A STUDY COMPARING THE EYEWITNESS ACCURACY OF POLICE OFFICERS AND CITIZENS

by

JOHN C. DECARLO

A dissertation submitted to the Graduate Faculty in Criminal Justice in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2010

This manuscript has been read and accepted for the
Doctoral Faculty in Criminal Justice in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Jennifer E. Dysart, Chair of Examining Committee
Date	Joshua Freilich, Executive Officer

Dennis Jay Kenney

Jon M. Shane

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

A STUDY COMPARING THE EYEWITNESS ACCURACY OF POLICE OFFICERS AND CITIZENS

by

John C. DeCarlo

Adviser: Professor Jennifer E. Dysart

Police officers are not only responsible for administering lineups and interviews where citizens' eyewitness memories are tested, they are also called upon to make arrests, write reports, prepare warrants and testify in court based on their own memories. Do police make better eyewitnesses than citizens? This study hopes to partially answer that question while contributing to the body of eyewitness research on weapon focus effect (WFE) and the new area of inquiry on trying to understand witness accuracy for multiple perpetrator crimes. This study investigates the effects of WFE and the presence of multiple perpetrators on eyewitness memory in two separate experiments. Two groups, police officers and citizens, were tested and compared in each experiment. One of the reasons that have been theorized to explain the presence of WFE is that weapons might hold a certain amount of contextual relevance or novelty that draws the attention of the observer when the weapon is present in a scene. Since police officers are commonly exposed to weapons and receive training on de-escalation of multi-person conflicts, the current study attempted to determine whether citizens and police would perform in similar or dissimilar ways to situations which might inhibit the observation and encoding of crime scene elements into memory. Two experiments were conducted to measure the arousal level of participants and assess the accuracy of police and citizen identification decisions in situations that potentially divert attention from the perpetrator in a simulated crime. Experiment 1 examined the effects of the presence of a weapon and the "weapon focus effect." The results showed that police officers tested lower on certain factors associated with stress and arousal than citizens but both police and citizens made more errors when a weapon was inferred or present. This is the first time that the inferred weapon condition has been experimentally explored. In addition, police made fewer filler identifications when the lineup target was present than when absent. Experiment 2 tested whether the presence of two culprits instead of one culprit affected identification rates. Both police and citizens experienced an increase in two factors that are associated with stress and arousal. Additionally, both police and citizens' identification accuracy was lower in the presence of two culprits; no accuracy differences were found between the police and citizen group. The results will add to the body of literature in eyewitness identification and contribute to the understanding of how stress or anxiety may or may not affect identification accuracy. In addition, it is hoped that elements of the study will be useable in police training contexts to help understand and improve the way that eyewitness evidence is processed and used by law enforcement agencies.

Acknowledgements

First and foremost, I would like to thank my wife, Catherine, and son, Evan, for enduring all the times that we had to put family time on hold during the years I was in the doctoral program. I have been extremely lucky to enjoy their support, patience and love over these last several years. I have a truly wonderful family to whom I owe a long, restful vacation.

Dr. Jennifer Dysart has acted as a mentor and friend during my time at John Jay College. I owe her a tremendous debt of gratitude for the constant guidance that she has provided and for the many thoughtful insights from her vast knowledge of eyewitness psychology that she has shared with me.

My committee, Dr. Dennis Kenney and Dr. Jon Shane have been generous with both their camaraderie and the help they have provided to me. John Jay College is a unique environment where many, whose passion is justice, have come together to share their thoughts and provide a place that is vital to the constant improvement of the criminal justice system.

List of Tables

List of Figures

Table of Contents

A Study Comparing the Eyewitness Accuracy of Police Officers and Citizens

Introduction

Police as Eyewitnesses: Some Examples

On September 2nd, 2009, in Cape Girardeau, Missouri, a police officer identified Ausha Sharp as a felony suspect. The officer reported that he witnessed and identified Sharp as she drove a getaway car for Zatrun Twiggs as the two fled police in a high speed chase. Based on the officer's observations, a warrant was issued for Sharp's arrest for hindering prosecution and resisting a lawful stop (DiCosmo, 2009). When Sharp read in the newspaper the next day that she was wanted, she went to police with an explanation and an airtight alibi.

It was later learned by police that an acquaintance of Sharp was the actual getaway driver who accompanied Twiggs during the pursuit. The prosecutor in the case, Morley Swingle, said after charges against Ausha Sharp were dropped, "The two women [were] similar in appearance and an officer mistakenly identified the woman who drove the car during the chase as Sharp instead of her friend… [T]his almost never happens, and I'm embarrassed when it does."

In LaCrosse, Georgia a woman was identified by police as the perpetrator of a larceny using a stolen credit card at a convenience store. A friend of the accused saw her in a video on a local crime stoppers TV show and told her that she was wanted. Further investigation showed that the woman had been confused with the real perpetrator by police who had viewed videotapes from the store where both the actual criminal and the innocent shopper had been present. When contacted by the press about the misidentification, LaCrosse Police officer Drew Gavrilos said, “We had two very similar looking people come to the register about 10 to 20 seconds apart, each buying similar things. It was a one-in-a-million type of thing.” Police sent the woman, who was not named, a note of apology for their incorrect identification (Associated Press, 2008).

In 1969, Laszlo Virag was arrested in Liverpool, England on two offenses of theft from parking meter coin boxes and for using a firearm to resist arrest when he wounded a constable (British House of Commons, 1974). Police found 17 eyewitnesses who had seen Virag commit the thefts, fire his weapon at police officers and subsequently escape. Six of the witnesses were police officers who had taken part in the fracas that ensued while trying to unsuccessfully apprehend the alleged thief turned cop shooter. Five officers and three citizen witnesses felt sure enough to pick out Virag from a photo lineup. One of the witness/police officers commented in an official report “[H]is face is imprinted on my brain” (Devlin, 1976). Subsequently, another man was arrested for an unrelated matter and found irrefutably to have committed the crimes for which Virag was convicted. Virag was exonerated and given an official apology by the government. One of the officers who had earlier identified Virag as the perpetrator went on record to say “[That the two men were] in no way similar in appearance and it is difficult to believe that all these witnesses could have been mistaken” (Bruce, 1988, p. 8).

Police Eyewitness Credibility and Accuracy

Several studies have found that police officers are accorded great credibility when they testify in court (Leippe, 1994). Approximately 66% of participants in a study by Yarmey and Jones (1983) believed that police officers would be superior eyewitnesses and better able to make identifications than citizens because of the training they receive and experience they gain on the job. Even though there is a belief that police officers are better witnesses than the general public, very little empirical research has been conducted to learn if the belief in officers' greater eyewitness credibility is based on fact. Indeed, some studies, such as Krouse (1981), have found that police were less capable than citizens at remembering faces after two or three days of delay after initial exposure. One of the questions to be explored in the current research is whether police officers do or do not demonstrate better eyewitness performance than citizens. Even though they are often attributed more credibility than members of the general public, are police more accurate at remembering crime details and perpetrators? Ultimately, the answer to this question is important because police are often called upon to identify criminals and make eyewitness observations. The accuracy of police memory is vital to the criminal justice system since it is inexorably linked to correctly identifying suspects and accurately reporting the events that lead to the arrests of those believed to have committed crimes.

Eyewitness Fallibility

Over the last 30 years, a body of literature has grown which clearly demonstrates that eyewitness memory is subject to many factors that can cause inaccuracy (Loftus, 2005a; NIJ Technical Working Group for Eyewitness Evidence, 1999; Penrod, Fulero, & Cutler, 1995; Phillips, 2001; Wells, et al., 2000). The Innocence Project (2010a) reports 75% of the DNA exonerations that have taken place across the country had been the result of erroneous eyewitness identification. False identifications not only confound police investigations in their very early stages, potentially influencing other evidence

(Hasel & Kassin, 2009), they also allow the actual perpetrator of a crime to remain free to commit other crimes.

Although a great number of experiments on eyewitness recall and recognition have been conducted over the last 30 years, police officers are an understudied group, despite the fact that they often observe crime and act as witnesses (Hulse & Memon, 2006). Further study of police as witnesses, such that is conducted here, could make a difference in law enforcement and judicial policy and give officers and courts the ability to give more accurate weighting and consideration to eyewitness reports including those made by law enforcement.

Statement of the Problem

When a crime is committed, police are routinely called upon to take statements, prepare reports, write warrants for searches and arrests, and make warrantless arrests, based largely on their observational skills and memory, usually under stressful circumstances. Ultimately, police are asked to testify in court about their observations as eyewitnesses to criminal activity and collect eyewitness information from victims and witnesses. Officers do all of this with little or no special training in eyewitness administration, memory, facial recognition, or other factors that might enhance or inhibit their efficacy in this role (Carden, 1989; Yuille, 1984).

Eyewitness identification is important to American jurisprudence because of its commonality and its perceived and often misconstrued accuracy. The implications of misidentifying a suspect are grave: Innocent citizens serving sentences for crimes they did not commit while the guilty remain free. This study seeks to answer whether differences exist in the eyewitness accuracy of police officers and citizens. Police officers represent government, if a lay person gets an identification wrong that is one thing. If a police officer gets it wrong, then that calls into question the veracity and credibility of the police as an institution. Society accords the police great deference so it is important to know both their limitations as well as their strengths.

Parameters of the Current Study

The current study, which is comprised of two experiments, conducts research on how police eyewitness abilities are affected by certain situational factors including the presence of a weapon and the presence of multiple offenders. These results will be compared with those of citizen eyewitnesses. If a difference between these groups is found, it may be possible to better understand and disaggregate certain situational impediments to eyewitness memory such as the weapon focus effect (WFE) and multiple perpetrator crimes. WFE is a phenomenon that refers to the concentration of a witness's or victim's attention on a weapon to the exclusion of other elements of a crime scene. It is not fully understood whether this concentration is caused because the weapon is threatening or merely novel. The presence of more than one perpetrator during the commission of a crime may adversely affect a witness's ability to make a correct identification of either the culprit or the accomplice at a later time. Like WFE, the exact cause(s) underlying lower identification rates in multiple perpetrator crimes are not completely understood. Both WFE and multiple perpetrators are discussed in greater detail in the literature review.

General Definitions

To fully understand the details of the study, it will be helpful to touch briefly on several concepts that are foundational elements in eyewitness identification research.

The first concept, which is relevant to any study regarding eyewitness memory, is what are known as system and estimator variables (Wells, 1978). Other elements that will be examined in this definition section are simultaneous and sequential lineups, target present and target absent lineups, and lineup diagnosticity

Estimator and system variables. When looking at these two bedrock elements of eyewitness research for the first time, they seem deceptively simple. Estimator variables describe the elements that make up the contextual and environmental setting of a crime scene. Estimator variables, rather than being predictive, are postdictive. This means that they are used after a witness has made an identification in an attempt to estimate the accuracy of the witnesses' memory. They include random, unsystematic factors such as the victim's age, victim's and offender's races, scene illumination, presence of weapons, number of offenders, victim's stress level, clothing and disguises, etc. (Wells, Memon, & Penrod, 2006). Witness confidence in their identification decision is another example (Smith, Lindsay, Pryke, & Dysart, 2001). A close consideration of the list reveals that these are variables that have great potential for change and are non-controllable by the criminal justice system.

One estimator variable that is particularly relevant to the current study is WFE. Recent research has shown that both the presence of a weapon and a witness's physiological arousal may be distinct parts of the WFE that combine to form a greater whole, (Hagan, 2009; Hope & Wright, 2007; Pickel, 2007, 2009; Pickel, Lindsay, Ross, Read, & Toglia, 2007; Steblay, 1992). In essence, the attention that is paid almost exclusively to a weapon when one is present, produces a psychological tunnel vision-like effect which blocks out the relative richness of peripheral features (Cooper, Kennedy, Hervé, & Yuille, 2002). Findings from research have suggested several causes for the attentional deficit that seems to occur during weapon focus. When moderately aroused, people will sometimes have a more directed focus on a task in which they are engaged. However, when over-arousal takes place, such as when a weapon is present, our observational and memory performance might suffer (Easterbrook, 1959; Tooley, Brigham, Maass, & Bothwell, 1987).

System variables, on the other hand are subject to being influenced by elements of the criminal justice system. Examples include the method of lineup presentation (e.g., live vs. photo, simultaneous vs. sequential) and the use of a double blind system in which neither the witness or the lineup administrator knows if the suspect is in the lineup or their position in the array (Wells, et al., 1998). Although how quickly a witness is able to view a lineup after a crime might be initially thought of as a system variable, is normally considered and estimator variable. Although this variable appears as though it could be strictly controlled for by the system, there are practical considerations that influence timing of a lineup procedure, including witness injury, police availability due to activity and staffing, technical considerations with photos, and finding a suspect. Certainly, the quickness with which the police show the lineup is considered a system variable but because of the considerations mentioned above, there are estimator elements which cause delays also. Thus, some aspects of the investigation may act as both system and estimator variables. Although time to the lineup was strictly controlled for in the current study, lineups taking approximately 15 minutes to be shown to the participants after they viewed the experiment videos, delays of up to a month have not been shown to significantly affect identification accuracy (Dysart, et al., 2007).

Simultaneous and sequential lineups. When conducting a "lineup", the vast majority of police agencies in North America do not bring live suspects into a room to be viewed by a witness (Turtle, Lindsay, & Wells, 2003). When conducting a lineup, most police departments use a photo-array consisting of a photograph of the suspect placed with photos of generally similar looking individuals. A simultaneous lineup is one in which all photos in the lineup are shown at once. In a sequential lineup, photos are shown to the eyewitness one at a time. There is support in the research literature that sequential lineups reduce the occurrence of misidentification of innocent suspects by encouraging witnesses to use absolute rather than relative judgments when making their lineup identification decision. Wells and Seelau (1995) explain that because of the relative judgment that witnesses utilize when viewing simultaneous lineups, witnesses compare lineup photos relative to each other, instead of to their memory of the perpetrator. This problem can be especially insidious if the perpetrator is not present in the lineup. In that case, the witness might be predisposed to select an innocent lineup member who most closely resembles the offender as opposed to rejecting the lineup altogether. Conversely, when witnesses view a sequential lineup, their decision making tends more toward absolute decisions because they must compare each successive photo to against the image they retain in memory of the perpetrator rather than the other photographs. The result is (often) a reduction in false identifications with this procedure.

Currently in the United States, most police departments utilize simultaneous lineups (Klobuchar, Steblay, & Caligiuri, 2006). As already mentioned, however, research has shown that sequential lineups result in fewer false identifications (Carlson, 2008). One of the possible reasons that simultaneous lineups are still in such wide use is that some studies have found that sequential lineups might also produce fewer true identifications of the perpetrator (Haw & Fisher, 2004). One such study is the Illinois Pilot Study (Mecklenburg, et al., 2006). However, this study has been criticized regarding methodological issues in the design of the study (Steblay, 2008; Steblay, Dysart, Fulero, & Lindsay, 2001; Wells, 2008; Winzeler, 2008) which have caused the study's conclusions to be highly disputed. These issues are discussed in detail in the discussion section of this study.

An additional factor, the importance of which is supported by the literature, is the advisory to a witness before they view a lineup. The admonition to a witness that the culprit may or may not be in a lineup is germane in setting the expectations of the witness (Phillips, 2001; Technical Working Group for Eyewitness Evidence, 1999; Turtle, et al., 2003; Wells, et al., 2000), with research showing a significant decline in false identifications when this warning precedes the viewing of the lineup (Connecticut State of, 2005).

Target-present and target-absent lineups. Police usually endeavor to include a person whom they believe is the perpetrator of a crime into a lineup (Wells & Olson, 2003). The construction of target-absent lineups is usually considered an anomaly in police field work. Reasons why the true perpetrator may not be included in a lineup might be a consequence of police not including that person due to having received inaccurate information from the witness. Other factors that might cause this condition are that police investigators have a belief that a person other than the actual culprit is the perpetrator of the crime. When the perpetrator is not present in the lineup the lineup is known as "target absent" (Steblay, Dysart, Fulero, & Lindsay, 2003) and is usually not a circumstance that is intended in a police investigation unless the lineup is intended to be exculpatory in nature, which most lineups are not.

The majority of police lineups are intended by investigators to be "target present" which is a lineup where the suspect is included. The other photos in a lineup are called fillers or foils. They are generally selected to be close in general appearance to the suspect's photo. As an example, if the suspect is a brown haired white male with a moustache, then the filler photos should also be brown haired mustachioed white males.

