. The program presents the results from the MCMC simulations as histograms. Each parameter estimate is derived from the confidence interval calculated from the raw data. This process is repeated 100,000 times during the MCMC simulation. The result is the histogram presented above. It is important to note that these are not histograms of the raw data with a single sample estimate. Rather, the histograms are representative of credible posterior parameter estimates. The MCMC simulation process helps control for data that are not normally distributed through this process.

Each distribution is shown with its mean and Highest Density Interval (HDI). An HDI is the Bayesian program's equivalent of a confidence interval. In order to make the interval analogous to frequentists confidence intervals, 95% HDIs were utilized. The HDI contains 95% of the credible parameter values as determined by the MCMC simulation process. Recall from the discussion of the Bayesian program in Chapter 3, the credible parameter estimates are supported, and estimated, by the collected data. As such, the HDI (aka: credible interval) is interpreted as the range of simulated parameter values where the true parameter point is likely located. This process is repeated for both groups. The difference of means is computed by subtracting the test group mean from the control group mean. This difference of means is then utilized to develop the effect size distribution. The effect size utilized by the program is a Cohen's *d*. The difference of means is divided by the pooled standard deviation to create the Cohen's *d* effect size.

Interpretation of Example Graphic. When interpreting the results, it is imperative to look at the HDI range of the difference of means and effect size estimates. If the difference of means HDI range crosses zero, one cannot be certain that a difference is actually present. Rather, the parameter value could be positive or negative. If this is the case, one may assume any difference present in the data was a result of random assignment error. The effect size HDI should also not cross zero. If the effect size HDI crosses zero one cannot be certain of the effect's magnitude.

In this example, the control group crops were 71.4 inches on average (95% HDI [70.5, 72.3]). The test group crops grew to an average height of 66.7 inches (95% HDI [65.6, 67.8]). The 95% HDI for the difference in means (4.65 inches) ranged from 3.26 to 6. This suggests that the observed difference was not a product of random assignment error. The effect size of this difference is considered large (Cohen's *d* = 2.16, 95% HDI [1.37, 3.05]). These data suggest the test intervention (i.e., new fertilizer) was more effective at reducing the overall height of the corn stalks than the control fertilizer. The hypothesis that test group's fertilizer will reduce the height of the corn stalk more than the control group's standard fertilizer is supported by these data.

Bayesian Analysis

Accuracy of Weapon Identification

Research Question #1: Will a computer-based, vision-training program improve the accuracy of weapon identification in a dynamic situation?

Hypothesis: Participant's receiving the test intervention will be more accurate at identifying the presence of a weapon than participants receiving the control intervention.

Data required to answer the first research question were derived from the previously described accuracy variable.

Pre-Test. The pre-test control group made 16 total errors, while the test group made 13 total errors. This corresponds with a control group mean of .37 (95% HDI [.18, .56]) and a test group mean of .30 (95% HDI [.16, .44]). The difference in means (.07 errors per person) ranged from -0.16 to .30, suggesting that any observed difference may be a product of random assignment error. The effect size of this difference was very small

(Cohen's *d* = .13, 95% HDI [-.30, .55]). Additionally, the effect size HDI crosses zero, suggesting the observed effect could plausibly be explained by random assignment error.

Post-Test. Figure 16 illustrates participants' accuracy results for the post-test. As described in Chapter 3, the Bayesian program allowed for the imputation of the pre-test results as the prior observation. This process allowed the observed data adjust the credible data from pre-test levels to the presented levels in Figure 16. The average group mean for the pre-test control and test group was utilized for the informed prior (μ = .335). The average group standard deviation for the pre-test control and test group was utilized for the informed prior standard deviation (σ = .545).

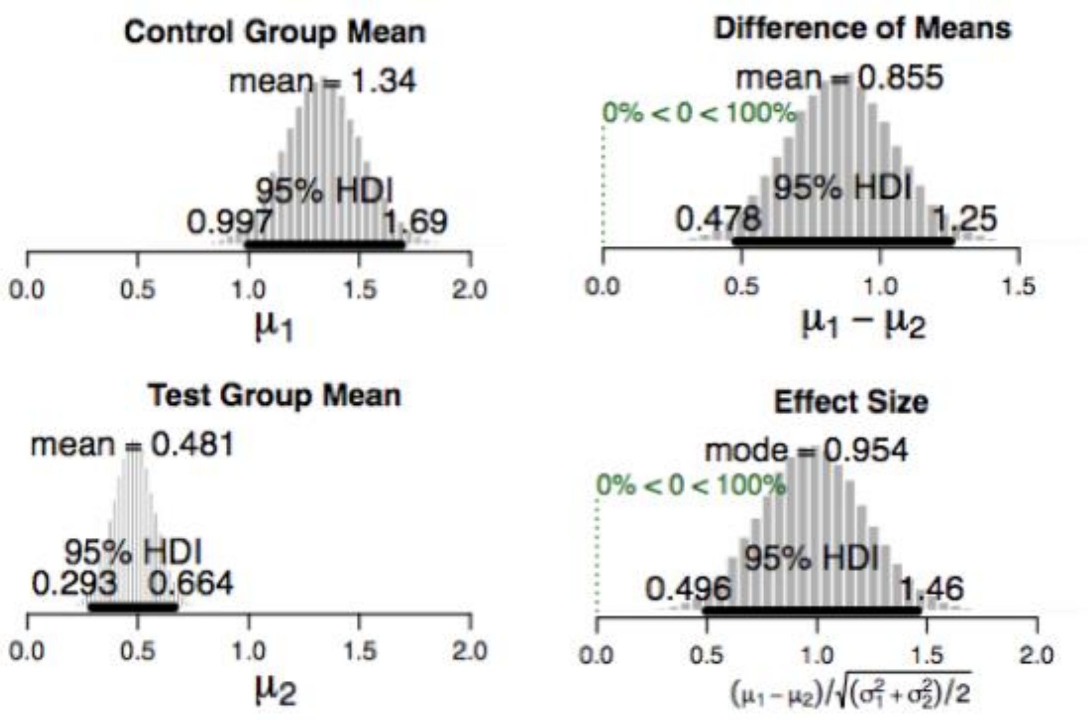

Figure 16. Post-Test Accuracy Results

The control group made 61 total errors, while the test group made 22 total errors during the post-test. This corresponds with a control group mean of 1.34 errors per person

(95% HDI [1.0, 1.69]) and a test group mean of .48 errors per person (95% HDI [.29, .66]). The 95% HDI for the difference in means (.86 errors) ranged from .48 to 1.25. One hundred percent of the HDI is above zero. This suggests that the observed difference was not a product of random assignment error. The effect size of this difference is considered large (Cohen's d = .95, 95% HDI [.50, 1.46]). The HDI ranges from a moderate effect size (.50) to a large effect size (1.46). These data suggest the test intervention was more effective at training participants to make accurate decisions regarding the presence of a gun. The hypothesis that test participants would make more accurate decisions than the control group is supported by these data.

These data were count data and not normally distributed according to a Shapiro-Wilks test of normality (w = .78, $p < .001$). While the Bayesian program corrects for non-normally distributed data by clustering credible values, a secondary analysis (i.e., test of proportions) was performed to further assess the participants' error contingency table. As seen in Figure 17, the relative frequency of an error for the control group (i.e., Group 1) was 14% while the relative frequency of an error for the test group (i.e., Group 2) was 5.6%. The results from the test of proportions were then converted to an odds ratio. The odds of having an incorrect decision for the control group were 14:86 (.162), while the odds of being incorrect were 5.6:94.4 (.059) for the test group. This corresponded with an odds ratio of 2.75. Therefore, the odds of making an error in the control group were 2.75 times the odds of making an error in the test group. These results are consistent in direction and magnitude with the previous error test presented in Figure 16.