When an eyewitness is called upon to view a target-present lineup in a laboratory/ research situation, they can make a correct identification in which they select the actual "culprit", a filler identification, or an incorrect rejection of the lineup. Under field conditions the outcome is not as cut and dry because police rarely know before a lineup if the suspect is the actual perpetrator of the crime being investigated.

Diagnosticity. The last of the several foundational elements to be discussed regarding eyewitness memory is the *diagnosticity* of suspect identifications in lineups (Clark & Wells, 2008; Malpass & Devine, 1981; Navon, 1990; Tredoux, 1998; Wells & Lindsay, 1980; Wells & Luus, 1990). Diagnosticity is a form of Bayesian statistics dealing with conditional probabilities. It asks the question: If a witness has made a lineup identification, what is the likelihood that the suspect is guilty? Wells and Luus (1990, p. 1) defined lineup diagnosticity "[A]s the probability that an innocent suspect is identified by the witness when lineup members resemble the eyewitness's pre-lineup description." Diagnosticity is defined by the *ratio* of correct lineup identifications in target-present lineups to the number of incorrect identifications when the target, or suspect, is absent from the lineup (Wells & Olson, 2003). One way of calculating diagnosticity is by dividing the correct identification rate for target-present lineups by the average false identification rate in the target-absent lineup (Owens, 2009).

The formula for calculating Diagnosticity is $\frac{P(ids|s=c)}{P(ids|s\neq c)}$ where ids = identifications; s= suspect and c= criminal. Table 1 illustrates how the diagnosticity ratio (DR) calculation is carried out. It shows the distribution of witness photo number selections on 4 lineups. Lineup A1 and Lineup B1 both contained photos of a guilty culprit in position 5 (target present condition). Lineups A2 and B2 were both target absent. The target in the A group is 14 times more likely to be chosen when he is guilty than when he is innocent. The B Group's diagnosticity value is 1.59 meaning that the culprit is only 1.59 times more likely to be chosen when guilty than

Table 1. *Example of Diagnosticity Ratios (DR) in Sample Photo Lineups*

Photo number (* indicates target photo)

Criminal		1	2	3	4	5*	6	7	8	N/P	N	Proportion	DR
1	P	7	12	2	11	56	20	9	10	3	130	0.430	14.33
	A	11	7	18	20	4	6	2	5	57	130	0.030	
2	P	4	3	18	13	51	8	14	11	8	130	0.392	1.59
	A	12	9	14	16	32	20	12	9	6	130	0.246	

Note: P=criminal present lineup. A= criminal absent lineup. Criminal 1 has a diagnosticity ratio of 14.33 (56/130)/(4/130). Criminal 2 has a diagnosticity ratio of 1.59 (51/130)/(32/130). N/P indicates that no choice was made.

when innocent. The B group has much lower diagnosticity than the A group. The higher the DR, the more discriminant, and less error prone, the lineup It is important to note that diagnosticity ratio is the percentage of correct responses in TP to incorrect choices in TA lineups. Since, under actual conditions in police departments, lineups are not used over again many times, diagnosticity ratio is only useful as an analytical tool and not in the real world.

Review of the Literature

Since the current study concerns itself with the possible differing abilities of police officers and citizens to act as eyewitnesses, several areas of the literature are reviewed. First general eyewitness memory research will be reviewed, which will serve to set a foundation to build on other topics that are relevant to the current study. This will be followed by examinations of Weapon Focus Effect, and the literature concerning multiple offenders and their affect on eyewitness performance.

General Research on Eyewitness Memory and DNA Exonerations

Wells (1993) observes that three general areas support the fact that a problem exists with eyewitness evidence. First, a large body of experimental research using mock crimes has found that false identifications occur at a high rate. Secondly, witnesses who make these false identifications affect their mistakes with a high degree of sincerity making them excellent and believable witnesses in the courtroom. Third, analysis of over 1,000 court cases, from as far back as 1973 where innocent people had been convicted and then cleared through subsequent investigation, listed eyewitness evidence as the biggest cause of the convictions (Brandon & Davies, 1973).

After two decades of systematic research, the tenuous nature of eyewitness identification was underscored in 1989 (Wells, 1993). In that year, the first exonerations by DNA were conducted, with literally hundreds to follow. For the first time, research findings were supported by real world evidence of the fallibility of eyewitness evidence. Huff (2004) reported that the results of a survey of 353 state level attorneys general, judges, police officers and prosecutors showed that the belief among 79% of the respondents was that incorrect eyewitness identification was the leading cause of wrongful convictions. Huff's study was especially telling when the conservative nature of the practitioners surveyed is taken into account. These were people who worked in the system and were able to see the problems with eyewitness evidence firsthand. In a 2010 report, The Innocence Project reported that 75% of the individuals that had their convictions overturned by DNA evidence had originally been sentenced using eyewitness identification as the prosecutors' primary tool. Unless there is a sea change in the way the criminal justice records are kept, and every wrongful conviction based on erroneous eyewitness information is caught and overturned, it is impossible to know accurately how many incorrect eyewitness identifications are made each year. Huff estimated a rate of incorrect eyewitness identifications of about a half of a percent. Out of the approximately 14 million felony arrests in the United States in 2008 (Bureau of Justice Statistics, 2008), there were about 1,051,000 convictions. Based on Huff's estimate, the potential for wrongful conviction due to faulty eyewitness evidence is over 5,000 people per year. Wells (1993) reports that even under optimal conditions, with all system variables controlled for, the base error rate for eyewitness convictions would still be approximately one in ten.

Although eyewitness identification has limitations, it still serves as a procedure that is used by law enforcement officers every day and is fundamental to the criminal justice system. Eyewitness identification serves as the sole type of evidence against a defendant in many criminal cases (Yarmey, 2004). Therefore, the chance of eyewitness identification disappearing as an evidentiary tool, despite its drawbacks, is slim. Eyewitness research is vitally important, since making the process less error-prone by being aware of estimator variables and managing system variables might significantly reduce the probability of misidentifications.

National Institute of Justice Guidelines

Wells (1993), in his seminal paper, *What do we know about eyewitness identification*, observed that psychological research is no longer so concerned notionally with the fact that eyewitness evidence can be unreliable so much as finding answers to what conditions can foster more consistent accuracy in the process. In 1999, the National Institute of Justice published a set of guidelines for law enforcement (NIJ Technical Working Group for Eyewitness Evidence, 1999). The recommendations that were made at that time were based on the then current state of research. Section V of that report, *Procedures for eyewitness identification of suspects*, included suggested policies on the composition of lineups, instructing the witness before viewing the lineup, conducting the identification procedure and recording the results.

The recommendations about the composition of the lineup now seem commonplace to police. The group responsible for the guidelines suggested that only one suspect be included in any lineup. This was consistent with what Wells (1993) had described as the single-suspect model. This type of lineup fills all but the target photo with innocent filler photos. In addition, the guide recommended that fillers generally fit the witness' description of the perpetrator so that the suspect does not stand out from the group by being the only one who matches the witness' description.

The section in the NIJ Guide that addressed instructing the witness prior to viewing a lineup included several recommendations. First is to set expectations that the lineup procedure is not an indictment. Officers were advised to tell the witness that the lineup might be exculpatory as well as inclusive. In addition, it was advised that the witness be told that the person who committed the crime may or may not be in the lineup. Where sequential and simultaneous lineups were concerned, the Guide stopped short of recommending either one of the procedures over the other. Also, it did not mention a double-blind procedure in which neither the lineup administrator nor the witness knew if the actual culprit was present in the lineup. Instead, it was suggested that for simultaneous lineups the investigator did not say anything which could possibly bias the witness and, among other recommendations, that sequential lineups be in a random order and all photos be shown even if an identification of one of the first photos was made.

A substantial amount of eyewitness research has been published since the publication of the Guide in 1999. The following sections of the literature review address some of that research which is relevant to the subject of the current study and that, in the future, might also inform practice through the inclusion of reports and recommendations to the criminal justice community. For the present, since the NIJ guidelines are what are being recommended to practitioners, the current study will adhere to them. We will first cover the research on the WFE followed by multiple perpetrator crimes.

Weapon Focus Effect - Theory

Historically, an increase in arousal has been associated with less accurate recall. As physiological stressors go up, recall tends to go down (Christianson, 1997; Cutler, 2004).

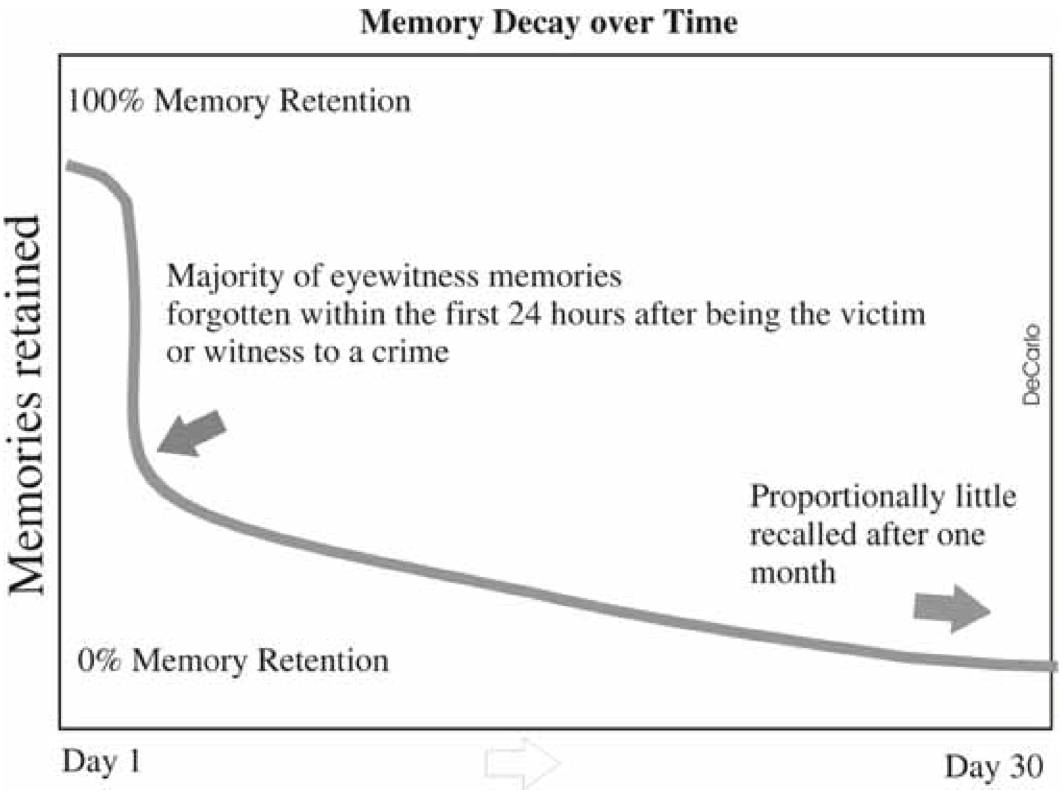


Figure 1. The Ebbinghaus Forgetting Curve applied to eyewitness memory

Certainly, the presence of a weapon being brandished by a criminal is threatening to feelings of well-being and might tend to contribute to heightened states of arousal and, consequently, poorer

memory. The Ebbinghaus forgetting curve comes under this category. Ebbinghaus (1885) conducted some of the first experiments on memory. He found that certain factors, including the difficulty of the material that the subject was attempting to memorize and physiological aspects, such as stress and lack of rest, affect memory retention. He theorized that memory naturally decays over time (See Figure 1).

Ebbinghaus speculated that the introduction of stress at the time a memory was initially being encoded would dilute the intensity of the memory and cause it to decay faster than memories learned under non-stressful conditions. Ebbinghaus' belief is consistent with the traditional view of memory in that forgetting was associated with associative bonds being weakened (Terry, 2000). The reason that the bonds might be weakened could be caused by processes that might interfere with or inhibit the encoding process.

Cue utilization theory. Cue utilization theory (Easterbrook, 1959) states that higher levels of arousal facilitate the processing of and response to stimuli (McCloskey & Egeth, 1983). Easterbrook first used his theory to describe memory performance in the realm of educational psychology. It was found that during an aroused state, school athletes tended to not pay attention to cues which were irrelevant to their performance, thereby focusing their mental energy on the task at hand. However, when over-arousal occurred, situationally relevant cues were also filtered out to the ultimate detriment of performance (See Figure 2). It is unknown whether this circumstance arises from a stimulation of the sympathetic nervous system, in a condition similar to a fight or flight response (Loftus, 2005b).

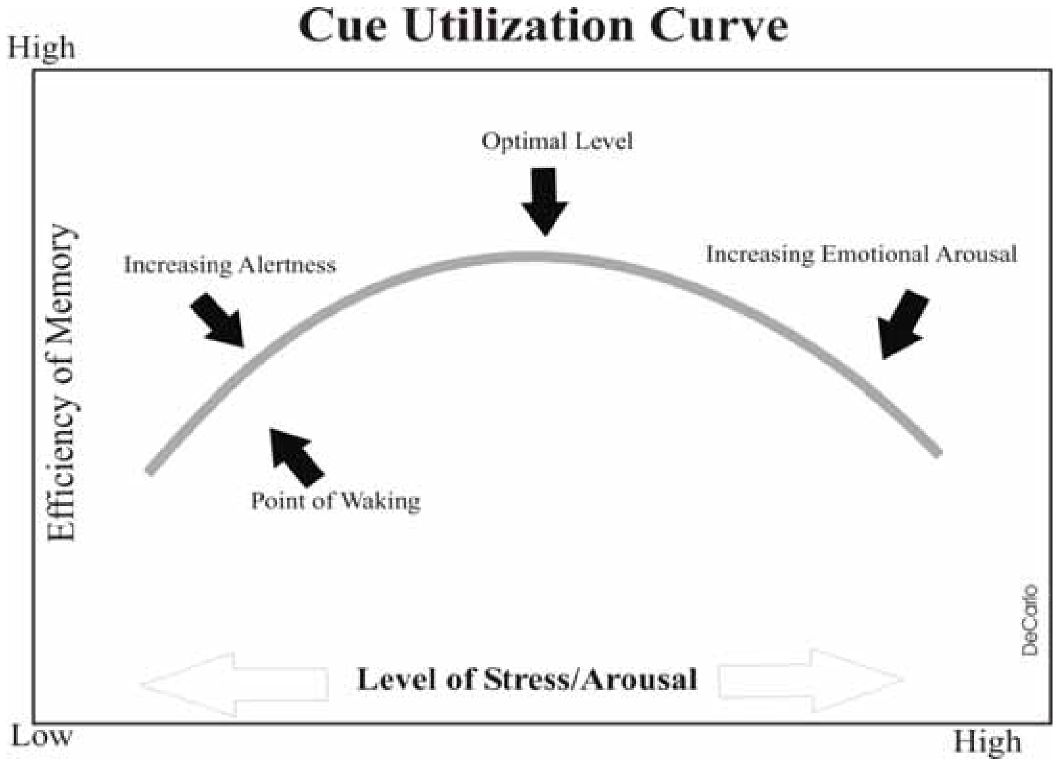


Figure 2. Visualization of Easterbrook's (1959) Cue utilization curve

Research conducted on the effect of stress on recall,(Kramer, Buckhout, Fox, Widman, & Tusche, 1991) in which participants were shown photos of autopsies that were mixed in with photos of travel scenery. Participants and control subjects showed no difference in retrograde recall before participants viewed the stress producing autopsy slide. After participants viewed the stressful autopsy slide, anterograde recall of the remainder of the slides they viewed significantly differed between participants and control subjects with control subjects recalling more than participants. A meta-analytic review of the effects of high stress on eyewitness memory (Deffenbacher, Bornstein, Penrod, & McGorty, 2004) reported that 27 independent tests of the effects of high stress had a negative impact on the recall of perpetrators or target persons in the tests. In addition, the effect sizes reported were larger for target present than for target absent lineups. The study also found that effect sizes were larger for studies employing a staged crime than for other methods that were used to create stress in participants.

Limited attentional capacity. Other reasons other than emotional arousal might be responsible for the suggested division of attention that the presence of a weapon causes. It was been observed by Revelle and Loftus (1990) that the mind is like a computer with only so many “CPU cycles”. The human analogue to this metaphor would occur when one is using a large amount of cognitive “cycles” to pay attention to a weapon. In this condition, fewer cycles would be available to devote to other particulars. Revelle and Loftus proposed a partial explanation for WFE which they called the “tick rate hypothesis.” Arousal's effect on memory again takes center stage in this theory. Revelle and Loftus described human attention as being analogous to the clock speed of a computer where there are only a finite number of cycles or “clock ticks” available for processing. They proposed that arousal increases the rate at which the environment is sampled and predicted that increased arousal leads to a faster “attentional tick rate” of response to environmental cues. Conversely, when the response to environmental cues is increased too much, it becomes a detriment to the availability of immediate memory. This model predicts that a high arousal at the time of observation should facilitate long-term retrieval. When the tick rate hypothesis model is associated with weapon effect it is found that the victim of a crime where a weapon is used often has greater long-term memory of the weapon rather than of peripheral elements of the crime scene, including their assailant. Currently, no additional research on this hypothesis has been published.

Weapon Focus Effect - Empirical Research

Research ethics constrains some methodologies that would seem to be suggested by theoretical research on the WFE. Certainly it would not be acceptable to conduct research with human subjects that substantially frightens participants or puts them in harm's way. Consequently, most empirical research on weapon focus is conducted under controlled laboratory conditions.

In their 1987 article, Loftus, Loftus and Messo defined and conducted the first empirical study of the WFE. The experiment, which consisted of two groups of university students, had five distinct parts. In the first experiment, participants observed two series of 18 photographic slides depicting a transaction involving several people at a fast food restaurant. In the control set of slides, a dinner check was handed to the clerk and in the second, known as the weapon condition, a gun was pointed at the clerk. After the slides were displayed, a questionnaire asking observers to describe estimator variables such as the race, age and general description of the person in the slides was completed. Next, a 12-person photo lineup was shown to participants and they were asked to identify the actor with the check or the gun, dependent on whether they were in the control or the weapon condition group.

On average, participants who had viewed the weapon condition slides were 5% less accurate in describing elements of the crime scene, such as the store environment and the description of the culprit. In the lineup stage of the experiment, 39% of the control group was able to pick the correct actor as opposed to only 11% in the weapon condition group. Analysis of the eye movements, as measured by corneal reflection eye tracking equipment, showed that the weapon group spent about 42 milliseconds longer fixated on the gun than the control group spent looking at the check. Similar results were reported in a second population using slightly different parameters.

Pickel (1999) conducted an experiment which investigated the influence of context on WFE. It was hypothesized that weapons are novel and would capture more of a witnesses' attention if viewed out of context. For example, a person viewed carrying a handgun at a delicatessen would theoretically draw more of an observer's attention to the weapon than would the same person at a shooting range. The context in which items are seen might be an important factor in how much awareness they draw away from the observation and memory encoding of other items in a scene. In Pickel's study, undergraduate students were shown a videotape which depicted a male carrying a weapon. The results showed that the descriptions of the armed man were less accurate if shown in a context in which a gun was not expected, such as an armed college student, than where one was expected to be seen, such as a police officer in uniform. In addition, eyewitnesses were worse at describing an individual if that person carried an object that was not consistent with their occupation. The results of the experiment imply that WFE may be a function of the novelty of weapons and because they are not expected in certain contexts. In addition, Pickel (2009) conducted a later study which examined the weapon focus effect for male versus male perpetrators. The study hypothesized that weapons were inconsistent with psychological activated by witnesses at the time of a crime. This condition would theoretically cause witnesses to focus more on weapons than on objects that were neutral. The study found that, as predicted, females holding handguns degraded eyewitness memory performance more so than when a male held a handgun. This study is a good example of context and expectations caused by psychological schema and how an out of context item not associated with preconception can refocus attention from a perpetrator and on to a weapon.

Hope and Wright (2007) conducted a similar experiment in which a threatening object, (weapon condition; gun), a novel object (unusual condition; multicolored feather duster) or a neutral object (control condition; man's wallet) was shown to participants. Results suggested that threatening objects command more attention and the significance of a weapon can lead to impaired performance on other aspects of a crime scene such as perpetrator description. The overarching question in the Hope and Wright study was whether WFE might be better understood by addressing the reasons that a witnesses' attention might be drawn to a weapon. Does the phenomenon of attentional diversion exist on a continuum that proceeds from an object being usual and in context to threatening and out of context? Do people, when exposed to a threatening object focus on it for some reason, to the partial exclusion of other factors in a scenario? It is suggested by the Hope and Wright study that witnesses paid more attention to weapons than the people that were carrying them. Participants were better at describing the weapon than the person holding the weapon.

The results of the Hope and Wright (2007) study set some of the groundwork for the current study. Are police officers, who handle weapons on a daily basis, less affected by the unusualness or novelty of a weapon than someone not inured to weapon presence? Will the results of the current study show that police officers are also affected by WFE and therefore suggest that threat rather than novelty is the condition at work as an attentional drain? Context and expected versus surprising conditions might all play a part in disambiguating the WFE.

A meta-analysis of the WFE was conducted by Steblay in 1992. Nineteen prior experiments on the WFE, using line-up accuracy as the dependent variable, were examined in the study. Steblay found that the presence of a weapon consistently led to lower identification accuracy. The data that was included in the meta-analysis was considered to be robust in WFE. It was noted that the experiments had been done with collegiate and non-collegiate populations and that WFE was evenly distributed over both groups in moderate to high arousal states. Steblay (1992, p. 422) reported that focus of attention, combined with high arousal produces a situation that "maximizes the potential for weapon focus effect." Steblay concluded by saying that the data was convincing that WFE did indeed exist and further research on the mechanics of the process would be relevant.

Weapon Focus Effect – Archival Research

Behrman and Davey (2001) conducted an experiment in which they analyzed 271 police cases to explore several prevalent issues in the eyewitness literature, including WFE. The researchers cited that other studies have supported the phenomenon of weapon focus. In their archival analysis of actual cases, there was no statistical significance in the difference between crimes that involved a gun and crimes that did not. The researchers opined that a possible explanation was that weapon focus might not be a real-life phenomenon, suggesting instead that the arousal level during a real crime is simply too high for there to be a substantive difference in the memory of non-weapon and weapon groups. In addition to the presence of a weapon, ambient conditions and lack of training/ experience could potentially raise stress levels to the point where the efficient encoding of information and decision making processes might be subject to interference.

These situations not only have the potential to inhibit clear memory of an event but, because the tunnel vision analogue and lessened situational awareness might cause officers to erroneously focus on a misperceived weapon and limit their ability to process other important cues.

Multiple Offender Research - Theory

Because inquiries into the effect that multiple offenders might have on eyewitness performance are a relatively new area of research, no central theory has been forwarded to explain the phenomena that might cause a witness's ability to make a correct identification if more than one perpetrator was present at the time the witnessed crime was committed. If there is a multiple perpetrator effect on memory, it is important that the magnitude and nature of the effect is understood in police identifications and related court proceedings so as to avoid incorrect identifications. Megreya and Burton (2006) suggest that although the facial recognition literature has helped explain applied issues in eyewitness research on single faces, there is no published theoretical research on viewing and encoding multiple faces simultaneously.

Three possible reasons why the number of offenders is of consequence to what an eyewitness is able to remember about more than one culprit involved in a crime, are diminution of attention, unconscious transference and stimulus sampling theory. It is obvious that when paying attention to more than one stimulus simultaneously, there must be a reduction in the amount of focus on any single object, as discussed in the WFE section above. It seems logical that this line of thought would extend to multiple offenders being viewed by an eyewitness.

Although division of attention is perhaps the most obvious impediment to accurate recall when multiple offenders are present, it is only one possible cause of confusion. Unconscious transference (Geiselman, Haghighi, & Stown, 1996; Kassin, Tubb, Hosch, & Memon, 2001; Phillips, Geiselman, Haghighi, & Lin, 1997) may also contribute to less eyewitness accuracy in multi-offender events. Unconscious transference is thought to happen when one image in memory becomes transferred to another image that is stored in memory. Functionally, this type if memory amalgamation might be compared to the meta-data which describes information such as name, type, content and the file's location(s) in memory that is recorded by a computer for each of its files. If the meta-data pointer for one file becomes inadvertently transferred, in whole or in part, to another file, there will be a mismatch of information. In a similar fashion, if the location pointer in a witness' memory meta-data becomes attached to another image, transference might occur.

Multiple Offenders - Empirical Research

Although there are substantial numbers of multiple offender crimes, the psychological literature on the effects of multiple perpetrators on eyewitness recollection, when compared to other areas of research, is relatively sparse. The main hypothesis of this line of research is that when two or more perpetrators are present at a crime, the victim will not remember either one of them as well as either if they had been alone (Brear, (2002); Clifford & Hollin, 1981; Dempsey & Pozzulo, (2008b); Owens, (2009); Wells & Pozzulo, (2006).

A 1981 study (Clifford & Hollin) combined the presence of multiple offenders and violence and found that eyewitness accuracy suffered to a greater degree than was observed in an earlier study by Davies et al. The authors of this study observed that some earlier researchers had speculated that as the number of offenders went up, accuracy would diminish due to a perceptual overload (Tickner & Poulton, 1975; Wall, 1965) while others opined that, because witnesses would be able to compare offenders physical traits, they would be able to remember physical details more accurately (Levine & Tapp, 1973). The Clifford and Hollin study ultimately showed an interaction effect between violence and number of offenders. Participants were able to describe one and three offender scenarios slightly more accurately when no violence took place. When violence was added to the scene accuracy dropped lower for one and three offender scenarios respectively and plummeted to below chance when a fifth offender was added.

A later experiment conducted by Geiselman et al. (1993) assessed the ability of 171 college students, who had viewed a videotaped purse snatching, to correctly identify an assailant and his accomplice. This study theorized that there would be unconscious transference causing a high number of identifications of the accomplice rather than the assailant. The study found that there was a 3:1 bias toward identifying persons other than the assailant and significantly more participants were able to identify the accomplice than the perpetrator.

A second experiment for unconscious transference (Geiselman, et al., 1996) used the same purse snatching videotape from their 1993 experiment. After viewing the mock crime videotape and waiting one week, the 238 study participants were shown two target-present simultaneous photo lineups, each containing six photos. One of the arrays contained the assailant and the other the accomplice. When a participant made an identification, they were asked to describe the actions of the person they selected from the lineup. The results showed that although the accuracy of identifications did not differ significantly from their 1993 study, there was a strong bias to label filler photos as the assailant. This experiment confirmed the 1993 experiment's conclusion, that there is little basis for unconscious transference being the reason for confusing the assailant with a bystander.

Although they are not semantically the same, multiple offender crimes in which an incorrect identification is made and single offender crimes where a bystander is misidentified, have certain commonalities. First, more than one person is present during a crime and second, it might be as common to misidentify a bystander as it is an accomplice.

An experiment centered on boundary conditions regarding misidentifications of bystanders by eyewitnesses was conducted by Philips et al. (1997). Although the research was directed at confusion of witnesses between bystanders and offenders, it is certainly possible that what they learned might also be applicable to multiple offender eyewitness identifications. Bystanders and accomplices might have similar attention dividing effects on witnesses.

Philips et al. (1997) used 650 college students who viewed a video of a mock theft that occurred at an ATM bank machine. The video had three versions. The first showed a male who, after staring at the camera for 10 seconds moved off screen. A female was then shown being robbed by a second male. The second version was identical instead the roles of the two males were reversed. The third version exposed the faces of the offender and the bystander for the same amount of time as they crossed paths before the robbery. Four photo lineups with six photos each were constructed. The first lineup contained both the offender and the bystander, the second only the offender, the third only the bystander and the fourth all fillers.

In the first lineup and first video, in which it was plausible that the men were the same, 72% of the participants selected the bystander instead of the offender and only 6% correctly identified the offender. When the bystander was not present in the array, more than half of the participants correctly identified the offender. It should be noted that this higher number of identifications might have been caused by the uniqueness of one male over the other due to stimulus sampling issues (Estes, 1950; Niemark & Estes, 1967; Wells & Windschitl, 1999). In subsequent videos, where the men were shown on screen simultaneously, misidentifications dropped considerably. The authors concluded that a memory blending effect might have been the cause of the unconscious transference that occurred during the experiment.

In a later experiment, two culprit crimes and lineup identification accuracy was examined by Wells and Pozzulo (2006). One hundred and fifty participants viewed a video of a theft in which an offender and an accomplice were present. In addition to testing the accuracy of the witnesses' memory with simultaneous and sequential lineups, a two-person sequential lineup was also tested. This lineup was based on the hypothesis that an offender and an accomplice will be a contextual tag between each other at the time of memory encoding. Half of the participants in the study were shown target absent lineups and the other half, target present lineups. The study found that significantly more people were able to identify the accomplice than the offender. There were no differences between sequential and simultaneous lineup types when it came to identifying the offender or the accomplice. In addition, when target absent lineups were shown, witnesses rejected offender and accomplice lineups at almost identical rates. The study reported that there was a higher incidence of the two-person sequential lineup and increased correct rejections for both offender and accomplice roles.

Unlike the finding in Philips et al. (1999), Dempsey and Pozzullo (2008a) found that when either the offender or the accomplice was present in a lineup, witnesses selected the offender at a higher rate. The authors note that the facial recognition literature suggests that an increase in the number of perpetrators results in diminished witness accuracy. They go on to observe that this axiom may not generalize well to eyewitness research because of fundamental differences in the way research is carried out in each domain (facial recognition v. eyewitness). This study compared eyewitness memory for multiple perpetrator crimes using both a simultaneous lineup and a novel method called an elimination lineup. In the elimination lineup, witnesses are asked to successively eliminate lineup photos before ultimately making an identification decision about the last remaining photo.

Dempsey and Pozzulo's (2008a) objective was to observe whether there was a difference in the lineup procedures on witness accuracy during multiple offender scenarios. One hundred thirty-two people were shown a two culprit crime on video and then observed either a target absent or present simultaneous or an elimination lineup. The lineups consisted of six photos. No significant difference was found between lineup types however, more participants correctly identified the offender rather than the accomplice. In the target absent lineups, participants correctly rejected accomplice lineups at a higher rate. Also, rejection rate varied with the lineup type with a higher percentage of correctly rejected offenders and accomplices when using the elimination rather than the simultaneous lineup.

Owens (2009) conducted a combination single/multiple offender experiment with a sample of 352 college students. The experiment was unique in that the participants were asked, in addition to describing the perpetrators, to also describe the victim. The study discussed system and estimator variables (Wells, 1978; Wells, Luus, & Windschitl, 1994) which brought to the fore the concept of defining police training as an important system variable. The proper training of police officers in eyewitness identification methods might, in itself, be considered a system variable that could be controlled and improved upon. The results showed that approximately 50% of participants were successful in identifying the victim. Correct rejection rates did not show a significant difference for the victim between lineup types. In general, the study found that, as the number of perpetrators rose so did the error rate in identifying offenders. Finally, this study showed that participants were more accurate at rejecting a lineup when the target was not present than identifying the perpetrator when he was present.

Multiple Offenders – Archival Research

In 1996 , 2,010,170 crimes of violence committed by multiple perpetrators were reported to the National Crime Victimization Survey (NCVS) (Bureau of Justice Statistics, 1998). By 2006 that number had decreased by 47% (Bureau of Justice Statistics, 2008). This reduction is slightly greater, but generally consistent with the overall national downward trend in crime since 1994 reported by the Uniform Crime Reports (Federal Bureau of Investigation Criminal Justice Information Services Division, 2008) during the same period. It is important to note that when each of the categories of violent crime including robbery, rape and assault, were examined in the NCVS data, all but rape showed a drop.

In fact, rape stands out as an important category of multiple offender violent crime that has increased in contrast to the general trend of diminishing crime. In 1996 there were 26,630 multiple perpetrator rapes reported by victims. In 2006, there were 38,520 reported, a 31% increase. These figures are especially relevant when studying misidentifications made by eyewitnesses because out of 200 Innocence Project exonerees, 141 (71%) were convicted of rape (Garrett). Further, a search conducted by Owens (2009; Wells, 2001) revealed that out of 164 cases of wrongful convictions based on eyewitness identifications, 15% were committed by multiple offenders. Both of these figures illustrate the potential that research has to shape policy on how public agencies conduct investigations that might have similar circumstances.