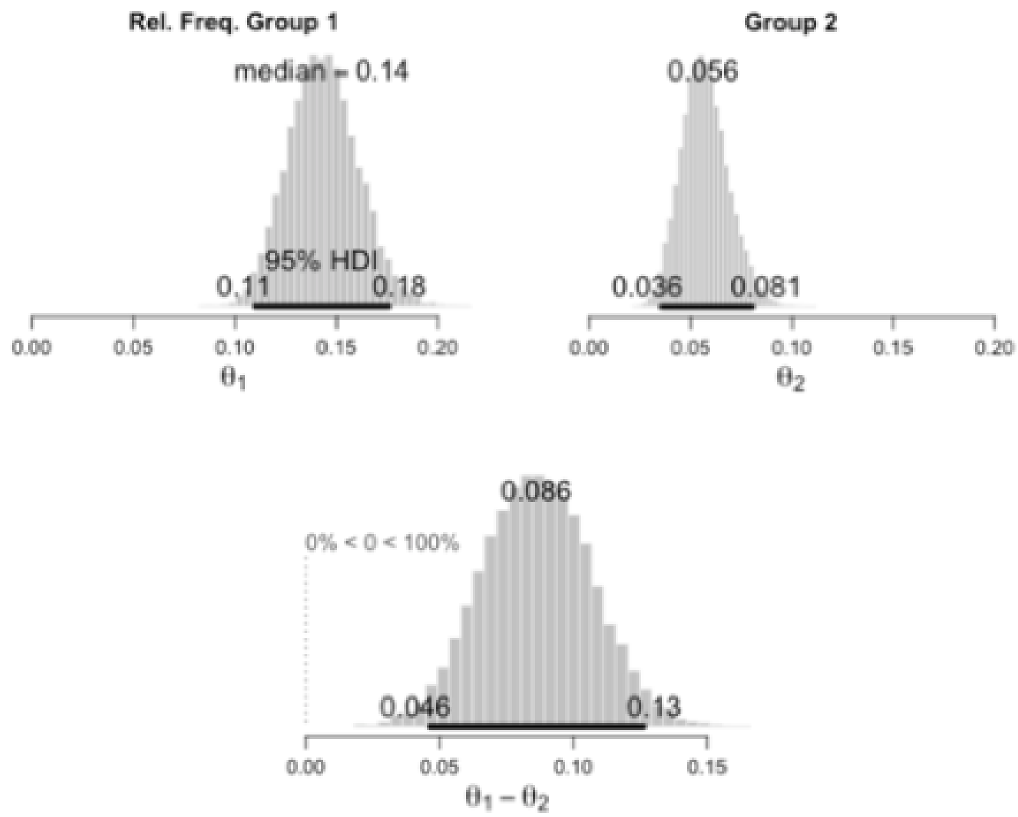

Figure 17. Relative Frequencies of Post-Test Errors

Speed of Weapon Identification

Research Question #2: Will a computer-based, vision-training program improve the speed of weapon identification in a dynamic situation?

Hypothesis: Participant's receiving the test intervention will verbally identify the presence of a weapon faster than participants receiving the control intervention.

To address the research question, participants were asked to verbalize when they saw a gun in each video.

Pre-Test. It took the control group .64 seconds (95% HDI [.57, .72]), on average, to verbalize the presence of a gun. The test group required .68 seconds (95% HDI [.61, .76]) to verbalize the presence of a gun. The difference in means (-.48 seconds) ranged from -.15 to .06, suggesting any observed difference may be a product of random assignment error. The effect size of this difference was small (Cohen's *d* = -.19, 95% HDI [-.62, .27]). The effect size HDI also crosses zero, suggesting observed differences are attributable to random assignment error.

Post-Test. Figure 18 presents the decision speed data for the post-test. There were no missing decision speed data to correct for within the decision speed variable. Even if the vision track was bad, the decision was captured via the captured audio. As with the previous section, the pre-test was utilized to inform the prior distribution in this analysis. As such, the average group mean for the pre-test control and test group was utilized for the informed decision speed prior ($\mu = .66$). The average group standard deviation for the pre-test control and test group was utilized for the informed decision speed prior standard deviation ($\sigma = .23$).

The control group took .4 seconds (95% HDI [.34, .46]), on average, to verbalize the presence of a gun across the four post-test videos with a gun present. The test group took .36 seconds (95% HDI [.30, .42]), on average, to verbalize the presence of a gun. The difference of means HDI (.04 seconds, 95% HDI [-.04, .12]) crossed zero as a possible parameter value estimate. This suggests any observed difference may have been a product of random assignment error. However, over 84% of the credible values were positive (i.e., the test group made a faster decision than the control group). The effect size of this difference was small and also crossed zero (Cohen's *d* = .26, 95% HDI [-.21, .69]). This suggests the observed effect could be attributed to random assignment error. The hypothesis that test participants would make faster decisions than the control group was not supported by these data.

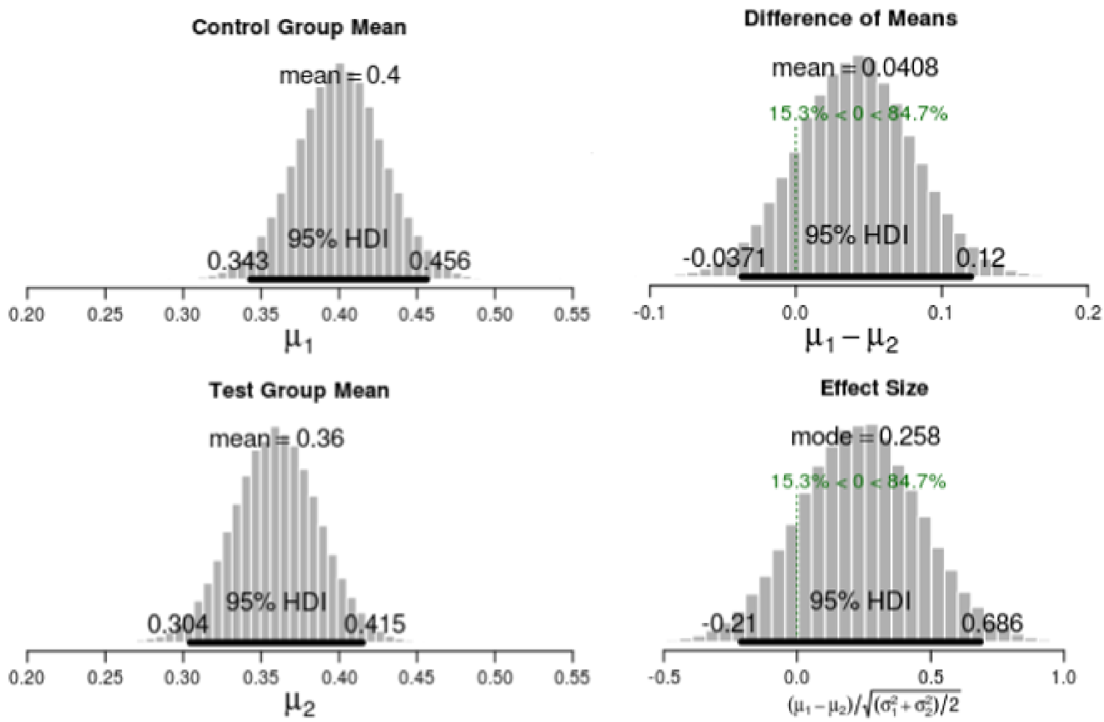

Figure 18. Post-Test Decision Speed Data

While the test intervention does not appear to improve participants' ability to find a gun faster, there are important findings regarding the speed of decision presented in an upcoming section titled, "Fixation vs. Accuracy". The effectiveness of the program to improve the participants search strategy is presented next.

Improvement in Search Strategy

Research Question #3: Will a computer-based, vision-training program improve the visual search strategy of participants in a dynamic situation?

Hypothesis: Participant's receiving the test intervention will visually fixate on the presented item faster than participants receiving the control intervention.

In order to answer this research question, data were gathered to measure how long it took participants to fixate on the item presented in the pre- and post-test videos. These data were derived from the vision tracking video and were also measured in frames per second. The data were transformed to seconds for analysis.

Pre-Test. It took the control group .55 seconds (95% HDI [.52, .59]) to visually fixate on the item. The test group required .54 seconds (95% HDI [.51, .58]) to visually fixate on the item. The difference in means (.01 seconds) ranged from -.04 to .06, suggesting any observed difference may be a product of random assignment error. The effect size of this difference was very small (Cohen's d = .10, 95% HDI [-.35, .55]). The effect size HDI also crosses zero, suggesting any observed differences are attributable to random assignment error.

Post-Test. As previously discussed, there were missing values within the computed fixation variable. These occurred when a participant blinked or squinted his/her eyes at the exact moment the item was being presented. Out of the 680 combined individual variables making up the index, there were 34 mean imputations (control group n=19, test group n=15). This section of the analysis is presented both with and without the mean imputation previously discussed. Figure 19 presents the fixation data for the post-test without mean imputation. Figure 20 presents the fixation data for the post-test with the mean imputation. For both graphics, the pre-test was utilized to inform the prior distribution. The average group mean for the pre-test control and test group was utilized for the informed fixation time prior (μ = .55). The average group standard deviation for the pre-test control and test group was utilized for the informed fixation speed prior standard deviation (σ = .13).