In 1976, there were 18,780 total homicides nationwide and in 2005 there were 16,692. The multiple offender homicides for those years were 2,159 and 3,388 respectively, a rise of 11% in multiple offender homicides although the total number was down 69% (Bureau of Justice Statistics, 2007). The fact that, even in light of the general downward crime trend in the United States, these two areas of violent crime are on the rise makes multiple offender research an important topic.

Police Memory v. Citizen Memory - Theory

The majority of experiments that have been conducted on police memory have not attempted to compare police and citizens. We were able to find only seven instances in the literature that attempted a direct comparison (Ainsworth, 1981; Christianson, Karlsson, & Persson, 1998; Lindholm, Christianson, & Karlsson, 1997; Pickel, 2007;

Stanny & Johnson, 2000; Yarmey, 1986, 1998) and none have measured memory on lineup performance. Since no theoretical research has been conducted in this area, any inference of an underlying theory for a difference between police and citizen memory is purely speculative.

One possible reason for the existence of a difference between police and citizen eyewitness memories, if one is found to exist, is Easterbrook's cue utilization theory

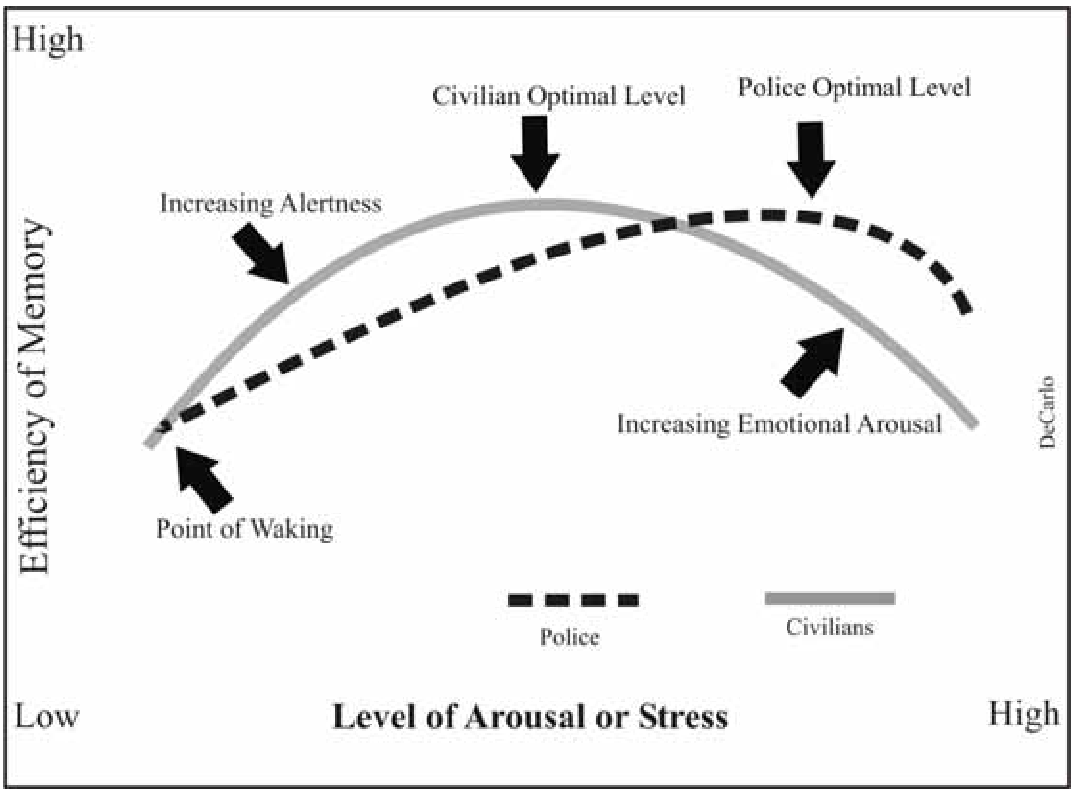


Figure 3. Hypothesized difference in cue utilization theory between police officers and citizens

(Easterbrook, 1959). The theoretical foundation of the current study is that cue utilization might be different for police and citizens because of the constant exposure that police receive to viewing stressful situations on their jobs. As discussed earlier, cue utilization theory proposes that elevated levels of arousal facilitate the processing of and response to stimuli (McCloskey & Egeth, 1983).

However, when over-arousal occurs, situationally relevant "cues" are filtered out to the ultimate detriment of cognitive performance. Hypothetically, if because of repeated exposure to crime scenarios and violence, police officers become desensitized to certain factors commonly attendant to criminal acts, two things might happen. First, if the officer never reaches a level of arousal that causes the focusing effects hypothesized by Easterbrook, then it is possible that citizens, because they become aroused more by crime, remember more. If, on the other hand, the arousal continues to the levels that cause cue filtering of the citizens' cognitive performance, the officers' desensitization could possibly lead to better memory performance. In effect the officers' arousal curve would be flatter (see Figure 3) which might cause them to reach the point of higher arousal and lower cognitive performance, predicted by cue utilization, later than a citizen counterpart.

Police Memory v. Citizen Memory - Empirical Research

One of the first investigations of a possible difference between police officer and citizen memory for certain elements of a crime was conducted by Ainsworth (1981). The study observed that, in Britain and the U.S., the media, public, and court systems all believed that police had greater powers of observation due to their training. The study focused on incident perception of British police officers. Groups of officers who had a number of years of experience on the job were compared to younger officers and citizens. Each group was asked to view several videos which contained staged criminal acts, suspicious circumstances and people, and traffic infractions. The participants (N=48) were instructed to note every incident that they saw. All three groups observed similar total numbers of incidents. The most experienced (mean service time 9 years) police officers however, reported noticing fewer traffic infractions than the newer officers (mean service time 11 months). The number of traffic incidents reported was inversely proportional to the number of criminal incidents noted. Citizens ranked almost as high as newer officers at noticing traffic infractions and matched experienced officers in the numbers of crimes observed. The authors observed: "…[T]he fact that no significant difference was found between the citizens and the two groups of police officers suggests that neither initial training nor subsequent experience significantly affects the police officers' ability to identify offenses." (Page 236)

A more recent experiment by Hulse and Memon (2006) took an approach to investigating police officers' memories of crime scenes which used emotional arousal and weapon focus together as elements that might inhibit memory. In this experiment, police officers were asked to view a third party point of view video (the officers watched other people act out the scenario) where a woman was confronted by two men. The eventual outcomes were that the scene either ended with little violence or the woman produced a weapon and shot both men. One finding in this experiment was that emotional arousal, as measured by a self report instrument called Spielberger's State Anxiety Inventory (1983 See appendix A.), inhibited officers' abilities to make correct line-up identifications as opposed to scenarios where lower arousal was experienced. The study found no evidence of weapon effect that impaired recall of the test scenario. Participants who witnessed the shooting video actually measured greater overall recall than did participants who did not witness a shooting. Findings from this study did not support attentional narrowing that had been observed by other researchers who had tested citizens (Hope & Wright, 2007; Kassin, et al., 2001; Loftus, et al., 1987; Steblay, 1992). The authors did not discuss whether there might be differences between police officers and citizens that might provide an explanation for the differences found between citizen and police study participants. A possible factor exists in this study that might have limited its generalizability. A pre-packaged firearms training scenario, commonly used for shoot, no-shoot police force continuum training on a simulator appears to have been used to present the scenarios to the participants. Because the authors' were unable to obtain clear photos of the actors, all lineups in the study were target absent.

Stanny and Johnson (2000) used several novel approaches in testing the premise that police make better eyewitnesses than citizens. Rather than show videos scenarios of crimes on a computer monitor, as other researchers had done, they used a police training device called a firearms training simulator (FATS), which is commonly used in police firearms decision making training.

Stanny and Johnson (2000) conducted two experiments. The first tested police officers' memories of people who they viewed in a FATS scenario of two different simulated calls for service. One situation was a domestic disturbance and the other an attempted abduction. In each, officers were presented with either a shoot or a no shoot situation. The officers were asked to fill in a survey form describing the actors in the scenario. The results showed that that their accuracy was discernibly worse when a shoot decision was called for in the scenario.

The second experiment compared police officers to citizens and used galvanic skin response (GSR) to measure levels of physiological stress. Other researchers had previously only measured perceived stress as a self report. One member from each of the groups was paired into a police-shooter and citizen-bystander team. When a shoot situation occurred in the FATS video, both members of each pair remembered less detail about the offender and both had better recollection of the weapon that the offender used than their recall of the offender. Interestingly, both officers and citizens showed similar levels of stress, as measured by galvanic skin response, during the scenarios. This finding, in which police and citizens had very similar results in identifying culprits, is contrary to several other studies (Christianson, 1997; Christianson, et al., 1998; Correll, Park, Judd, & Wittenbrink, 2002; Correll, et al., 2007) which found police officers to have greater success acting as eyewitnesses than citizens. Stanny and Johnson (2000) were the first to measure physiological stress levels during an experiment of this type.

Another study (Christianson, Karlsson, & Persson, (1998) hypothesized that police officers might be better at observing details because of their repeated need to do so combined with an interest in improving their efficiency at those types of tasks. They tested 61 college students, 31 teachers, 60 police recruits and 59 more experienced officers at observing and remembering details of a simulated crime shown to the participants as a slide presentation. Their findings showed that the group of more experienced police officers was more precise in remembering details of the crime simulation than the police recruits and citizen groups. The authors concluded that experienced police officers, because of their professional knowledge and experience of violent crime situations, were more able to isolate and analyze information from observing crimes than members of the general public. Even though the study opined that experienced police are probably superior in high stress situations, a limitation of the study was no stress was introduced during the experiment other than the inferred stress of watching slides of a crime occur.

Police Memory v. Citizen Memory – Archival Research

There has been no archival research in this area of inquiry; all of the research that has been conducted has been of an empirical nature. The essential nature of comparing citizens to police officers on memory performance usually requires a side by side measurement. It would be difficult or impossible to research records of incorrect police identifications. Perhaps as the number of DNA exonerations grows and a number of identifications made purely by police is represented among the exonerees, it will be possible to compare citizens and police groups on incorrect identifications. So far, all DNA exonerations on eyewitness evidence have been of persons misidentified by citizens.

Rationale and Hypotheses of the Current Research

Do people who regularly work with, and who are familiar with, certain types of activities and situations possess greater abilities to process and recall details of those activities and situations? Is it possible that police officers can remember the events and people that comprise a crime-scene with a higher percentage of accuracy than non-police witnesses (Stanny & Johnson, 2000)?

When an eyewitness identifies a suspect in a criminal case, courts, police and conventional wisdom attribute a great deal of credibility to the identification. However, an examination of cases in the United States has discovered that mistaken identification by eyewitnesses is the single most common error leading to the arrest and conviction of innocent citizens (National Institute of Justice, 1996). Conversely, because the most efficient and empirically supported methods to effect eyewitness identifications are not always employed, potentially valuable information is at times not learned by investigators. In either instance, the person who actually committed the crime remains unidentified because the wrong person was selected in the identification or, because less than the best possible procedures were followed, no one is identified.

Although DNA evidence has been responsible recently for the exonerations of a small number of wrongfully convicted individuals, it is no panacea to the problem. It is estimated that every year, approximately 77,000 people are convicted of crimes based solely on eyewitness evidence (Wells et al, 1998). Since there are normally only a small number of cases where perpetrators leave DNA evidence at crime scenes, eyewitness testimony is often the only type of evidence that exists. If only ten percent of the eyewitness identifications are incorrect (which research suggests is wildly conservative) (Brigham & Cairns, 1988; Buckhout & et al., 1974; Cutler, Penrod, & Martens, 1987; Lindsay & Wells, 1980; Maass & Köhnken, 1989; Wells & Turtle, 1988)we are potentially left with 7,700 wrongfully convicted people every year (Winzeler, 2008). Eyewitness evidence is very common in court cases and the production of warrants and is often the only evidence available. Eyewitness evidence is given great weight in the criminal justice system. When a police officer takes the stand as a witness, he is accorded even more credibility due to his profession than citizen eyewitnesses (Yuille, 1993). Do both police and citizens have the same potential for error? It is vitally important to learn as much as possible about it through research.

This study tests whether certain situational distractions will affect the recall of crime scenes by police officers and community members differently. The main hypothesis of the research is that when situations which may induce distraction or stress are introduced into a scene, eyewitness memory might be affected by situational elements that might vie for cognitive attention.

Although the concept of WFE has been around as an anecdotal observation for some time, it was first empirically explored by Loftus and Messo (1987) . Broadly, weapon focus effect is a phenomenon in which elements of a crime scene are remembered on an unequal basis dependent on whether or not a weapon is introduced by the perpetrator(s) as the crime is occurring. A knife, gun or other object that a victim perceives is apparently observed with more intensity than other features of a crime scene. The attention that is paid to the weapon usually interferes with clear memories of other elements of the scene (Loftus, et al., 1987).

The perpetrator's facial features, height, race, hair-color, clothing and other salient factors may not be remembered as well when a weapon is present as when one is not (Steblay, 1992). WFE is widely believed by psychologists to be a negative factor in the reliable recall of crime scenes and perpetrator features. Indeed, 87% of psychologists surveyed by Kassin, Tubb, Hosch and Memon (2001) thought that this phenomenon was reliable enough to be reported in court. An older survey by the same authors conducted in 1989, reported only 56% of respondents believed this to be the case. It appears that the continuing research in this area along with consistent empirical results supporting weapon focus have markedly increased the perception that this effect exists as an interfering factor in eyewitness recall.

Theoretical foundation of the current study.

The theoretical foundation of the current study is that cue utilization might be different for police and citizens because of the constant exposure that police receive to viewing stressful situations on their jobs. As discussed earlier, cue utilization theory proposes that elevated levels of arousal facilitate the processing of and response to stimuli (McCloskey & Egeth, 1983).

However, when over-arousal occurs, situationally relevant "cues" are filtered out to the ultimate detriment of cognitive performance. Hypothetically, if because of repeated exposure to crime scenarios and violence, police officers become desensitized to certain factors commonly attendant to criminal acts, two things might happen. First, if the officer never reaches a level of arousal that causes the focusing effects hypothesized by Easterbrook, then it is possible that citizens, because they become aroused more by crime, remember more. If, on the other hand, the arousal continues to the levels that cause cue filtering of the citizens' cognitive performance, the officers' desensitization could possibly lead to better memory performance. In effect the officers' arousal curve would be flatter (see Figure 3) which might cause them to reach the point of higher arousal and lower cognitive performance, predicted by cue utilization, later than a citizen counterpart.

Research Questions

The current study asks the following research questions: 1. Is the eyewitness accuracy (the ability to make correct identifications of culprits in photo lineups) of police and citizens affected differently by the WFE? 2. Is the eyewitness accuracy of police and citizens affected differently by the presence of weapons or multiple perpetrators?

In addition to the traditional manipulations of weapon present and weapon absent, a third weapon condition will be introduced to the literature with this study: a weapon inferred condition. No experiment has addressed this condition before and, in an archival review by the author of police reports of 312 robberies, over half of them had an inferred weapon with no actual weapon being shown. It is unknown if an inferred weapon might add enough distraction, focus or novelty to a situation that it might cause a perceptual narrowing as the victim searches their visual field for the weapon. Another possibility is that the inference increases stress to the point that it interferes with memory encoding. Both police and citizens will be tested to examine if they are affected differently by this manipulation. It is hoped that, if there is a difference between the weapon present condition and the inferred condition, that it might be a step in the direction of disambiguating the causal factors of the WFE by separating the novelty and threat theories in this phenomena by studying the differences between police officers, who handle weapons daily, and citizens to whom weapons are both novel and threatening.

The hypotheses that will be evaluated by the experiments are: 1. that there will be a difference between the number of correct photo identifications that police officers and citizens make; 2) that police officers and citizens will be differently affected by weapon focus effect and multiple perpetrators; 3) that police officers' and citizens' perceived stress levels and physiological stress levels will differ. The perceived stress that a police officer might feel when witnessing a violent scenario might be less than a citizen. This might occur because police officers might become desensitized from repeated exposure to violence. It is theorized that even though perceived stress might differ, physiological stress, e.g., pulse rate and galvanic skin response, might be similar in officers and citizens. In other words, officers might perceive less psychological stress because of desensitization but their bodies would still produce an autonomic response to the violence (Gilmartin, 1990); 4) that perceived stress will coincide with correct lineup identification differently than physiological stress; and 5) that multiple perpetrators will degrade witness memory equally for both groups of participants.

Overview of Methodology

The current study examines galvanic skin response (GSR; see Stanny & Johnson 2000) to measure levels of stress during the video viewing phase of the experiments. In addition, pulse rate, as a secondary covariate, will also be monitored to validate GSR stress levels and explore whether a difference exists between GSR and heart rate measured stress. Unlike Hulse and Memon (2006) whose study measured only perceived stress and only in police officers, the current study will collect both self reports of stress levels and physiological stress measures. This decision was made based on the possibility that perceived and physiological stress levels may be quite different between citizens and police officers when confronted with crime situations. Lindholm, Christianson and Karlsson (1997) reported that police officers, after gaining experience, might become inured or de-sensitized to violence, and therefore have lower physiological and/or perceived stress than citizens who observe the same scenario. The survey instrument used to measure perceived stress measures both state stress for the incident and general stress as a baseline for each participant.

Hulse and Memon (2006) and Stanny and Johnson (2000) used Firearms Training Simulators (FATS) in their experiments. Since there is no shoot/no shoot decision making element in the current study, the branching component of these machines serves no useful purpose in the current study. We instead borrowed the first person point of view of the simulators along with their life size projections and added clearer picture quality and sound for a more realistic, immersive experience. Hulse and Memon (2006) experienced a situation in which they were unable to acquire lineup photos of offenders used in the FATS scenarios because of quality issues when the offenders in these scenarios were photographed from the screen. This caused them to have to use target absent lineups exclusively. The current study avoids that problem by authoring all video scenarios and photo mug shots used in the lineups and using the same actors. High quality photos, produced in the same police environment and on the same police equipment, that the foils were produced on were used throughout the current study. Only caucasians were used in the current study in order to remove the possible confounding of the cross race effect (CRE). A potential problem with eyewitness memory is found in CRE or own-race bias. Cross race effect occurs because people are better at discriminating between peoples' faces of their own racial/ethnic group. Research has shown that the majority of mistakes that occur because of own-race bias are false alarms in which innocent people are misidentified (Anderson, 1999; Meissner & Brigham, 2001). It is interesting to note that in research conducted by Cross, Cross and Daily (1971), it was observed that white people make more misidentifications (45%) when viewing African-American faces than they do with other whites (about 39%). African-Americans, on the other hand, make about the same amount of errors when viewing black or white faces (40%). One of the theories that attempts to explain cross race effect is called the differential experience hypothesis. This theory states simply that not the frequency of exposure to other races but meaningful exposure is the determining factor in becoming familiar with characteristics from other groups. The current study uses only whites in the scenarios so hopefully, will not add a confound to the experiment based on cross race effect.

Statistical Treatment

A statistical power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2008; Faul, Erdfelder, Lang, & Buchner, 2007) was conducted based on the number of conditions in each experiment with the desired outcomes set to at least 80% power (α=.05, two tailed) to detect two outcomes: an accurate or inaccurate lineup identification. G*Power indicated required sample sizes of between 118 and 290 for the statistical tests mentioned above. The sample sizes in both Experiments 1 and 2 are consistent with the statistical treatments used, specifically, t-tests, one way and two ANOVA, chi-square tests for independence and logistic regression and, in fact, exceed the predicted power requirements.

Several statistical treatments were used to construct the analysis plan using data from the two experiments in the current study. The first examination was a broad look at the data using descriptive statistics to learn the overall identification rates in the study and whether they were consistent with prior research. Next, chi-square tests for independence were used to analyze whether a significant association exists between police and citizen groups and making correct identification on lineups.

As discussed earlier, the main research question asked in these two experiments was whether there was a difference in the correct lineup identification rates between citizens and police officers. The main dependent variable in both experiments is dichotomous, an identification or rejection is correct or it is not. The major statistical analysis was conducted using logistic regression (Tabachnick & Fidell, 2007). It examined what variables, including weapon presence, the number of perpetrators, perceived stress and physiological stress, have the largest predictive value on a witness's ability to make a correct or incorrect categorical decision about a lineup after controlling for all the other variables in the model. Unlike multinomial logistic regression where the dependent variable is a nominal variable with more than two categories which cannot be ordered meaningfully, binary logistic regression was chosen for the current study because the dependent variable is restricted to two categories – correct or incorrect (Hosmer & Lemeshow, 2000; Menard, 2001; SPSS Inc., 2010; Tabachnick & Fidell, 2007).

Unlike Ordinary Least Squares (OLS) regressions, logistic regression has the ability to use dependent variables with limited numbers of outcomes as are found in the current study (Menard, 2001). Logistic regression does not have the same assumptions of homoscedasticity, linearity and normality (Pampel, 2007) .

In addition to the regular log-linear functions of this type of regression, the logit model also allows for analyzing the odds of the expected cell frequencies of the dependent variable (Aldrich & Nelson, 2007; Hosmer & Lemeshow, 2000; Knoke & Burke, 2007). Although one prior study, examining non-identifications in actual criminal cases, used a simple binomial logistic regression (Hagan, 2009), no study in the present eyewitness literature has combined the current variables and type of statistical analysis in an attempt to examine the relationships that might add to the understanding of certain interactive factors affecting eyewitness efficacy which are important to the criminal justice field. When they are combined to test for interactions among them, our predictor variables might produce non-linear results. This makes the ability of logistic regression to work in this situation especially important. The model produced by logistic regression is non-linear and the outcome variable, $\hat{Y}$, shows what the probability of having either of the outcomes of the dependent variable might be in a non-linear fashion. The logistic prediction equation is: log (p/1-p) = b0+b1*x1+ b2*x2+ b3*x3+ b4*x4.

Because our primary dependent variable only has one category, it was not necessary to create dummy variables to utilize a reference variable (Hardy, 2007; Hosmer

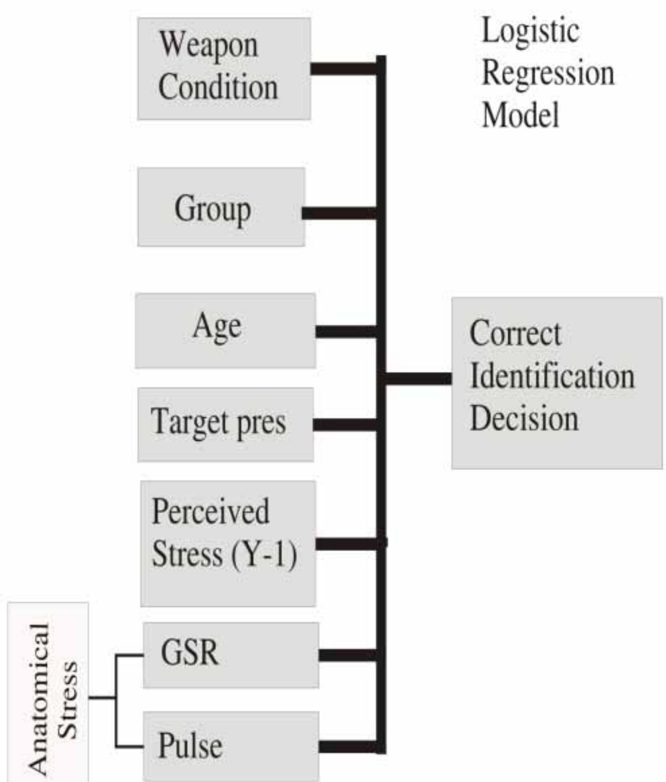


Figure 4. Logistic Regression Model for Experiment 1

& Lemeshow, 2000) The interpretation of the logistic regression results should indicate if our model(s) can significantly predict the chances of success of making correct eyewitness identification, the relative contribution of each of the independent variables and how strongly the model(s) predict the probability of our dependent variable (see Figures 4 and 5).

Experiment 1 – Weapon Focus Effect

Method

Participants

The sample size for Experiment 1 was 323 individuals (165 police and 158 citizens). The mean age of police officers was 37.7 years (SD =12.88) and the mean age for citizens was 40.7 years (SD =12.88). All participants were from the same Northeastern state in the United States. Police officers came from 6 different agencies and the size of the departments ranged from 49 to 611 sworn officers (M = 173.3, SD = 224.6). The mean length of service for police officers was 10.6 years (SD = 5.95). Police participants took part in the experiment as they participated in their regular in-service training. Citizen participants were volunteers recruited from the communities in which the police departments are located and one university from the Northeastern United States.

Although the citizen group included participants of both genders (72 females and 86 males), the police group only included males for two reasons. First, the departments that agreed to take part in this study have less than a total of 3% female officers. Second, the percent of female police officers in the United States is, according to the latest census by the Department of Justice (Federal Bureau of Investigation Criminal Justice Information Services Division, 2008), approximately 11%. Because of these low numbers, it was not possible to oversample female police officers so they were not included in the sample. In addition, as mentioned earlier, it was decided that in order to avoid the possible confounding by cross race effects in identification accuracy (see Meissner & Brigham, 2001) we would test only Caucasian participants[1]. None of the participants had any prior knowledge of the purpose of the experiment.

Design

Experiment 1 was a 3 (Weapon condition: present, absent, inferred) x 2 (target present/target absent lineup) x 2 (police and citizen) between subjects factorial design. Age and sex (of citizens) was also recorded and examined (Campbell & Stanley, 1963, p. 7; Cook & Campbell, 1979; Trochim, 2001). Other demographic data that was recorded for police officers was the length of service time, current rank, size of agency and when, during their tour of duty, the experiment took place. An eight hour tour of duty was broken down into two half segments numbered 0 and 1, with 0 being the first four hours of the tour and 1 being the final four hours.

Participants' ability to accurately make a correct identification decision in a photo lineup was the primary (categorical) dependent variable. In addition, pulse rate and galvanic skin response (GSR) were tracked as consequent dependent variables in response to the possible stress induced by the robbery scenario.

---

[1] Cross race effect occurs because people are better at discriminating between peoples' faces of their own racial/ethnic group. Research has shown that the majority of mistakes that occur because of own-race bias are false alarms in which innocent people are misidentified (Anderson, 1999; Meissner & Brigham, 2001). It is interesting to note that in research conducted by Cross, Cross and Daily (1971), it was observed that white people make more misidentifications (45%) when viewing African-American faces than they do with other whites (about 39%). African-Americans, on the other hand, make about the same amount of errors when viewing black or white faces (40%). Due to the fact that weapon focus effect is what is being tested for, having mixed ethnicities take part in the experiment might have added potential problems in isolating the independent variable.

Materials

Three video-taped scenarios were filmed involving a robbery being committed in a busy environment. The videos were filmed from the point of view of the victim. In addition, to decrease the likelihood that the results might be due to a specific perpetrator being particularly memorable (a key assumption of Stimulus Sampling Theory; (Estes, 1950; Niemark & Estes, 1967; Wells & Windschitl, 1999), two different perpetrators were filmed in all three versions of the scenario, which were: A) *No weapon condition*, where a male perpetrator was shown holding up the victim without displaying or inferring that he was in possession of a gun; B) *Inferred weapon condition*, where a male perpetrator held up the victim with an inferred weapon. This was achieved by having the perpetrator hold his hand under his shirt and referencing it as a gun, but no gun was actually displayed; and C) *Weapon condition*, where a male perpetrator held up the victim with a gun displayed the entire time the perpetrator was in view. In each 45 second scenario, the perpetrators' faces were visible for approximately 30 seconds. All perpetrators were Caucasian males with short dark hair and no discerning features, such as scars, marks or tattoos. The height of all actors who played perpetrators was approximately 69-70 inches and the general body configurations were similar. All wore dark colored shirts and trousers.

All videos used in the experiment were recorded in color, high definition digital video on a Sony DCR-PC110 digital video camera. Sound was recorded through a supplemental Sony high gain directional microphone. Once the videos were captured to tape, they were transferred to a hard disk storage device and were cued from computer. The video output of the computer was routed through a high intensity LCD projector which enlarged and projected the video onto a semi-circular screen. This was intended to produce a more immersive, ecologically valid experience than viewing on a computer monitor might be able to deliver. The videos that were viewed by participants were life size and from the victim's point of view. This point of view is more consistent with a person actually experiencing a robbery than a third person perspective which is typically used in eyewitness experiments. Participants viewed the crime simulation videos projected by a Sanyo Mark 120 LCD Video projector connected to a Dell Laptop computer via its VGA port. Resolution was set at 1280x1024 to produce maximum resolution. The screen used to project the videos at approximately life-size was a semi-circular design, for use in a firearms training simulator. Speakers used to deliver sound were Altec-Lansing 8 ohm high-fidelity midrange speakers with added tweeters for high frequencies and a sub-woofer for low level bass. After the simulated crime scene videos were viewed, a computerized sequential lineup was presented on a laptop computer (MacLin, Zimmerman, & Malpass, 2005).

All lineups, containing eight digital photos each, were displayed in a double-blind, sequential method on a laptop computer controlled by the participant. Lineup fillers were chosen based on the general appearance of the simulated perpetrators. Wells' (1993) recommendations on propitious heterogeneity were observed when selecting fillers and no foils with gratuitous similarities were included. The clothing worn by the lineup members was whatever they were wearing when they came in to be photographed. We purposely did not photograph the lineup actors wearing the same clothes as they did in the video since this might have been suggestive.

During Experiment 1, participants had their galvanic skin response (GSR) recorded, as a measure of stress (Stanny & Johnson, 2000), on a GSR Temp 2X Biofeedback device that was connected to the participant's left hand index and middle fingers. GSR is also known as electro-dermal response (EDR) or psycho-galvanic reflex and is used to measure stress levels in lie-detection equipment and in psychotherapy to measure depth of hypnotic trance (Baker & Taylor, 1955; Cutler, et al., 1994). It has been validated as a measuring device in research for over fifty years (Office of Technology Assessment, 1983) and is commonly used to measure levels of stress in psychological experiments (Baker & Taylor, 1954, 1955; Boggs, 1904; Bruning & Frew, 1987; Coppock, 1955; Du Toit, 1956; Geer & Klein, 1969; Horvath, 1978, 1979; Kimmel & Hill, 1961; Pishkin & Hershiser, 1963). The specific model of GSR measurement device used for this experiment was identical in output to other models commonly used to measure GSR and was validated against a larger biomedical research model used for medical research at Yale University. The model used for this research was able to measure skin resistance between a range of 1,000 ohms and 3,000,000 ohms. The scale on the meter was read directly during the manipulation and was gradated from -2 to +2 in ten point increments. The device is pictured in Appendix C.

Another measure of stress, pulse rate, was measured using a finger pulse oximeter connected to the ring finger of the participant's left hand. This device is designed to measure pulse rate and blood oxygen saturation which both rise when a person feels stress (Awad, 2001). Finger pulse oximeters are commonly used in medical offices and hospitals to collect pulse and oxygen saturation data. They have been also been used in biomedical research for the same purpose (Screbo, 1999). Only pulse was recorded for this experiment, no oxygen saturation levels data was collected.

In order to measure participants' level of perceived stress, the 20-item Spielberger's State Anxiety Inventory, Form Y-1 was administered to each participant after viewing the video and before being shown a lineup. The possible range of scores for the test is between 20 and 80 with higher scores indicating higher state anxiety. A sample of the State Anxiety Inventory form Y-1 is shown in appendix A. Spielberger's state anxiety inventory is a generic instrument used to measure the state of stress a person is feeling. No specific instruments, specifically to measure stress in police officers, are known to exist. It was considered important when selecting this instrument that the same test be used to measure perceived stress in both police and citizens. In addition, the validity and reliability of the Y-1 test has been shown by research to be extremely high (Quek , Low, Razack, Loh, & Chua, 2004; Smeets, Merckelbach, & Griez, 1996). Many examples of the inventory's construct validity have shown consistently valid results (Spielberger, 1983; Tilton, 2008). Spielberger (1995) reports that concurrent validity between the Y-1 and other scales that measure anxiety, such as the Anxiety Scale Questionnaire and the Manifest Anxiety Scale, have produced correlation scores of .72 and .85 respectively. These results reinforced the decision to select the Y-1 as a highly valid perceived stress metric for the current research.

Procedure

Participants, randomly assigned to all conditions, were shown one of three video scenarios projected onto a large screen. Randomization was accomplished by the computer randomly selecting the video to be shown based on the generation of a random seed algorithm programmed by the researchers in C++. After each participant viewed their randomly assigned video, they were immediately asked to complete Spielberger's (1983) State Anxiety Inventory (Form Y-1). Y-1 Pretesting was not conducted during our experiment due to time restrictions on participants' availability. At this time, no questions were asked by the experimenter.