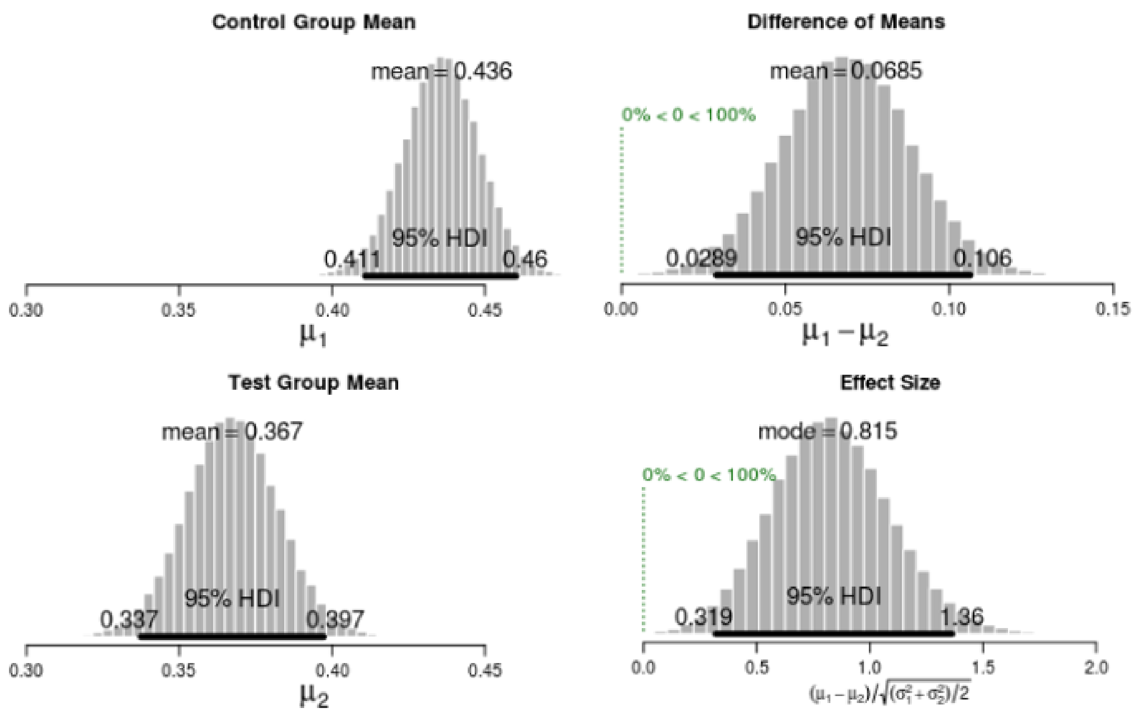

Figure 19. Post-Test Fixation Times: Without Mean Imputation

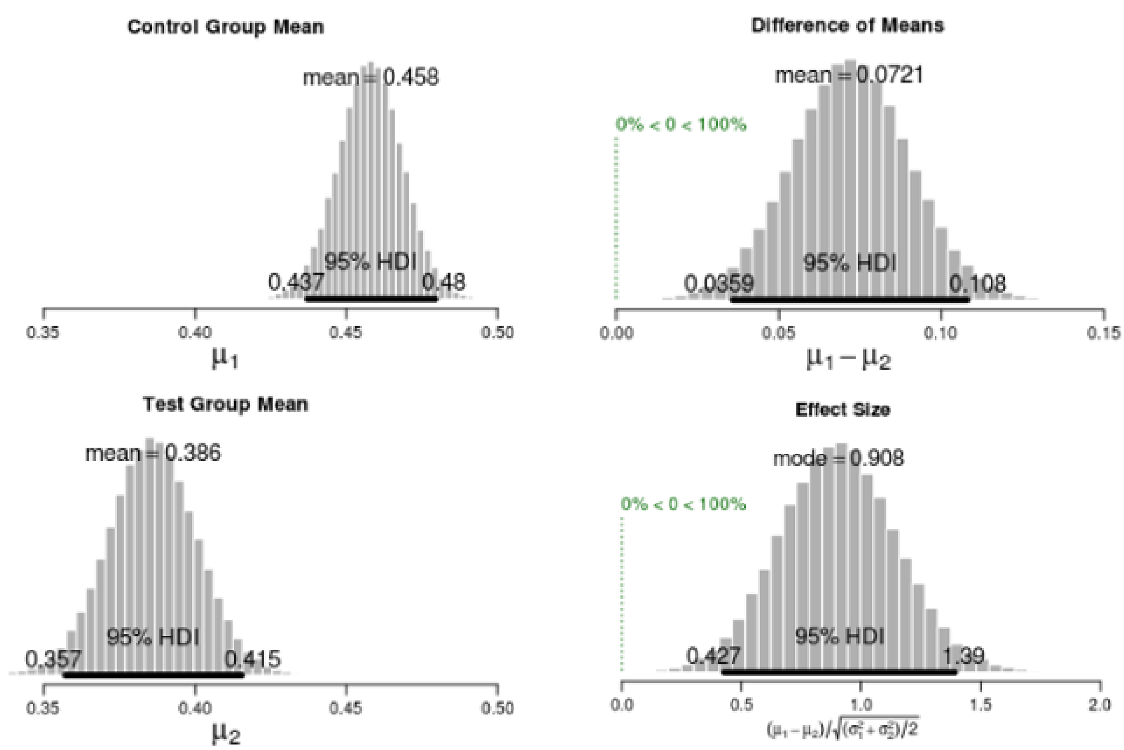

Figure 20. Post-Test Fixation Times: With Mean Imputation

As shown with both graphics, there was minimal change when performing the mean imputation for the fixation time data. The group means did slightly increase. However, this was to be expected as the imputation added additional fixation data to the participant's fixation variable score. Even though the group means increased, the difference in means only increased slightly (i.e., .004 seconds per video). For this reason, the below presentation of the data will discuss Figure 20 that includes the mean imputed data as it is a more accurate representation of the data (i.e., all values are present for all participants via mean imputation).

The control group took .46 seconds, on average, to fixate on the items (95% HDI [.44, .48]). Alternatively, the test group required .39 seconds, on average, to fixate on the items (95% HDI [.36, .42]). The difference of means for the two conditions was .07 seconds. The difference of means HDI ranged from .04 to .11, suggesting the observed difference is not a product of random assignment error. The effect size for the observed difference is large while the HDI ranges from a medium effect to large (Cohen's d = .91, 95% HDI [.43, 1.39]). This suggests the test intervention was more effective at improving the participants' search strategy by reducing the amount of time required to fixate on the item. The hypothesis that test participants would visually fixate on the presented item faster than the control participants is supported by these data.

Fixation vs. Accuracy

The following section is a supplementary analysis that developed after analyzing data for the research questions and hypotheses. Specifically, the post-test results allowed for a further analysis of the visual fixation versus accuracy dynamic. Utilizing the vision tracker, data were gathered to assess participants' visual fixation on the item related to their verbal identification of the gun. That is, did the participant actually *SEE* the gun before making a decision? Recall that visual acuity is closely tied to foveal vision. So, utilizing foveal vision to gather information about the object would aid in the correct identification of the object as a gun or not. If the participant only utilized the less accurate peripheral vision, it is possible that the participant would have incomplete information to base their decision on. If incomplete data are utilized, the brain will have to interpolate beyond the scope of the visual data obtained. Because of this, the more accurate question to ask is, was foveal vision associated with participant accuracy?

The following data are utilized to provide a cursory look at this question (see Table 8). First, decision and fixation speed data were pulled for all the times a decision was made. This included not only correct decisions, but it also included all the times the individual incorrectly identified an object as a gun. The fixation time data were subtracted from the decision time data. If the resulting value was a negative integer this indicated the individual verbalized their decision before visually fixating on the object. If the resulting value was a positive integer this indicated the individual visually fixated on the object before verbalizing their decision. Specifically, decision first refers to the participant verbalizing the presence of a gun before visually fixating on the object. Fixation first refers to the individual fixating on the object before verbalizing the presence of a gun. Errors that occurred during videos where a gun was present were excluded. This is because an error occurs when the individual doesn’t verbalize the presence of a gun. Therefore, there was no decision time present to include in the table.

Table 8. Fixation vs. Decision: Descriptive Data

Video Name	Control		Test	
	Decision First	Fixation First	Decision First	Fixation First
Posttest #1	14	12	8	7
Posttest #2	4	0	1	0
Posttest #3	28	13	23	16
Posttest #4	16	23	4	34
Posttest #6	8	35	4	35
Posttest #7	17	2	2	0
Posttest #8	1	42	0	40
Posttest #10	0	2	0	2
Total	88 (41%)	129 (59%)	22 (14%)	134 (86%)

*Percentages in parentheses

As seen in Table 8, it appears the test group was more likely to fixate on the item before voicing their decision. However, this table does not relate these data to a participant's ability to make a correct decision. To explore this relationship further, Kruschke's Bayesian logistic regression program was utilized (Kruschke, 2014). The dichotomous dependent variable was participant error (no error = 0, error =1). There were two primary independent variables. First, a dichotomous fixation variable (fixate). This variable was created from Table 8 above. If the participant fixated on the item first, he/she was coded as a 1. If the participate made a decision before fixating on the item he/she was coded as a 0. The second independent variable was a dichotomous test condition variable (test). If the participant was a member of the test group, he/she was coded as a 1. If the participant was a member of the control group, he/she was coded as a 0. Figure 21 shows the distributions for the model's coefficients.

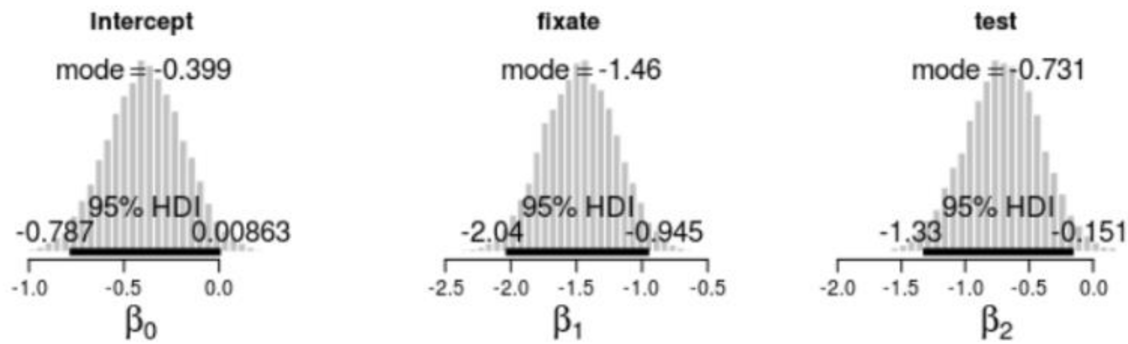

Figure 21. Fixation vs. Accuracy – Main Effects

Figure 22 illustrates the coefficient distributions once an interaction term was added to the prior analysis. Multiplying the fixation variable with the test condition variable created the interaction term. As such, the interaction term for participants that fixated first (coded as 1) and were in the test group (coded as 1) was coded as a 1. All the other combinations were coded as 0 as one of the dichotomous variables in the term would have a value of zero. The interaction term was created in this fashion as the previous analysis suggested this combination had the lowest probability of making an error. For this reason, the interaction term allowed for an assessment of this combination (test group and fixation first) as it compared to the other three coefficient combinations.

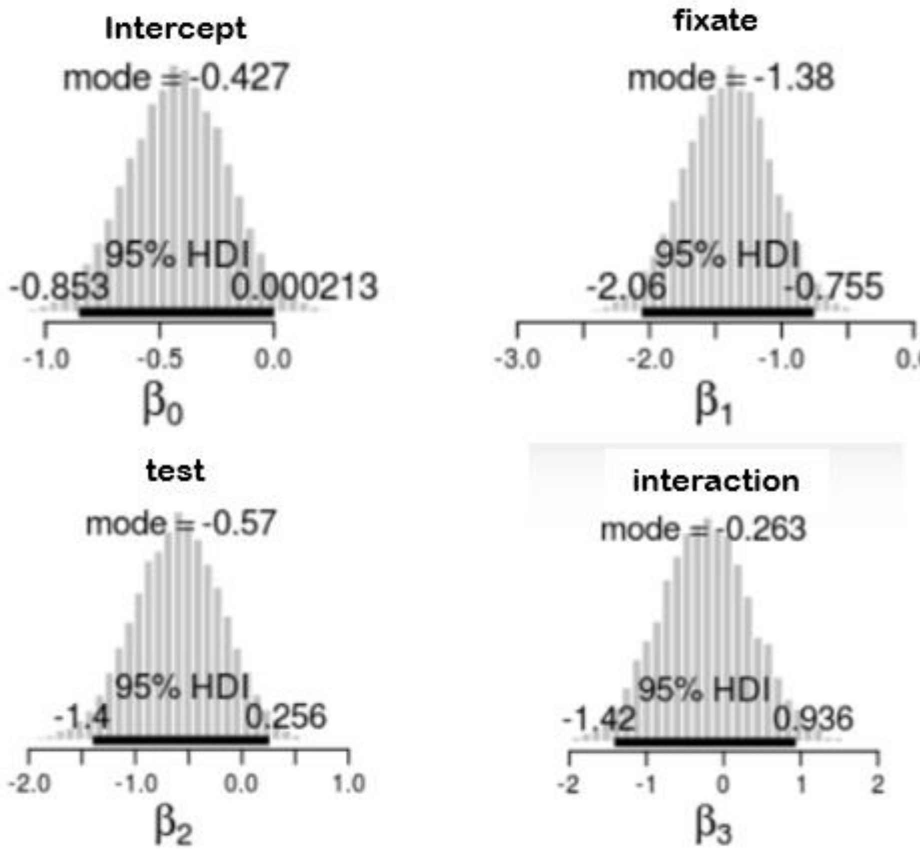

Figure 22. Fixation vs. Accuracy – Main Effects & Interaction

As seen in Figure 22, once the interaction term was added, the test condition variable's 95% HDI crossed zero (95% HDI [-1.4, .26]). As such, any observed impact based on test condition may be attributable to random assignment error. However, the majority of credible estimates were still in the direction as predicted. That is, the majority of credible estimates still indicate that being a member of the test group reduces one's probability of making an error.

The numerical magnitude for each slope coefficient was interpreted through the application of the logistic function ($\beta_0 + \beta_1 x_1 + \beta_2 x_2$) to determine the log odds of making an error given the visual fixation variable and experimental condition variable. As the independent variables are dichotomous, the log odds were calculated for each combination. The log odds were then converted to probabilities to aid in comprehension of the values. These data show that participants in the test group that fixated on the item before making a decision had the lowest probability of making an error (p = .07). Furthermore, participants in the control group that did not fixate on the item before making a decision had the highest probability of making an error (p = .40). The interaction term coefficients were added in the second row. The interaction slightly reduced the log odds of the condition coefficient, but did not appear to have an impact on the overall probability of error. This suggests that the interaction between the main effects may be negligible. The results are presented below in Table 9.

Table 9. Main Effects: Predicted Log Odds

	β_1	β_2	β_3	Log Odds	Probability of Error
Fixate – 1, Test – 1	-1.46	-0.73	--	-2.59	.07
Fixate – 1, Test – 1, Int – 1	-1.38	-0.57	-0.26	-2.64	.07
Fixate – 1, Test – 0	-1.46	0.00	--	-1.86	.13
Fixate – 0, Test – 1	0.00	-0.73	--	-1.13	.24
Fixate – 0, Test – 0	0.00	0.00	--	-0.40	.40

The next chapter will provide a discussion of these findings and their implications for law enforcement training. Furthermore, the limitations and future direction of this research agenda will be presented.

V. DISCUSSION

In today's environment where police actions are under intense scrutiny, improving the ability of law enforcement officers to correctly identify the presence of a gun in a dynamic use of force situation is a small, but important, step toward protecting the community. As such, a new training program was developed and tested. The training program tested within this dissertation was developed by bringing together multiple disciplines, including adult learning, sports science, visual processing, decision-making, and use of force correlates. It was my hope that bringing together aspects of multiple disciplines could result in a training program that effectively trained a specific task in the most efficient way possible. In developing this new training program, the present study was a proof of concept. Utilizing the key principles of the learning model known as deliberate practice, the training program provided participants with increasingly complex media—both still image and video. Throughout the process, the participants received feedback and several repetitions. This method of training was meant to improve participant performance. Specifically, the study sought to find if this new way of training could improve the speed and accuracy of use of force decisions.

The most important thing to consider when assessing the results of the study is the participants received only a small amount of training. The participants completed each level of training only one time, and most of the participants finished in about 25 minutes. Recall that the research into deliberate practice estimates 10,000 hours of effortful training to become an expert in a task (Ericsson, 1993). The goal of the training program was not to get participants to an expert level. Rather, the goal of the program was to exhibit improvement during short periods of training. In fact, the training program was successful at improving participant decision-making accuracy. Participants exposed to the newly developed training program made approximately 1/3 the total number of errors of the control group. The effect size of this finding was large. The test group also visually fixated on the items faster than the control group. In fact, the test group was about 20% faster than the control group at fixating on the item—and the difference equated to another large effect size.