Once the participant had completed the Y-1 form, they were shown either a randomly selected (by the same procedure above) target-present or target-absent sequential lineup on a laptop computer. Prior to viewing the lineup, the following instructions were given to participants: "The individual that you observed in the video may or may not be present in the photos that you are about to see. Even if you think that you recognize an individual, please continue looking at the other photos until there are none remaining." If the subject identified any of the lineup members, their identification decision was recorded. Participants were debriefed, thanked for their time and dismissed.

Results

Pulse, GSR and Perceived Stress

Overall. Measurements were taken of three factors that have been associated with stress and which increase during arousal: Galvanic Skin Response (GSR), reported state anxiety (perceived stress), and participants' maximum pulse during the experiment. The mean pulse at the start of the experiment, before viewing the robbery videos, was 72.1 (range = 65 to 80; SD = 4.33). During the viewing of the robbery video, which took approximately 45 seconds, the pulse was continuously monitored and the highest pulse attained was recorded. The mean high pulse was 73.9 (range = 65 to 85; SD = 4.73). The mean GSR attained during the test was .83 (range = 0.0 to 2.00; SD = .583). Spielberger's

State Anxiety Inventory mean score was 44.8 and the range was 54 with a minimum of 22 and a high of 76 (SD = 12.76). This is compared to a sample of 1,883 working adults tested by Spielberger (1983), who reported normative scores on the state anxiety test of M=35.72 (SD = 10.40) . The increase in arousal between the normative and the test groups seems consistent with the test group having just witnessed a simulated robbery. An independent samples t-test was conducted to compare the overall Y-1 scores for police and citizens. There was a significant difference in scores for police (M = 43.33, SD = 10.91) and citizens (M = 46.57, SD = 14.39); t (277) = 2.23, p =.027. The magnitude of the differences in the means (mean difference = 3.2, 95% CI: -6.0 to -.37) was small (eta squared = .01). Neither the GSR (t (321) = .943, p =.347) or the pulse data (t (321) = .960, p =.338) was significantly different between groups.

State anxiety Y-1. A two-way between groups analysis of variance was conducted to explore the potential interaction between being in the police or citizen group and weapon condition on state anxiety scores as measured using form Y-1 of Spielberger's State Anxiety Inventory. In the initial analysis, Levene's test for equality of error variances was significant at the < .05 level suggesting that the variance of the dependent variable was not equal across groups (homogeneity of variance was violated). To counter this condition, we re-ran the analysis at a more stringent <.01 level. The interaction effect between group and weapon condition on Y1 score was not statistically significant at this level, F (2,309) = 3.435, p = .033. Also, at the < .01 level, none of the variables reached a significant main effect (see Table 2).

A one-way between groups analysis of variance was conducted to explore the impact of weapon condition on state anxiety scores as measured using form Y-1 of

Spielberger's State Anxiety Inventory. Subjects were divided into three groups according to the weapon condition they experienced, no weapon, weapon inferred and weapon present. There was not a statistically significant difference at the $p < .05$ level in Y-1 scores, F (2, 312) = .148, p = .863.

Table 2. *Y-1 state anxiety mean scores from each weapon manipulation.*

	N	*M*	*SD*
Overall			
Police	165	43.43	10.91
Citizens	158	46.57	14.39
No weapon			
Police	48	44.75	10.17
Citizens	46	44.30	15.16
Inferred weapon			
Police	61	43.88	11.10
Citizens	56	45.51	13.76
Weapon Present			
Police	56	41.53	11.25
Citizens	48	49.97	14.01

A one-way between groups analysis of variance was conducted to explore the impact of being in the police or citizen group had on Y-1 score. Y-1 scores during the experiment showed a significant difference between police (M = 43.33, SD = 10.91) and citizens (M = 46.67, SD = 14.39), f (1, 313) = 5.10, p = .34 ($\eta 2$=.02). It is unknown if the limited time of exposure to the stimulus caused by the videos or the fact that the videos did not induce an intense stimulus, was a factor in the Y-1 scoring.

Pulse. A two-way between groups analysis of variance was conducted to explore the potential interaction between being in the police or citizen group and weapon condition on pulse during the manipulation. The interaction effect between group and weapon condition on Y1 score was not statistically significant, $F\,(2,323) = .639$, $p = .528$. Also, at the < .05 level, none of the variables reached a significant main effect meaning that police and citizens did not significantly differ in their pulse scores (see Table 3).

One-way between groups analysis of variance was conducted to explore the impact of weapon condition on pulse. Pulse during the experiment also showed no significant difference between police (M = 73.6, SD = 4.7) and citizens (M = 74.2, SD = 4.7), t (320) = .960, p = .34 (η2=.099). It is unknown if the limited time of exposure to the stimulus caused by the videos or the fact that the videos did not induce an intense stimulus, was a factor in the non-significance of pulse rate.

Table 3. *Summary of independent sample t-tests showing difference between police and citizen groups on a computed variable which was based on the* ***difference*** *in starting pulse and high pulse attained during the manipulation.*

	N	*M*	*SD*	*F*	*df*	*p*	η^2
Overall							
Police	165	+1.88	2.36	.783	321	.424	.002
Citizens	158	+1.93	1.93				
No weapon							
Police	48	+1.79	1.45	.352	98	.725	.001
Citizens	52	+1.63	2.75				
Inferred weapon							
Police	61	+1.68	3.28	.301	76.26	.764	.001
Citizens	56	+1.55	1.17				
Weapon Present							
Police	56	+2.17	1.71	.807	104	.421	.006
Citizens	50	+1.92	1.56				

Galvanic Skin Response (GSR). A two-way between groups analysis of variance was conducted to explore the potential interaction between being in the police or citizen group and weapon condition on GSR measurement during the manipulation. The interaction effect between group and weapon condition on GSR measurement was not statistically significant, *F* (2,323) = .994, *p* = .371. Overall, there was no main effect of GSR scores between police (M = .86, SD = .578) and citizens (M = .79, SD = .58), t (319) = .943, p = .34 (η2=.099 with the weapon condition main effect result being *F* (2,322) = .43, *p* = .090.

Table 4. *GSR mean readings during weapon manipulations*

	N	*M*	*SD*
All conditions			
Police	165	.86	.57
Citizens	158	.79	.58
No weapon			
Police	48	.88	.61
Citizens	46	.70	.58
Inferred weapon			
Police	61	.80	.50
Citizens	56	.74	.56
Weapon Present			
Police	56	.90	.62
Citizens	50	.95	.59

Identification Accuracy

Examining identification accuracy as a whole, participants made correct identifications across all conditions 65.6% of the time. Looking at the identification accuracy of police versus citizens, it was found that no statistically significant difference existed in overall accuracy between the groups. Police made correct identifications 70.3% of the time and citizens 60.8% of the time, χ^2 (1, n=323) = 3.25, p = .071, phi = .071.

Identification accuracy and weapon presence overall. The distribution of correct identification decisions was compared across all three weapon conditions to determine if the presence of an inferred or actual weapon affected the accuracy of correct identification decisions. A chi-square test for independence indicated a statistically significant association between weapon presence and correct decisions with accuracy falling off 35.9% from the no weapon condition to the weapon present condition, χ^2 (2, n=323) = 29.38, p = .019, Cramer's V = .302 The percentage of correct identifications in the no weapon condition was 84%; in the weapon inferred condition, 65.8% and the weapon present condition, 48.1%.

Police v. citizens weapon presence moderated by target condition. When comparing the identification performance (a correct identification or rejection) of police officers versus citizens, no statistically significant difference was found to exist between the groups in the target absent condition. When the target was absent from a lineup, police made a correct decision 79.8% of the time and citizens 74.4%, χ^2 (1, n=161) = 0.385, p = .535, phi = .064. When the target was present in a lineup, a significant difference existed between police and citizens in identification accuracy. In these conditions, police were correct 61.6% of the time compared to citizens at 46.1% of the time, χ^2 (1, n=162) = 3.342, p = .058, phi = .156. As can be seen, the phi correlation coefficient indicates a small effect as defined by Cohen (1988).

Weapon presence moderated by target condition- overall. Next, we examined whether an inferred weapon or actual weapon would have an effect on the accuracy of identification decisions when moderated by having the target absent or present in the lineup. To make a correct identification decision, a participant was required to identify the target when he was present in the lineup or reject the lineup if he was absent.

Target absent. A chi-square test for independence indicated a statistically significant association between weapon presence and identification decisions when the perpetrator was absent from the lineup, χ^2 (2, n=161) = 7.88, p = .019, Cramer's V = .221. In this condition, all participants combined were able to make a correct rejection 88.86% of the time. When a weapon was inferred (held under a garment and referenced) also in the target absent condition, identification accuracy dropped to 76.27% and when a weapon was actually shown, again in the weapon absent condition, correct rejections further dropped to 65.31% for a total decline in accuracy of 23.29%.

Target present. When the perpetrator was *present* in the lineup, a chi-square test for independence indicated a significant association between weapon condition and identification accuracy, $\chi2$ (2, n=162) = 21.41, $p < .001$, Cramer's V = .364. In the no weapon condition, participants were able to correctly identify the target 78.7% of the time. When a weapon was inferred, accuracy dropped to 55.2% and when a weapon was actually shown, accuracy dropped to 33.33%.

Table 5. *Percentage breakdown of identification decisions (frequency in parentheses) by group, lineup condition and weapon condition. Last rows shows diagnosticity ratios*

		Group	
	Overall	Police	Citizen
Target absent			
Correct Rejections			
No weapon	88.7 (47)	92.3 (24)	85.2 (23)
Inferred Weapon	76.3 (45)	79.3 23)	73.3 (22)
Weapon Present	65.3 (32)	66.7 (16)	64.0 (16)
Filler Identifications			
No weapon	11.3 (6)	7.7 (2)	14.8 (4)
Inferred Weapon	23.7 (14)	20.7 (6)	26.7 (8)
Weapon Present	34.7 (17)	33.3 (8)	36 (9)
Target present			
Correct Identifications			
No weapon	78.7 (37)	90.9 (20)	68.0 (17)
Inferred Weapon	55.2 (32)	65.6 (21)	42.3 (11)
Weapon Present	33.3 (19)	37.5 (12)	28.0 (7)
Filler Identifications			
No weapon	10.6 (5)	4.5 (1)	16 (4)
Inferred Weapon	5.2 (3)	3.1 (1)	7.7 (2)
Weapon Present	15.8 (9)	3.1 (1)	32 (8)
DR			
No weapon	6.96	11.81	4.59
Inferred Weapon	2.33	3.17	1.58
Weapon Present	0.96	1.13	0.78

Police v. citizens moderated by weapon condition. There were no significant differences between police and citizens across the three weapon conditions (see Table 5). When no weapon was present there was no significant difference between the groups. Police made correct identification decisions 92.3% of the time and citizens 85.2% of the time, χ^2 (1, n=161) = .146, p = .701, Phi = .112. When a weapon was inferred there was still no difference in the identification decision rates between the groups. Police were correct 79.3% of the time and citizens 73.3% of the time, χ^2 (1, n=161) = .054, p = .815, Phi = .070. When a weapon was shown, the same trend continued. Police were correct 66.7% of the time and citizens 64.0% of the time, χ^2 (1, n=161) = .038, p = .845, Phi = .112.

Police v. civilians - Filler identifications. Police made significantly fewer filler identifications in TP lineups than citizens. When filler identifications by police or citizen group, across all weapon conditions, were examined we found that when a target was absent, police made incorrect filler identifications 20.7% of the time and citizens made incorrect filler identifications 26.7% of the time. When a target was present, police made incorrect filler identifications 3.5% of the time and citizens made incorrect filler identifications 18.4% of the time. The target absent condition showed no significant difference between groups in filler identifications, $\chi2$ (1, n=323) = .652, p = .419, Cramer's V = .964 but when the target was present, the police and citizen group did show a significant difference, $\chi2$ (2, n=323) = .10.38, p = .006, Cramer's V = .064.

Diagnosticity ratio. The diagnosticity ratio was computed for the lineups by dividing the correct identifications in the target present condition by the percentage of incorrect identification decisions in the target absent condition. The police group had higher diagnosticity ratios across all three weapon conditions which might possibly mean that police were more discriminant in their selections. See table 5 for the actual diagnosticity figures. It is important to note that diagnosticity ratio is the percentage of correct responses in TP to incorrect choices in TA lineups. Since, under actual conditions in police departments, lineups are not used over again many times, diagnosticity ratio is only useful as an analytical tool and not in the real world.

Factors Contributing to Identification Decision Inaccuracy

Direct logistic regression was performed to assess the impact of factors on the likelihood that participants would make a correct identification decision (see Table 6). The model contained seven independent variables: (age; GSR, pulse rate; Y-1 score; target presence; weapon condition and group.) The full model containing all predictors was statistically significant, χ^2 (8, N= 315) = 53.160, p<.001, indicating that the model was able to distinguish between participants who made and who did not make a correct identification. A Hosmer and Lemeshow test, which is indicative of the goodness of model fit, was conducted and indicated that the model was worthwhile (χ^2 = 5.851, p = .664). The model, as a whole, explained 21.5% (Nagelkerke R Square) of the variance in identification correctness, and correctly classified 73% of cases. As shown in Table 6, three of the independent variables made unique, statistically significant contributions to the model.

On average, the odds of making an incorrect identification decision (identification or rejection), after holding all other variables constant, when there is a weapon inferred are 3.066 times higher than when no weapon is present. When a weapon was actually shown, after holding all other variables constant, the odds of making an incorrect identification are 6.161 times higher than when a weapon is absent. On average, the odds of making an incorrect identification decision after holding all other variables constant, when a lineup target is present, are 3.276 times higher than when the target is present. On average, after holding all other variables constant, the odds of making an incorrect identification decision were 1.821 times greater in the citizen group than in the police group. In summary, when a weapon is present, the likelihood of making an incorrect identification decision increases. Also, when a target is absent from a lineup the odds of making an incorrect identification decision are less than when the target is present (only counting rejections obviously). Finally, police generally make fewer identification decision errors than citizens.

Table 6. *Logistic regression model showing impact of factors on correct identifications*

Variable	β	S.E.	*Wald*	*df*	*p*	*Odds ratio*	95.0% C.I	
							Lower	Upper
Age	-.002	.012	.018	1	.892	.998	.976	1.022
GSR	.046	.238	.038	1	.846	1.047	.657	1.671
Pulse	.015	.029	.257	1	.612	1.015	.959	1.074
Y1Score	-.006	.010	.323	1	.570	.994	.974	1.014
No Weapon			24.933	2	.000**			
Inferred Weapon	1.120	.358	9.814	1	.002	3.066	1.521	6.179
Weapon present	1.818	.365	24.872	1	.000	6.161	3.015	12.590
Target condition	1.187	.270	19.373	1	.000**	3.276	1.932	5.558
Group -Police or Citizen	.599	.268	4.986	1	.026**	1.821	1.076	3.082

** p<.001

An additional direct logistic regression was performed to assess the possible impact of other factors on the likelihood that participants would make a correct identification decision (see table 6). The model contained five independent variables: (age; agency size, length of service; rank and tour segment.) The full model containing all predictors was not statistically significant, χ^2 (7, N= 323) = 2.32, p<.940, indicating that the model was not able to distinguish between participants who made and who did not make a correct identification. A Hosmer and Lemeshow test, which is indicative of the goodness of model fit, was conducted and indicated that the model was not worthwhile (χ^2 = (8,321) 11.66, p = .167). The model, as a whole, explained 2% (Nagelkerke R Square) of the variance in identification correctness, and correctly classified 70% of cases. As shown in Table 7, none of the independent variables made unique, statistically significant contributions to the model.

Table 7. *Logistic regression model showing no impact of certain variables on identification decisions*

	β	S.E.	*Wald*	*df*	*p*	*Odds ratio*
Age	.005	.031	.030	1	.862	1.005
Agcy Size 50 to 100			1.050	2	.591	
Agcy Size 101 to 200	.086	.452	.036	1	.850	1.089
Agcy Size 201>	.438	.429	1.046	1	.306	1.550
Length of service	-.038	.051	.562	1	.453	.963
Rank Ptlm			.244	2	.885	
Rank Sergt	-.027	.600	.002	1	.965	.974
Rank Lt>	.636	1.322	.231	1	.631	1.888
Tour Seg	.174	.354	.242	1	.623	1.190

CI not computed because model was not significant

Discussion

When the data were examined in the context of our three hypotheses, several interesting findings emerged. Our first hypothesis, that there would be a difference in the number of correct identification decisions in different weapon conditions, was supported by our results. There was a significant reduction in accuracy for both police and citizen groups as weapons were first inferred and then made present in the experimental scenarios. Another hypothesis, that police officers and citizens would be differentially affected by weapons focus effect was rejected. There was not a significant difference between groups either overall or across weapon conditions for both target absent and target present lineups with the exception of police making fewer incorrect filler identifications when a target was absent in the no weapon condition.