It is worth noting that the differences observed in this study involved fractions of seconds. For instance, the test group visually fixated on the item .07 seconds faster than the control group, on average. While these differences are short in terms of speed, use of force encounters unfold quickly. The smallest of differences could result in an encounter changing from an improper use of too much force to the proper application of force necessary to control a dynamic situation. All of these findings are extremely exciting and bode well for future research endeavors.

While the test group made fewer errors, they did not make their decisions faster than the control group. In fact, there was no discernable difference between the control and test group regarding decision speed. So, while the control group was equally as fast at verbalizing their decision as the test group, they were not as accurate. The visual fixation versus decision accuracy relationship was not hypothesized and required further exploration.

Fixation vs. Accuracy

The prior chapter identified when participants verbalized their decision in relation to the moment they visually fixated on the item. These data were pulled for further analysis.

The new variables were run through a Bayesian logistic regression. Those results (see Figure 18) suggested that participates that fixated on the item before making a decision were less likely to make an error. Additionally, the first model suggested that participants in the test group were less likely to make an error. However, once an interaction term was added in the second model this relationship was removed. The second model suggested that fixating on the item alone predicted whether or not the participant was likely to make an error. These analyses were supplementary to the primary analysis. Specific data concerning this foveal vs. peripheral vision dynamic were not captured. For instance, there was not a decision verbalized if the participate correctly identified the item as a non-gun. In these cases, the only data gathered is fixation data. Future iterations of the training program will need to take this dynamic into consideration and gather more complete data. This will be discussed further in the "Future Direction" section.

Regardless, the finding on item fixation is interesting. With participants not utilizing foveal vision, they were not likely to have gathered the necessary information to make the correct decision. Recall the decision-making process known as the OODA loop described in Chapter 2 (p. 23). During the *Orient* phase of the OODA loop the individual gathers as much information as possible and sends this information to the brain for processing. As such, foveal vision is imperative to gathering pertinent information needed to make an accurate decision. Without foveal vision, an individual is required to process interrupted, partial visual data attributed to the less accurate peripheral vision. This foveal vs. peripheral vision dynamic likely explains why there was no difference in decision-making times between the test and control group. While fixation times were faster for the test group, the control group likely was utilizing peripheral vision to make the decision in a “gun” video without fixating on the item. This would result in their decisions being made equally as fast as the test group even though their fixation times were not as fast. In terms of fixation speed vs. accuracy, these initial data suggest that utilizing foveal vision prior to making a decision can reduce the number of errors made.

For all of the reasons mentioned thus far in this chapter, the findings from this study appear to suggest a training program following the key principles of deliberate practice is effective at improving individuals’ ability to detect a weapon with greater accuracy. While no difference was found between groups regarding speed of decision, the program was effective at improving the visual fixation times of the test group.

Future Direction of Research

The current dissertation research served as a proof of concept for a much larger series of studies. This study has shown that a training program based on the concept of deliberate practice can effectively improve participant performance in terms of accuracy and visual verification of an item. As this was a proof of concept, the training time was short and I tried to limit the number of force correlates presented in the training and test videos. The next phase of the project will begin to add in additional correlates of force. Specifically, civilian race will be added in the next immediate phase. The current state of policing and race relations within the media makes this the next logical correlate to study.

As this will only be the second correlate added to the training program, the outcome measure will remain the accurate identification of a weapon. As the training program advances and more correlates are added, the user will not be simply identifying the presence of a gun. Rather, participants will work through use of force encounters and have the option of moving through the force continuum. This will allow officers to decide on an appropriate action from verbal commands to deadly force. This process will also give law enforcement officers training on a wider range of situations. In order to capture the complex number of potential actions, future iterations of the research will need to be done in a use of force training simulator. A training simulator allows the officer to utilize a variety of force options (e.g., gun, conducted energy device, OC spray) beyond simply verbally identifying an object. The increased number of system-supported options automates the data collection process and increases specificity. Specifically, the simulator systems capture the exact moment an individual deploys a system-supported force option.

Furthermore, adding a race variable into the training program increases the needed training images and video in order to include all the possible combinations of guns, non-guns, and race. Inherent with increasing the number of training videos, the amount of time for completing the training will also increase. Instead of completing the training over the course of two days, the training will take place over several days. This process will be further extended as more correlates are added.

As the program is moving beyond a proof of concept, the sample will need to be adjusted. The training is specific for law enforcement. As such, future participants will be law enforcement officers. Officers that attend ALERRT courses, as well as local law enforcement officers, will be solicited to participate in future iterations of the study. It is anticipated that the effect sizes will shrink. This is because law enforcement officers will have a baseline of training beyond what the non-law enforcement sample had for this study.

The other force correlates will be systematically added to the training program following the race correlate. This will allow for a stepwise assessment of each of the correlates. This process will also allow for refinement of the training program as needed to ensure the best training is provided to law enforcement officers. If this training program is able to help law enforcement officers make fast, accurate decisions, keeping both law enforcement officers and innocent civilians safe, it is a success. This goal of improved safety provides the motivation and guidance to continually improve the training program.

Limitations

Much like all research, the current project had limitations that need to be discussed. First, this project was a proof of concept. Therefore, the training program attempted to utilize a single correlate of force to train the participants in order to prove the basis of the program worked before moving forward. While it is anticipated that these results will also be present once more correlates of force are added, there is no guarantee. The presence of a gun is objective…there either is or is not a gun. This made the presence of a gun a clear first correlate to introduce into the training program. Some other correlates of force are subjective and more difficult to measure. For instance, does each person see a certain demeanor as hostile, or can it be interpreted in multiple ways?

Second, the participants were asked to verbally identify a gun in a series of videos. While they were also asked to identify other objects as “not gun”, the participants did not do this. This is likely because they were solely focused on finding a gun; thus, the participants likely skipped all other objects. This limited the study in that the only measurable decision occurred if they verbalized a gun, either correctly or as an error. Future iterations of this study need to incorporate improved methods for collecting decision data. One such method may be incorporating a handheld device where the participant pressed a button for gun or no gun based on what they are seeing at the moment. This decision could then be synced with the vision tracking video for a more accurate representation of the decision-making process.

The study was also limited in that participants only received about 25 minutes of training. Ideally the participants would receive continued training over a longer period of time. However, as the participants were student volunteers, it would have been difficult to have them participate for a long period of time. Furthermore, participant mortality would be high. While the low amount of training time is a limitation to the study, the positive results are amplified when one considers the small amount of training time for the participants. Future iterations of this research should, and will, incorporate longer training periods for participants.

Lastly, the media utilized for the study limited the effectiveness of the training and evaluation. Amateurs with no photography/cinematography experience shot the still images and videos utilized for the training program. As such, the lighting and quality of the images/videos can be much improved. In the future, professionally trained individuals will create the media. Further, the models for the training program were co-workers and friends of the author. I limited the models to white males. All models wore a blue shirt and jeans to limit variation and focus on the presence of a gun. These precautions were taken to eliminate extraneous variance (e.g., demeanor, demographics). Future iterations of the training program will be actors from the Department of Theater at Texas State University. The author has been given approval to recruit students of the Theater Department for future projects. These students will be representative of all races and genders. It is anticipated that the student actors will be trained and capable of performing various states of emotion needed to measure other correlates of force.

Beyond the training program, the pre- and post-test videos are also a limitation to address. MILO Systems provided the videos for both the pre- and post-test. While there were many videos to choose from, few could be used. An attempt was made to limit the number of additional correlates visible in the videos while having a clear view of the item to be identified. Many of the videos were made in prior decades. As such, the quality of the video was slightly blurry and the audio was less than clear. The videos used in this study worked; however, future studies need to have professionally made videos that are clear. Additionally, the creation of videos will allow for specific features to be present in each video. This level of control will aid in ensuring the key concepts being training are also being assessed in the evaluation period. Pre-test videos can be made with equal correlates as the post-test videos for a more accurate representation of improvement. This level of control will also allow for a repeated measures approach when analyzing the findings. The current project was unable to utilize repeated measures as the videos varied significantly in terms of all the correlates present, complexity, length, etc.

While there are several limitations, these do not reduce the importance of the current research. The limitations were acceptable as the research was an initial look at incorporating several aspects of training from different disciplines into one training program. It is my belief that the positive outcome of the research highlights the importance of drawing from multiple disciplines to solve complex problems, such as developing a comprehensive use of force training program to improve law enforcement performance in dynamic situations.

APPENDIX SECTION

APPENDIX A

Description of Pre-Test Video

Video 1 (No Gun)
Scenario: pts courtesy stop
Officer is talking with driver. The driver turns to the passenger seat with her back to the officer. She turns back around with a wallet in her hand.

Video 2 (Gun)
Scenario: School shooting threat assessment
While approaching an active shooter a man darts in front of the officer. He is an officer, but he is not in uniform and has a gun in his hand.

Video 3 (Gun)
Scenario: Room clearing
Officers are clearing rooms in a building looking for a murder suspect. A man approaches with his hands in the air. He moves his hand down and pulls out a pistol.

Video 4 (No Gun)
Scenario: Car stop 2
Officer approaches car at night. The driver reaches up to the visor and pulls down a badge without a gun.

APPENDIX B

Description of Post-Test Video

Video 1 (No Gun)
Scenario: Alarm Call 1
Suspect appears in a car workshop. While speaking with the officer the suspect puts a tool in the toolbox.

Video 2 (No Gun)
Scenario: Crazy Brother
Respond to a call and a woman approaching telling the officer that her brother has gone crazy. The brother appears with a knife in his hand. He then runs away.

Video 3 (Gun)
Scenario: dfmp902
The officer approaches the vehicle. The individual reaches into his pocket and gives the officer is license. However, after a few second he gets out of the car and pulls a weapon.

Video 4 (Gun)
Scenario: Baby Hostage
The officer approaches a house with a couple fighting over a child. The woman pulls a gun and points it at the child.

Video 5 (No Gun)
Scenario: Fired employee escort
A recently terminated employee paces back and forth among several objects. He reaches behind, however, he never comes out with a gun.

Video 6 (Gun)
Scenario: DV Daddy hurt mommy
The officer moves through the home. When the officer turns a corner he is confronted with a man holding a gun to his wife’s head. The gun is first visible through the stairway.

Video 7 (No Gun)
Scenario: Active shooter-suspect in hall
The officer approaches a male in the hall after an active shooter event. The man turns around quickly with a cell phone in his hand.

Video 8 (Gun)
Scenario: Court Disturbance
The officer enters a court room where a man is visibly upset. He leaves the room and comes back with a gun in his hand.

Video 9 (No Gun)
Scenario: ticket complaint
A man quickly approaches a police desk with an object in his hand. He is carrying a parking ticket that he pushes up against the window.

Video 10 (No Gun)
Scenario: 200
The officer approaches a man in a rundown bus. The door is covered with fabric. He is yelling and sticking a black object out of the door. It ends up being a black cane.

APPENDIX C

Demographic Information

Age: ____

Sex (please circle)
Male Female

Race/Ethnicity (please circle)
African American/Black
Asian
Caucasian/White
Latino/Hispanic
Other

Prior Law Enforcement Experience (please circle)
Yes No

Prior Military Experience (please circle)
Yes No

APPENDIX D

Example Program Instructions

Thank you for volunteering to participate in my research project. Please follow the below instructions to complete the training program.
You will be completing four levels of training over the next two days. On day one, you will complete Level I, II, & III. When you return for day two, you will complete Level IV.
Your participant number is printed on the piece of paper stapled to the top right corner of this sheet. Please remember/keep this number as you will need to enter it when logging in on the software.

Opening the Program (Insert Screenshots)
You will be completing the levels in order. So, locate Level I on your desktop and double-click. You will need to enter your participant number and sequence number. The sequence number is simply the level you have opened.
Please read the displayed instructional page prior to beginning each level. Once you press the button listed on the instructional page you will begin the level.

Closing the Program
Once you have completed the designated level you will simply press the Spacebar to exit. At this point you are able to open the next level and continue.
Once you have completed all the designated levels for the day please inform the graduate student in the room. He/she will inform you of the next step.

APPENDIX E

IRB Consent Form

Consent Form (IRB Approval # 2016O7324)

M. Hunter Martaindale of the ALERRT Center at Texas State University – San Marcos is conducting this research project. The researcher can be contacted at (512) 245-1332 or hunter.martaindale@txstate.edu. You were selected to participate in this project because you are a criminal justice student at Texas State University. It is estimated that your participation will take approximately two hours over two days (i.e., one hour a day). This research project is not externally funded.

The purpose of this study is to test the effectiveness of a computer based visual training program to improve the decision making process. The study will be conducted at the ALERRT Training Center. Should you choose to participate, you will be asked to complete the training program and be evaluated following the completion.

These sessions will utilize vision tracking technology that is incapable of capturing identifiable information. You will be able to withdraw your tape from the study if you desire at no penalty to yourself. The data gathered from the videotapes will be confidential. That is, the researcher will be able to identify your specific data according to a random number, however the results will not link your video to your identity. These data will be maintained in the researcher's office for the duration of the research project. The signed copies of the consent forms will be kept in a locked cabinet for three years. Once this time period is over they will be destroyed. If you wish to be provided with a summary of the results, the researcher will provide one to you upon request. Please contact the researcher at the e-mail above if you wish to receive a summary.

You may experience some psychological stress from performing the decision making assessment. Should you need to speak to someone about psychological issues after participation, the following mental health resources are available in the area: Texas State Counseling Center (512) 245-2167, Hill Country Community MHMR Center (830) 792-3300, and Counseling and Assessment Clinic on campus in the Education building (512) 245-8349. If you are registered Texas State student, the Texas State Counseling Center services are free, but the number of sessions allowed may be limited. You will be responsible for costs incurred at the other treatment centers.

You may benefit in that you will have the opportunity to experience working through a problem and improve your decision making process. The findings of the study will be used to inform future studies regarding the decision making process. These findings may improve decision making training policies and procedures.

Your participation is voluntary and you may choose to withdraw from the study at anytime without prejudice or jeopardy to your standing with Texas State University. You may also choose not to answer any of the demographic questions for any reason. This project [IRB Approval # 2016O7324] was approved by the Texas State IRB. Pertinent questions or concerns about the research, research participants' rights, and/or research-related injuries to participants should be directed to the IRB chair, Dr. Jon Lasser (512-245-3413 - lasser@txstate.edu) and to Becky Northcut, Director, Research Integrity & Compliance (512-245-2314 - bnorthcut@txstate.edu).

Participant Signature: ___________________ Researcher Signature: __________________
Date: ___________ Date: __________

LITERATURE CITED

Aamodt, M.G. (2004). *Research in law enforcement selection.* Boca Raton, FL: Brown Walker Press.

Ahmad, K. M., Klug, K., Herr, S., Sterling, P., & Schein, S. (2003). Cell density ratios in a foveal patch in macaque retina. *Visual Neuroscience*, *20*(02), 189-209.

Alpert, G.P., & Dunham R.G. (2010). Policy and training recommendations related to police use of CEDs: Overview of findings from a comprehensive national study. *Police Quarterly, 13,* 235-259.

Alpert, G. P., & MacDonald, J. M. (2001). Police use of force: An analysis of organizational characteristics. *Justice Quarterly, 18*(2), 393-409.

Alpert, G.P., Dunham, R.G., & Stroshine, M.S. (2006). *Policing: Continuity and Change,* Waveland, Long Grove, IL.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review, 89*(4), 369.

Anderson, D. R., Burnham, K. P., & Thompson, W. L. (2000). Null hypothesis testing: problems, prevalence, and an alternative. *The journal of wildlife management*, 912-923.

Artwohl, A., & Christensen, L. W. (1997). *Deadly force encounters: What cops need to know to mentally and physically prepare for and survive a gunfight*. Boulder, CO: Paladin Press.

Bahill, A. T., & LaRitz, T. (1984). Why can't batters keep their eyes on the ball? *American Scientist*, *72*(3), 249-253.

Bard, C., & Fleury, M. (1981). Considering eye movement as a predictor of attainment. *Vision and sport*, *20,* 28-41.

Bard, C., Fleury, M., Carrière, L., & Hallé, M. (1980). Analysis of gymnastics judges' visual search. *Research Quarterly for Exercise and Sport*, *51*(2), 267-273.

Bayley, D.H., & Mendelsohn, H. (1967). *Minorities and the Police: Confrontation in America,* Free Press, New York, NY.

BBC News. (2014) Sean Groubert charged in shooting of Levar Jones - BBC News. Retrieved September 12, 2015, from http://www.bbc.com/news/world-us-canada-29373115

Bédard, J., & Chi, M.T., (1993). Expertise in auditing; Discussion. *Auditing*, *12*, 21.

Blair, J. P., Pollock, J., Montague, D., Nichols, T., Curnutt, J., & Burns, D. (2011). Reasonableness and reaction time. *Police quarterly*, *14*(4), 1-21.