Lastly, we hypothesized that police officers' and citizens' perceived stress levels and physiological stress levels would differ. What we found was that measures associated with physiological stress (i.e., pulse and GSR) did not vary significantly between groups. Police had a mean stress score on the inventory that was lower than citizens. This finding is generally consistent with our hypothesized difference in cue utilization between police and citizens, possibly due to some degree of desensitization to stressful scenarios having taken place in police officers.

The fact that we found that police made significantly fewer filler identifications in target present lineups than did citizens was unanticipated. When a target was present, police made incorrect filler identifications 3.5% of the time and citizens made incorrect filler identifications 18.4% of the time. This finding is also supported by the diagnosticity ratios that were computed for each group across weapon conditions.

The propensity to make fewer filler selections may emanate from several different sources. Police may be more familiar with the lineup identification task and thus this experience might have had an influence on their choosing behavior. Alternatively, police might have a greater knowledge of the criminal justice system (Hulse & Memon, 2006) and the penalties associated with a wrong decision than citizens. This may inhibit them from guessing when they do not think they will be accurate. The target absent condition showed no significant difference between groups in filler identifications. This might indicate that when the target is absent both police and citizens guess at the same rate. Looking back at the filler identifications in the target present category when a weapon was present, we recorded police choosing fillers 28.9% fewer times than citizens. At this time, we are uncertain as to why this difference emerged and recommend that future researchers investigate target present lineups in more detail.

Integration of Findings with Past Literature

Utilizing the results of the Spielberger's State Anxiety Inventory, we found that the levels of stress, recorded just after the experimental videos were viewed, were higher than the normative values for the inventory. This heightened level of stress, and the difference recorded between police officers and citizens, is consistent with the lower rates of correct identification decisions when more stressful conditions, such as an inferred or actual weapon, were introduced. Although the difference in correct identification decisions was not significantly different between police and civilians, the difference in anxiety inventory scores between the groups might well be demonstrative of our hypothesized difference in the shape police and citizen cue-utilization curves (see Figure 3).

As we hypothesized, if because of repeated exposure to crime scenarios and violence, police officers become desensitized to certain factors commonly attendant to criminal acts, two things might happen. First, if the officer never reaches a level of arousal that causes the focusing effects hypothesized by Easterbrook (1959), then it is possible that citizens, because they become aroused more by crime, remember more. If, on the other hand, the arousal continues to the levels that cause cue filtering of the citizens' cognitive performance, the officers' desensitization could possibly lead to better memory performance. In effect the officers' arousal curve would be flatter which might cause them to reach the point of higher arousal and lower cognitive performance, predicted by cue utilization, later than a citizen counterpart. If stress were then, the only contributor to accuracy, police would theoretically have been more accurate. We believe that it may be possible that novelty also contributes to WFE and discuss this concept further below,

Weapon focus effect. Of all the tests conducted during the experiment, none were more consistently apparent than those of weapon focus. Hope and Wright (2007) conducted a study which asked if police officers, who handle weapons on a daily basis, might be less affected by the unusualness or novelty of a weapon than someone not inured to weapon presence. The results of the current study show that police are clearly affected by WFE. This might suggest that they are not immune to the novelty explanation of the phenomenon.

Perhaps the novelty aspect of a weapon does not reside in whether a gun is a common item in one's daily environment. Conceivably, the novelty aspect might come from a weapon being pointed at an individual which places it in an entirely new context. Support for this hypothesis is found in the current study in that both citizens and police reacted with similar identification rates across all of the weapon conditions.

Putting the current study into perspective with other WFE studies shows some consistencies and some inconsistencies. A study by Hulse and Memon (2006) using only a police sample and only a target-absent lineup condition had divergent findings for the existence of WFE. Their study did not show a lessened ability, as found in the current study, to make accurate identification decisions in the presence of a weapon. There are three factors that might account for the divergent WFE findings in the Hulse and Memon study. First, they used a domestic violence scenario in which the participants were virtual observers rather than virtual victims. Secondly, the weapons in the scenario were only visible for 5 seconds and were not directed at the experiment's participants. Lastly, Spielberger's Y-1 scores were lower than in the current study indicating that the participants experienced less stress.

Similar to our study, Bartlett, et al. (2003) found that civilian witnesses were more likely to make a filler identification, especially as the seriousness of a crime increased. In addition, Hulse and Memon proposed that police officers would be less prone to making filler identifications because they would be more inherently aware of the "…[C]onsequences of making an incorrect identification, particularly when a crime was serious…" (p. 322). In future studies, we would recommend that researchers ask both police and citizens participants about the consequences of making an incorrect identification decision in real cases in order to test this hypothesis.

Police – citizen memory comparisons. No discernable difference, except the previously noted tendency for police in the current study to choose fillers at a lesser rate, was found. Stanny and Johnson (2000) conducted an experiment comparing police officers to citizens and also used galvanic skin response (GSR) to measure levels of physiological stress. During their study, as in the current research, both officers and citizens showed similar levels of stress, as measured by galvanic skin response and also showed diminished accuracy rates. Stanny and Johnson did not use lineups but survey details, instead, as a measure of accuracy. Recall memory such as this may actually use a different type of encoding than the current study so may not generalize the effect of stress on memory.

Other studies have investigated whether police and citizens differ in terms of their recall of events during the commission of a crime and found that police are sometimes more accurate (Christianson, 1997; Christianson, et al., 1998; Correll, et al., 2002; Correll, et al., 2007). These results cannot be directly compared to the results of the current study because we examined face identification accuracy from lineups. The research literature on eyewitness identification accuracy shows little relationship between recall accuracy and identification accuracy (see Wells et al., 2006). Similar to the current study, however, Christianson (1998) found that civilian and police groups were not significantly different when attempting to make lineup identifications.

The findings in the current study contribute to the literature in several ways. First, it has been assumed that police make better eyewitnesses than citizens (Cutler & Penrod, 1995; Deffenbacher & Loftus, 1982; Yarmey, 1998). This belief has been based largely on the assumption that police are "trained observers" who receive special instruction at the police academy or on the job in eyewitness techniques. This assumption, in reality, is mostly false.

Checking the curriculum of the police academy system in the state where the officers in this study were from, no instructional content was found in observation or memory and the factors that affect eyewitness abilities either negatively or positively. In personal communications with 41 police administrators from 32 states, the answer to the question on whether special training on eyewitness memory was given to police officers was always the same. No jurisdiction included any but the most rudimentary type. Usually, cadets at the police academy were asked, after the fact, to describe a person who had walked across the classroom while the cadets were receiving a lecture.

As a group, police officers are often called upon to act as witnesses to crimes and be the administrators of lineups, showups and other procedures that put citizens' eyewitness memory to the test. It seems obvious that police need to be made partners in eyewitness research so that wide dissemination of reliable eyewitness procedures can take place and valid training procedures for police and court officials can be developed and instituted. Another contribution to the literature of the apparent phenomenon of weapon focus effect was the introduction of a third condition, weapon inferred. The original two conditions were weapon present or absent. This third condition exists in many actual police cases but had never been tested experimentally before. By changing the paradigm of weapon focus effect from a binary state to a more continuous form, we were able to quantify what appears to be an intermediate stage of attentional diversion that had not been observed or measured previously. The results of this study will no doubt have implications for future WFE studies, regardless if police officers are a subject of investigation.

Experiment 2

Experiment 2 is designed to test whether having 1 as opposed to 2 perpetrators present during the commission of a crime affects the witness' ability to make a correct identification decision when shown a lineup. The hypotheses that will be evaluated by Experiment 2 are: 1. that there will be a difference between the number of correct photo a

identifications that police officers and citizens make; 2) that police officers and citizens will be differently affected by the presence of multiple perpetrators; 3) that police officers' and citizens' perceived stress levels and physiological stress levels will differ as result of officers being desensitized. The perceived stress that a police officer might feel when witnessing a stressful scenario might be less than a citizen. This might occur because police officers might become desensitized from repeated exposure to this type of situation. It is theorized that, even though perceived stress might differ, physiological (Gilmartin, 1990); 4) that perceived stress will coincide with correct lineup identification differently than physiological stress; and 5) that multiple perpetrators will degrade eyewitness memory equally for both groups of participants. Overall identification rates were tested as well as whether any differences existed between police officers and citizens. Several descriptive and parametric statistical tests were used to describe the data. Direct logistic regression was also used to examine the possible impact of certain factors on correct identification decisions. An illustration of the regression model is shown in Figure 5.

stress, e.g., pulse rate and galvanic skin response, might be similar in officers and citizens. In other words, officers might perceive less psychological stress because of desensitization but their bodies would still produce an autonomic response to the situation

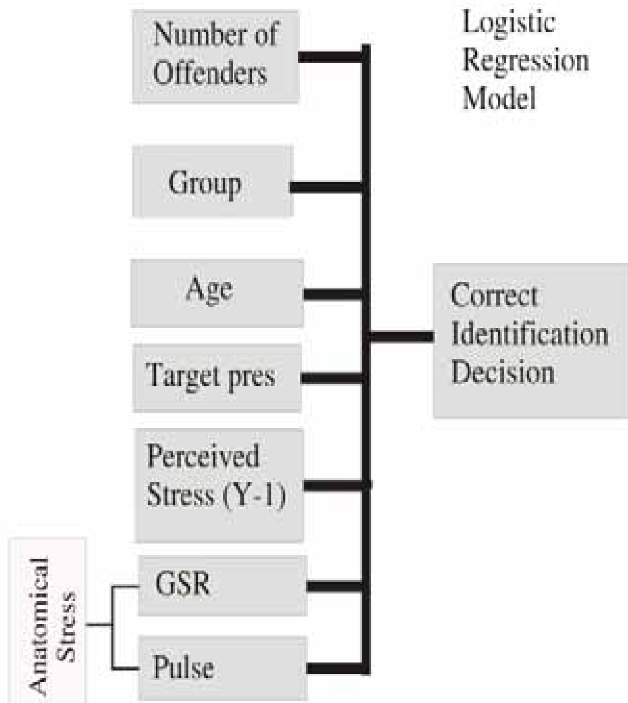


Figure 5. Logistic Regression Model for Experiment 2

Method

Participants

Three hundred and eight participants, 154 police officers and 154 citizens, were recruited for Experiment 2. Overall, the mean age for the two groups was 39.87 years old (SD =11.15). Citizen mean age was 41.44 years old (SD =12.80) and police officer mean age was 38.27 years old (SD =8.96). The citizen and police respondents were all from the same Northeastern state in the United States. Police officers came from 6 different agencies and the size of their departments' ranged from 49 to 611 sworn officers (M = 173.3, SD = 224.6). Mean length of service for police officers was 10.6 years (SD = 5.95). Police volunteers took part in the experiment as they participated in their regular in-service training. Citizen participants were volunteers recruited from the communities in which the police departments are located and also, undergraduate college students from a Northeastern university. All police officers were male and the citizen group consisted of 67 males and 87 females.

Design

Experiment 2 is a 2 (single offender or offender plus accomplice) x 2 (target present/target absent lineup) x 2 (police and citizen groups) between subjects factorial design. Correct lineup identification is the major dependent variable with additional dependent data to gauge participant stress and arousal.

Materials

Two videos were created showing which displayed either a scenario with a single offender or a two perpetrator scenario, with an offender plus an accomplice. The scenarios all involved a robbery being committed by either the lone offender or the pair. The robbery took place from the point of view of the victim. Two photo-lineups containing 8 photos each were created and shown to the participants. One lineup was target present for the perpetrator and the other was target absent. All lineups contained the accomplice whether he was involved or not. All other materials were the same as used in Experiment 1. The perpetrator or the replacement photo was always placed in lineup position 2 and the accomplice was always in lineup position 7. All other materials were identical to what was used in Experiment 1. As in Experiment 1, Y-1 pretesting was not conducted during our experiment due to time restrictions on participants' availability.

Procedure

Each participant was randomly assigned to view one of the two different video scenarios. After each person viewed their randomly assigned video, they were asked to complete Spielberger's State Anxiety Inventory, Form Y-1 to gauge their perception of the stress that they felt. No questions were asked by the researchers.

Once the individual completed the form, they were randomly shown one of the photo lineups described in materials above. As in Experiment 1, all lineups were displayed in a double-blind, sequential method on a laptop computer controlled by the participant. Participants were told that they were to attempt to identify the lone culprit in single culprit scenarios and the main culprit in two culprit situations. In the videos the main culprit talked and the accomplice did not except to say "yeah" several times. The instructions given to the subject were: "The individual(s) that you observed in the video may or may not be present in the photos that you are about to see. Even if you think that you recognize an individual, please continue looking at the other photos until there are none remaining." If the subject identified any of the lineup members, the data was recorded the subject identified any of the lineup members, their identification decision was recorded. Participants were debriefed, thanked for their time and dismissed.

Results

Pulse, GSR and Perceived Stress

Perceived stress, measured by administering Spielberger's State Anxiety Inventory directly after participants viewed the staged robberies, was significantly different between groups although not with a strong effect size (η^2=.01). Police had a mean stress score on the inventory that was lower (M = 43.3, SD =10.9) than citizens (M = 46.5, SD =14.3). This finding is generally consistent with our hypothesized difference in cue utilization between police and citizens, possibly due to some degree of desensitization in police officers.

In addition to Spielberger's State Anxiety inventory, two additional measurements were taken of factors that have been associated with stress and which increase during arousal: Galvanic Skin Response (GSR), reported state anxiety (perceived stress) and participants' maximum pulse during the experiment. The mean pulse at the start of the experiment, before viewing the robbery videos, was 72.8 (range = 62 to 85; SD = 4.3). During the viewing of the robbery video, which took approximately 45 seconds, pulse was continuously monitored and the highest pulse attained was recorded. The mean high pulse was 75.17 (range = 65 to 88; SD = 4.2). The mean GSR attained during the test was .94 (range = 0.0 to 2.00; SD = .610). As mentioned above, Spielberger's State Anxiety Inventory mean score was 43.2 and the range was 54 with a minimum of 22 and a high of 76 (SD = 13.06). This is compared to a sample of 1,883 working adults tested by Spielberger (1983), who reported normative scores on the state anxiety test of M=35.72 (SD = 10.40).

Stress and arousal. One-way between groups analyses of variance were conducted to explore the impact of the number of perpetrators on the three factors, Y-1; pulse and GSR associated with stress. No significant differences were recorded. The means are presented in Table 8. Y1, F (1,308) = .217, p=.642; GSR, F (1,308) = .781, p=.378; Pulse, F (1,308) = .1.28, p=.258

Table 8. *Mean stress values for perpetrator conditions.*

Condition		Y1	GSR	Pulse
1-perp	Mean	43.63	.911	74.91
	N	156	156	156
	SD	12.858	.616	4.20
2-perp	Mean	42.94	.973	75.45
	N	152	152	152
	SD	13.30	.60	4.21

An independent samples t-test (two-tailed) was conducted to compare the scores on our three measures associated with stress and arousal for police and citizens. Of the three measurements, only the GSR measurement did not show a significant difference between police (M = 39.38, SD =.769) and citizens (M = 47, SD =.617), t (305) = 1.03, p = .301. There was a significant difference in Y-1 scores between police (M = 39.38, SD = 9.5) and citizens (M = 47.2 SD = 14.8), t (319) = 5.495, p = .000. Pulse during the experiment also showed a significant difference between police (M = 74.44, SD = 9.54) and citizens (M = 75.9, SD = 14.8), t (282) = 3.14, p = .000. Because of the significant results for pulse and Y-1 anxiety, an independent samples t-tests (two-tailed) was conducted to gauge the effect of the absence or presence of an additional perpetrator on both pulse and anxiety scores. Pulse was not significantly different whether there were one or two perpetrators present, pulse, (M = 74.9, SD =4.2), t (305.7) = 1.133, p = .258, nor was the state anxiety score, (M = 42.9, SD =13.3), t (306) = .465, p = .642.