Boyd, J. R. (1996). The essence of winning and losing. *Unpublished lecture notes.*

Brebner, J. T., & Welford, A. T. (1980). Introduction: A historical background sketch. In A. T. Welford (Ed.), Reaction times (pp.1-23). New York, NY: Academic Press.

Brinkman, J. A. (1993). Verbal protocol accuracy in fault diagnosis. *Ergonomics*, *36*(11), 1381-1397.

Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). *Make it stick*. Harvard University Press.

Burroughs, W. A. (1984). Visual simulation training of baseball batters. *International Journal of Sport Psychology*.

Camerer, C. F., & Johnson, E. J. (1997). 10 The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In K.A. Ericsson and J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (pp. 195-217). Cambridge: Cambridge University Press.

Campbell, J. H. (1992). *A comparative analysis of the effects of post-shooting trauma on special agents of the Federal Bureau of Investigation*. Unpublished Ph.D. Dissertation. Department of Educational Administration, Michigan State University.

Carpenter, R. H. (1988). *Movements of the eyes.* London: Plion.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive psychology*, *22*(2), 225-271.

Charness, N., Krampe, R., Mayr, U., & Vasiukova, K. (1996). Deliberate practice in chess skill acquisition. *International journal of psychology*, *31*(3-4), 4662-4662.

Chi, M.T.H., Glaser, R., & Farr, M.J. (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Chinellato, E., Grzyb, B. J., Marzocchi, N., Bosco, A., Fattori, P., & Del Pobil, A. P. (2009). Eye-hand coordination for reaching in dorsal stream area V6A: Computational lessons. In *Bioinspired Applications in Artificial and Natural Computation* (pp. 304-313). Springer Berlin Heidelberg.

Christina, R. W., Barresi, J. V., & Shaffner, P. (1990). The development of response selection accuracy in a football linebacker using video training. *The Sport Psychologist*, *4*(1), 1.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (rev*. Lawrence Erlbaum Associates, Inc.

Ditchburn, R. W. (1973). *Eye-movements and visual perception*, London: Oxford University Press.

Denney, N. W. (1984). A model of cognitive development across the life span. *Developmental Review*, *4*(2), 171-191.

Doane, S. M., Pellegrino, J. W., & Klatzky, R. L. (1990). Expertise in a computer operating system: Conceptualization and performance. *Human-Computer Interaction*, *5*(2), 267-304.

Duncan, J., & Humphreys, G. (1989). A resemblance theory of visual search. *Psychological Review*, *96*, 433-458.

Duncan, J., & Humphreys, G. (1992). Beyond the search surface: visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 578-588.

Durose, M.R., Schmitt, E.L. & Langan, P.A. (2005). *Contacts between police and the public.* Washington, D.C.: Bureau of Justice Statistics.

Eith, C., & Durose, M. R. (2011). Contacts between police and the public, 2008. *Washington, D.C*.

Eckner, J. T., Kutcher J. S., & Richardson J. K. (2010). Pilot evaluation of a novel clinical test of reaction time in National Collegiate Athletic Association Division I football players. Journal of Athletic Training 45(4), 327-333.

Engel, R.S., & Silver, E. (2001). Policing mentally disordered suspects: A reexamination of the criminalization hypothesis. *Criminology, 39,* 225-252.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic medicine*, *79*(10), S70-S81.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual review of psychology*, *47*(1), 273-305.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

Ericsson, K. A., & Oliver, W. L. (1989). A methodology for assessing the detailed structure of memory skills. *Acquisition and performance of cognitive skills*, 193-215.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological review*, *87*(3), 215.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In Ericsson, K.A. & Smith, J. (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1-38). Cambridge: Cambridge University Press.

Festinger, L. (1971). Eye movements and perception. In P. Bach-y-Rita, C.C. Collins, & J.E. Hyde (Eds.) *The control of eye movements* (pp. 259-273). Academic Press New York.

Fitts, P.M., & Posner, M.I. (1967). *Human performance.* London: Prentice Hall.

Friedrich, R.J. (1980). Police use of force: Individuals, situations, and organizations. *Annals of the American Academy of Political and Social Sciences, 452*, 82-97.

Fritsch, J. (2000). The Diallo verdict: The overview; 4 officers in Diallo shooting are acquitted of all charges. *The New York Times*, *26*.

Fyfe, J.J. (1979). Administrative interventions on police shooting discretion: An empirical examination. *Journal of Criminal Justice, 7,* 309-324.

Fyfe, J.J. (1980). Geographic correlates of police shooting: A microanalysis. *Journal of Research in Crime and Delinquency*, *17,* 101-113.

Fyfe, J.J. (1988). Police use of deadly force: Research and reform. *Justice Quarterly, 5,* 165-205.

Galton, F. (1869). *Hereditary genius: an inquiry into its laws and consequences*. Macmillan and Company.

Garner, J.H., Maxwell, C.D., & Heraux, C.G. (2002). Characteristics associated with the prevalence and severity of force used by police. *Justice Quarterly, 19,* 705-746.

Garner, J.H., Schade, T., Hepburn, J., & Buchanan, J. (1995). Measuring the continuum of force used by and against the police. *Criminal Justice Review, 20,* 146-168.

Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory studies*, *12*(4), 267-284.

Geller, W., & Scott, M. (1992). *Deadly force: What we know.* Washington, DC: Police Executive Research Forum.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, *103*(4), 650.

Graham v. Connor, 490 U.S. 386 (1989).

Green, A. (1995). Verbal protocol analysis. *The Psychologist 8*(3): 126-129.

Grossman, D., & Christensen, L. W. (2008). *On combat: The psychology and physiology of combat*.

Gustin, W. C. (1985). The development of exceptional research mathematicians. In BS Bloom (Ed.) *Developing talent in young people*, 19-67. New York: Ballantine.

Haskins, M. J. (1965). Development of a response-recognition training film in tennis. *Perceptual and Motor Skills*, *21*(1), 207-211.

Hickman, M.J., Piquero, A.R., & Garner, J.H. (2008). Toward a national estimate of police use of nonlethal force. *Criminology and Public Policy, 7,* 563-604.

Hoffman, R.R. (1992). *The psychology of expertise: Cognitive research and empirical AI.* New York: Springer-Verlag.

Holmes, M.D. (2000). Minority threat and police brutality: Determinants of civil rights criminal complaints in US municipalities. *Criminology, 38*(2), 343-367.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1985). Interactive processes in perceptual organization: Evidence from visual agnosia. *Attention and performance XI,* 301-318.

International Association of the Chiefs of Police (IACP). (2001), *Police Use of Force in America, 2001*. Alexandria, Virginia, 2001.

Johnson, R. (2011). Suspect mental disorder and police use of force. *Criminal Justice and Behavior, 38*(2), 127-145.

Johnson v. Glick, 481 F. 2d 1028 (2d Cir. 1973).

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology*, *8*(4), 441-480.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263-291.

Kaminski, R. J., Edwards, S. M., & J. W. Johnson. (1999). Assessing the incapacitative effects of pepper spray during resistive encounters with the police. *Policing: An International Journal of Police Strategies and Management*, *22*, 7-29.

Kalinowski, A. G. (1985). The development of Olympic swimmers. In BS Bloom (Ed.) *Developing talent in young people*, 139-192. New York: Ballantine.

Kaufman, J. (2013). *The First 20 Hours: How to Learn Anything... Fast!* Penguin.

Klinger, D.A. (1997). The micro-structure of nonlethal force: Baseline data from an observational study. *Criminal Justice Review*, *20,* 169-186.

Klinger, D. A., & Brunson, R. K. (2009). Police officers' perceptual distortions during lethal force situations: Informing the reasonableness standard. *Criminology & Public Policy*, *8*(1), 117-140.

Kolers, P. A. (1975). Specificity of operations in sentence recognition. *Cognitive Psychology*, *7*(3), 289-306.

Kruschke, J. (2010). *Doing Bayesian data analysis: A tutorial introduction with R*. Academic Press.

Lawton, B.A. (2007). Levels of nonlethal force: An examination of individual, situational, and contextual factors. *Journal of Research in Crime and Delinquency, 44*, 163-184.

Lee, H., Jang, H., Yun, I., Lim, H., & Tushaus, D. (2010). An examination of police use of force utilizing police training and neighborhood contextual factors: A multilevel analysis. *Policing: An International Journal of Police Strategies & Management*, *33*(4), 681-702.

Lesgold, A. M. (1984). Human skill in a computerized society: Complex skills and their acquisition. *Behavior Research Methods, Instruments, & Computers*, *16*(2), 79-87.

Lewinski, W., & Hudson, B. (2003a, September/October). Time to start shooting? Time to stop shooting? The Tempe Study. *Police Marksmen, 7,* 26-29.

Lewinski, W., & Hudson, B. (2003b, November/December). The impact of visual complexity, decision making and anticipation. *Police Marksmen*, 28, 24-27.

Litman, L. & Davachi, L. (2008). Distributed learning enhances relational memory consolidation. Learn. Mem., 15, 711-716.

Luce, R. D. (1986). Response times: Their role in inferring elementary mental organization. New York, NY: Oxford University Press.

Marion, N. (1998). Police academy training: Are we teaching recruits what they need to know? *Policing: An International Journal of Police Strategies & Management, 21*(1), 54-79.

Massaro, D. (1975). *Experimental psychology and information processing.* Chicago: Rand McNally College Processing.

Mastrofski, S. D., Reisig, M. D., & McCluskey, J. D. (2002). Police disrespect toward the public: an encounter-based analysis. *Criminology*, *40*(3), 519-552.

McElvain, J.P., & Kposowa, A.J. (2008). Police officer characteristics and the likelihood of using deadly force. *Criminal Justice and Behavior, 35,* 505-521.

Meehan, N. C., Strange, C., & McClary, M. (2015). *Behavioral indicators during a police interdiction* (No. NRL/MR/5508--15-9598). Naval Research Lab Washington Dc Information Technology Div.

Meyer, G. (1992). Nonlethal weapons vs. conventional police tactics: Assessing injuries and liabilities. *The Police Chief, 59*, 10-17.

Morrison, G.B. (2006). Police department and instructor perspectives on pre-service firearm and deadly force training. *Policing: An International Journal of Police Strategies & Management, 29*(2), 226-245.

Moulton, C. A. E., Dubrowski, A., MacRae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching surgical skills: What kind of practice makes perfect? A randomized, controlled trial. *Annals of surgery*, *244*(3), 400.

Murray, K. R. (2004). *Training at the speed of life: The definitive textbook for police and military reality based training*. Gotha, FL: Armiger Publications.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, *1*.

Niederhoffer, A. (1967). *Behind the shield: The police in urban society*, 103-151. New York: Doubleday.

Paoline, E. A., & Terrill, W. (2007). Police education, experience, and the use of force. *Criminal Justice and Behavior*, *34*(2), 179-196.

Payne, B. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*, 181-192.

Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: A critical look. In In K.A. Ericsson and J. Smith (Eds.). *Towards a general theory of expertise: Prospects and limits* (93-123). Cambridge: Cambridge University Press.

Petrakis, E. (1986). Visual observation patterns of tennis teachers. *Research Quarterly for Exercise and Sport*, *57*(3), 254-259.

Petrakis, E. (1987). Analysis of visual search patterns of dance teachers. *Journal of Teaching in Physical Education*, *6*(2), 149-56.

Phelps, R. H., & Shanteau, J. (1978). Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance*, *21*(2), 209-219.

Ripoll, H. (1991). The understanding-acting process in sport: The relationship between the semantic and the sensorimotor visual function. *International Journal of Sport Psychology*. *22*: 221-243.

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: a meta-analysis. *Frontiers in psychology*, *5*.

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, *35*(6), 481-498.

Rosenbaum, D. A. (2009). *Human motor control*, San Diego CA: Academic Press.

Ruch, T.C. (1965). Vision. In T.C. Ruch, & H.D. Patton (Eds.) *Physiology and biophysics.* Philadelphia: Saunders.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & cognition*, *17*(6), 759-769.

Salmela, J. H., & Fiorito, P. (1979). Visual cues in ice hockey goaltending. *Canadian Journal of Applied Sport Sciences*, *4*(1), 56-59.

Shadish, W.R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Simon, H.A. & Chase, W.G. (1973). Skill in chess. *American Scientist*, *61*(4), 394-403.

Singer, R. N., Cauraugh, J. H., Chen, D., Steinberg, G., Frehlich, S. G., & Wang, L. (1994). Training mental quickness in beginning/intermediate tennis players. *Sport psychologist*, *8*.

Smith, D.A. (1986). The neighborhood context of police behavior. Reported in Reiss, A.J., & Tonry, M. (Eds.) *Communities and Crime*, University of Chicago Press, Chicago, IL: 313-341.

Smith, M., Kaminski, R., Alpert, G., Fridell, L., MacDonald, J., & Kubu, B. (2010). *A multi-method evaluation of police use of force outcomes: Final report to the national institute of justice.* U.S. Department of Justice. National Institute of Justice.

Smith, M.R. & Petrocelli, M. (2002). The effectiveness of force used by police in making arrests. *Police Practice and Research*, *3*, 201-215.

Sokolov, E. N., Spinks, J. A., Näätänen, R., & Lyytinen, H. (2002). *The orienting response in information processing*. Lawrence Erlbaum Associates Publishers.

Solomon, R.M., & Horn, J.H. (1986) Post-shooting traumatic reactions: A pilot study. In J.T. Reese, & H.A. Goldstein (Eds.), *Psychological services for law enforcement officers.* Washington, D.C.: U.S. Government Printing Office.

Sosniak, L. A. (1985). Learning to Be a Concert Pianist. In BS Bloom (Ed.) *Developing talent in young people*, 19-67. New York: Ballantine.

Strote J., & Hutson, H.R. (2006). Taser use in restraint-related deaths. *Prehospital Emergency Care, 10*, 447-450.

Tennessee v. Garner, 471 U.S. 1 (1985).

Terrill, W. (2009). The elusive nature of reasonableness. *Criminology and Public Policy*, *8*, 163-172.

Terrill, W., Leinfelt, F.H., & Kwak, D.H. (2008). Examining police use of force: A smaller agency perspective. *Policing: An International Journal of Police Strategies and Management*, *31*, 57-76.

Terrill, W., & Mastrofski, S.D. (2002). Situational and officer-based determinants of police coercion. *Justice Quarterly, 19*, 215-248.

Terrill, W., & Paoline , E. (2012). Examining less lethal force policy and the force continuum from a national use-of-force study. *Police Quarterly, 16*(1), 38-65.

Terrill, W., & Reisig, M.D. (2003). Neighborhood context and police use of force. *Journal of Research in Crime and Delinquency*, *40*(3), 291-321.

Tetlock, P. E. (1991). An alternative metaphor in the study of judgment and choice: People as politicians. *Theory & Psychology, 1*(4), 451-475.

Thomas, K.J., Collins, P.A., & Lovrich, N.P. (2010). Conducted energy device use in municipal policing: Results of a national survey of policy and effectiveness assessments. *Police Quarterly, 13,* 290-315.

Treisman, A.M. (1999) Feature binding, attention, and object perception. In G.W. Humphreys, J. Dundan, & A. Treisman (Eds), *Attention, space and action.* Oxford: Oxford University Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97-136.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, *79*(4), 281.

Tyldesley, D. A., Bootsma, R. J., & Bomhoff, G. (1982). Skill level and eye movement patterns in a sport oriented reaction time task. *Motor learning and movement behavior: contribution to learning in sport*, 290-296.

Ungerleider, L.G., & Mishkin M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Boston: MIT Press.

Vickers, J.N. (1988). Knowledge structure of elite-novice gymnasts. *Human Movement Science 7*, 4-72.

Vickers, J. N. (1992). Gaze control in putting. *Perception*, *21*(1), 117-132.

Vicente, K.J., & Wang, J.H. (1998) An ecological theory of expertise effects in memory recall. *Psychological Review, 105,* 33-57.

Von Der Malsburg, C. (1995). Binding in models of perception and brain function. *Current opinion in neurobiology*, *5*(4), 520-526.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic bulletin & review*, *14*(5), 779-804.

White, M.D. (2001). Controlling police decisions to use deadly force: Reexamining the importance of administrative policy. *Crime and Delinquency*, *47*(1), 131-151.

Williams, H.E. (2008). *TASER electronic control devices and sudden in-custody death: separating evidence from conjecture*. Charles C Thomas Publisher.

Williams, A. M., & Burwitz, L. (1993). Advance cue utilization in soccer. *Science and football II*, *2*, 239-244.

Wolfe, J.M. (1998). Visual Search. In H. Pashler (Ed.), *Attention.* Hove, UK: Psychology Press.

Worden, R.E. (1995). The "causes" of police brutality: Theory and evidence on police use of force. In W.A. Geller & H. Toch (Eds.), *Police violence: Understanding and controlling police abuse of force* (pg. 23-51). Washington, DC: Police Executive Research Forum.