A two-way between groups analysis of variance was conducted to explore the potential interaction between being in the police or citizen group and the number of perpetrators on state anxiety scores as measured using form Y-1 of Spielberger's State Anxiety Inventory. In the initial analysis, Levene's test for equality of error variances was significant at the < .05 level suggesting that the variance of the dependent variable was not equal across groups (homogeneity of variance was violated). To compensate for this condition, we re-ran the analysis at a more stringent <.01 α level. The interaction effect between group and number of perpetrators on Y1 score was not statistically significant at this level, *F* (1,307) = .115, *p* = .735. Also, at the < .01 level, none of the variables reached a significant main effect.

To continue to explore interactions between group and number of perpetrators two additional two-way analyses of variance were conducted. The first, on pulse, showed similar results to what we earlier found for Y-1 in that Levene's test for equality of error variances was significant at the < .05 level suggesting that the variance of the dependent variable was not equal across groups (homogeneity of variance was violated). To again compensate for this condition, we re-ran the analysis at a more stringent <.01 α level. The interaction effect between group and number of perpetrators on pulse rate was not statistically significant at this level, *F* (1,307) = .469, *p* = .494. Also, at the < .01 level, none of the main effects were significant.

Finally, GSR was examined. The interaction effect between group and number of perpetrators on GSR measurement was not statistically significant at the .05 level, *F*

(1,307) = .239, p = .625. Also, at the < .05 level, none of the main effects were significant.

Identification Accuracy

Presence of additional perpetrator. A chi-square test for independence indicated a significant difference in accuracy between single and multiple perpetrator conditions, χ^2 (1, n=308) = 8.5, p = .003, Phi = .173. In single perpetrator scenarios, participants made correct identification decisions 73.7% of the time whereas, in two perpetrator scenarios, correct identification decisions were made 57.2%

A chi-square test for independence indicated no significant difference in accuracy between multiple perpetrator conditions with target absent lineups, χ^2 (1, n=154) = 3.92, p = .082, Phi = .160. In the multiple perpetrator condition, all participants were able to make a correct rejection 81.6% of the time and in the single perpetrator condition they correctly rejected the lineup 92.3% of the time.

When the perpetrator was present in the lineup, a chi-square test for independence indicated a significant difference in correct identification between the single (55.1%) and multiple perpetrator (32.9%) conditions, χ^2 (1, n=154) = 6.841, p = .009, Phi = .224. Everyone, police and citizens had higher accuracy when the target was present in multi-perpetrator condition.

Out of the eleven filler identifications that occurred in multi-perpetrator scenarios participants chose the accomplice 45.5% of the time. Recall that the accomplice was always in lineup position 7. This was the highest percentage of all filler photos with the next most frequently selected filler in a target present situation at 18%. In addition, when police selected a filler photo, they chose the accomplice 33% of the times and civilians 40% of the time.

Police v. citizens. When we examined identification decision performance between police and citizens using a chi-square test for independence , we found that there was no overall difference between police and citizens' identification decisions, χ^2 (1, n=308) = 1.741, p = .187, Phi = .082.

A chi-square test for independence showed no significant difference in identification decision performance when target absent lineups were shown to police and citizens, χ^2 (1, n=308) = .517, p = .472, Phi = .77. Additionally, when police and citizens, as one group, were shown target present lineups, no statistically significant difference emerged. Once again, a chi-square test for independence was used and yielded the following results: χ^2 (1, n=308) = 1.290, p = .256, Phi = .105.

When filler identifications were compared by police or civilian group across all perpetrator condition we found that when a target was absent, police made incorrect filler identifications 10.4% of the time and citizens made incorrect filler identifications 15.6% of the time, χ2 (1, n=154) = .517, p = .472, Phi =.077. When a target was present, police made incorrect filler identifications 7.8% of the time and citizens made incorrect filler identifications 13.0% of the time. χ2 (2, n=154) = 2.170, p = .338, Phi =.119. There were no significant differences between police and citizens in this comparison.

Next, we examined whether an additional perpetrator affected the accuracy of identification decisions in both target present and absent lineups (to make a correct identification decision, a participant was required to identify the target when he was present in the lineup or reject the lineup if he was absent). In a single perpetrator situation, irrespective of lineup target presence, police made incorrect filler identifications 5.1% of the time and citizens made incorrect filler identifications 9.0% of the time. χ^2 (2, n=156) = 1.43, *p* = .487, Phi =.096. When a second perpetrator was added police made incorrect filler identifications 13.2% of the time and citizens made incorrect filler identifications 19.7% of the time. χ^2 (2, n=152) = 1.39, *p* = .497, Phi =.095. There were no significant differences between the police and citizen groups in this comparison.

Table 9. *Percentage of correct identification decisions (frequency in parentheses) by group, lineup type and number of perpetrators – chi square between police and citizens.*

		Group				
	Overall	Police	Citizens	χ^2	*Df*	*p*
Target absent						
One perp	92.3 (72)	94.9 (37)	89.7 (35)	.722	1	.395
Two perp	81.6 (62)	84.2 (32)	78.9 (30)	.350	1	.554
Target present						
One perp	55.1 (43)	61.5(24)	48.7 (19)	1.29	1	.250
Two perp	32.9 (25)	36.8 (14)	28.9 (11)	.536	1	.464

Factors Contributing to Identification Decision Inaccuracy

Direct logistic regression was performed to assess the impact of several factors on the likelihood that participants would make a correct identification decisions while controlling for all the other variables. The model contained one dichotomous dependent variable, identification decision, and seven independent variables: (age; Y-1 score; pulse; GSR, target presence; number of perpetrators and group.) The full model containing all predictors was statistically significant, χ^2 (7, N= 315) = 86.261, *p*<.001, indicating that the model was able to distinguish between participants who made and who did not make a correct identification. The model, as a whole, explained 33.7% (Nagelkerke R Square) of the variance in identification correctness, and correctly classified 65.6% of cases. As shown in Table X, two of the independent variables made unique, statistically significant contributions to the model.

On average, the odds of making an incorrect identification decision after holding all other variables constant, when a lineup target is absent, are 5.440 times higher than when the target is present. In Experiment 1, there was a significant difference between the identification decisions of police and citizens. On average, the odds of making an incorrect identification decision (identification or rejection), after holding all other variables constant, when there are two perpetrators is 2.55 times higher than when one perpetrator is present..

As will be recalled, in Experiment 1 police showed a significant difference in choosing fewer incorrect fillers. In Experiment 2, that difference was not evident. In summary, when a lineup target is present in a lineup, the odds of making an incorrect identification decision are less. Also, having two perpetrators present causes a higher rate of incorrect identification decisions than when only one perpetrator is present.

Table 10. *Logistic regression mode showing the impact of various factors on correct identifications*

	β	S.E.	*Wald*	*df*	*p*	*Odds ratio*	95.0% C.I	
							Lower	Upper
Age	.020	.013	2.37	1	.123	1.02	.99	1.04
Y1 Score	-.022	.012	3.50	1	.061	.97	.95	1.00
Pulse	.002	.034	.00	1	.960	1.00	.93	1.07
GSR	-.004	.228	.00	1	.988	.99	.63	1.55
Target condition	2.303	.311	54.829	1	.000**	10.009	5.44	18.41
Multi Perp	.936	.287	10.66	1	.001**	2.55	1.45	4.47
Group	.573	.305	3.53	1	.060	1.77	.976	3.22

** p=<.001

An additional direct logistic regression was performed to assess the possible impact of other factors on the likelihood that participants would make a correct identification decisions (see table 11). The model contained five independent variables: (age; length of service; agency size; rank and tour segment.) The full model containing all predictors was statistically significant, χ^2 (9, N= 308) = 23.11, p<.006, indicating that the model was able to distinguish between participants who made and who did not make a correct identification. A Hosmer and Lemeshow test, which is a more strident test than the omnibus test (Hosmer & Lemeshow, 2000; Menard, 2001; Pallant, 2007) that is indicative of the goodness of model fit, was conducted and indicated that the model was not worthwhile (χ^2 = (8,321) 11.66, p = .154). The model, as a whole, explained 19.7% (Nagelkerke R Square) of the variance in identification correctness, and correctly classified 70% of cases. Based on the Hosmer and Lemshow test, the model is not significant. If however, we chose the omnibus test results, only one of the independent variables, age, would make a unique, statistically significant contribution to the model. Holding all other variables contant and within the range of ages tested, for every year of age, correct identifications would increase 1.068 times. This is interesting because length of service did not contribute to the model which indicates the effect occurs in both police and citizens.

Table 11. *Logistic regression model for additional variables in Experiment 2, showing possible impact of the factors on correct identification decisions.*

	β	S.E.	*Wald*	*df*	*p*	Odds ratio
Age	.066	.03	4.33	1	.037*	1.068
Length of service	-.081	.05	2.25	1	.133	.922
Agncy Size 50 to 100	-22.91	22788.63	5.15	3	.161	.000
Agncy Size 100 to 200	-21.53	22788.63	.00	1	.999	.000
Agncy Size > 200	-22.77	22788.63	.00	1	.999	.000
Rank Ptlmn	.544	.86	1.00	3	.800	
Rank Sergt	.578	.577	1.00	1	.316	1.783
Rank Lt and >	-19.78	22571.935	.00	1	.999	.000
Tour Segment	-.146	.388	.14	1	.706	.864

* $p < .05$

Discussion

Our first assumption, that there would be a difference between the number of correct identification decisions that police officers and citizens made, was unsupported by our results. Our second hypothesis, that police officers and citizens would be differentially affected by the number of perpetrators, was also rejected. There was not a significant difference between how the police and citizen groups' identification accuracy was affected, across both perpetrator conditions, for both target absent and target present lineups. Although police and citizens did not differ in eyewitness accuracy across conditions, there was a significant reduction in accuracy for both groups as the number of culprits in the experimental scenarios went from 1 to 2.

We hypothesized that police officers' and citizens' perceived stress levels and physiological stress levels would differ. Unlike having a gun pointed at you as in Experiment 1, it was unclear whether the number of perpetrators confronting a victim would cause a greater level of stress and consequent reduction in identification accuracy. What we found was that measures associated with physiological stress, pulse and GSR did not vary significantly between the numbers of perpetrators but that police had significantly lower pulse and Y-1 scores.

When the data from Experiment 2 were examined relative to our original assumptions, several interesting findings became clear. Our first hypothesis, that there would be a difference between the numbers of correct identification decisions made in the presence of multiple perpetrators, was supported by our results. As the number of perpetrators increased from one to two, identification accuracy diminished. The second hypothesis, that police officers and citizens would make correct identification decisions at different rates in multiple perpetrator situations, was not supported by our results. Indeed, identification decision performance for both groups was similar. Finally, we hypothesized that police officers' and citizens' perceived stress levels and physiological stress levels would differ. What was revealed in the experiment was that two of the three measures associated with stress and arousal, pulse and Y-1 state anxiety score did significantly differ between groups and GSR did not.

It is interesting to note that there were only three results that were statistically significant in Experiment 2. The first was a difference between police and citizen state anxiety scores with police showing lower state anxiety. The state anxiety results were not significantly affected by the presence or absence of an additional perpetrator. This result is similar to the difference in state anxiety in Experiment 1 where officers also had anxiety states which were elevated from normative, baseline samples, but lower than the citizen group.

The second item of significance, unlike the prior experiment, was a difference in pulse rate between groups during the viewing of the videos. Citizens' pulse rate was slightly but significantly higher while viewing a video although still within normally acceptable ranges for the population. It is interesting to note that the number of perpetrators did not statistically affect pulse rate.

The other finding that showed statistical significance was a difference in correct decisions when there was more than one perpetrator present. Participants made more correct identifications when there was only one perpetrator present. Unlike having a weapon present and pointed at the victim as in Experiment 1, the number of people present would not seem to have a contextual or novelty component. What it does seem to have is an environment in which the victim would need to encode and accurately remember two faces rather than one. The amount of time needed to accomplish efficient encoding of two faces or possibly confusing or combining the faces in memory might all be possible contributors to why accuracy was lower with two perpetrators present.

It is notable that none of the identification accuracy tests showed any difference between police officers and citizens. The current study had similar results to a study by

Owens, Rainey and Dysart (2008) in that, as the number of perpetrators rose, so did the inaccuracy of identification decisions. While examining the subject of accuracy, it should be noted that, according to Steblay (1997), because of the unbiased instructions (the perpetrator may or may not be in the lineup) given to participants in the current study, a lower error rate than when these instructions are not given, might by possible. Steblay reported that unbiased instructions could be responsible for approximately 250 fewer errors per 1,000 target absent lineups. This is important because Steblay also reports in the same study that biased instructions do not have the same impact on accuracy in target present lineups. The fact that the current research found no difference in accuracy between target present and target absent conditions might be a consequence of our participants receiving unbiased instructions exclusively. This might have caused a lower error rate between the target conditions if biased instructions had been give. From a policy standpoint, of the six police departments that took part in Experiment 2, only one routinely administers unbiased lineup instructions to citizens viewing lineups. It can be assumed that those departments that are not administering unbiased instructions, which are recommended by the U.S. Department of Justice (Technical Working Group for Eyewitness Evidence, 1999) and which could easily be implemented with a written policy, might be experiencing higher error rates in real world lineup identification decisions.

The fact that there were no significant differences between police and citizens' overall identification rates in Experiment 2, suggests that the processes that both police and citizens use to acquire and store information at a crime scene are more similar than different, even in light if police officers' more frequent exposure to crime scenes. From a theoretical perspective, both of the groups were distracted equally by more than one perpetrator being present. Since there is no published theoretical research on viewing and encoding multiple faces simultaneously, we must look to possible reasons for the effect. Some possible contributory factors as to why the number of offenders is of consequence to what an eyewitness is able to remember about more than one culprit involved in a crime, are diminution of attention, unconscious transference and stimulus sampling theory. Out of the these three possibilities, we have reduced the chances of stimulus sampling by ensuring that both culprits in the videos were similar in facial appearance, coloring and build and dressed the same. Unconscious transference was controlled for by the high number of participants that took part in the experiment. If one person thought, subconsciously, that our culprits looked like someone they saw on a routine basis but did not know well, certainly over 300 participants did not experience similar confusion. This leaves us with diminution of attention. Simply, trying to view and store more information at one time, within a short temporal frame, is harder to do and results in seeing less detail and remembering less than one would if only a single perpetrator was present. Both the facial recognition literature and research on weapon focus effect might make some contribution to understanding why our participants made fewer correct identification decisions as the number of culprits rose. Bruce (1988), in discussing how faces are recognized, suggests that the process of how faces are encoded for memory storage is not well understood. Whether the face is viewed as a pattern, a collection of parts or as an entirety and whether context affects recognition might be factors in later identification. If our participants had the pattern recognition leading to facial encoding process interfered with in some way by the presence of another culprit, it might be a contributing factor to the lower identification rates that we recorded when a second perpetrator was present. On the other hand, the research that has been done on weapon focus effect might also offer some explanation for our multiple perpetrator results. Revelle and Loftus (1990) proposed a partial explanation for WFE which they called the “tick rate hypothesis.” They described human attention as being analogous to the clock speed of a computer where there are only a finite number of cycles or “clock ticks” available for processing. When those attentional “ticks” are theoretically divided between more culprits within a short amount of time, it would be logical to assume that each culprit would receive less encoding. As mentioned earlier, with multi-perpetrator crimes on the rise, it is important to know what effect this situation might have on victims and witnesses so that lineup policies and procedures could be informed by research to arrive at the most accurate results.

Limitations of Experiments 1 and 2

Wells (1978) described several methodological biases that limit the external validity or generalizability of any staged crime used in eyewitness research to the real world. The richness and complexity of a true crime scene as opposed to a mock crime shown on a screen is subject to the same attenuation that might be present during in-vitro experimentation relative to a true in-vivo procedure. Any experiment using staged crimes cannot fully duplicate the look and feel of a real crime scene in which a person might feel that their life is threatened. Because weapons are the subject of Experiment 1, it would have been extremely difficult - from an IRB perspective - to present participants with live perpetrators who were brandishing actual firearms. In addition to the fact that the safety of the confederate perpetrator may be at issue, ethical regulations of research studies state that participants may not be placed in situations which will likely cause them harm or put them in a position of extreme stress or fear. Therefore, in Experiment 1, we used a well-accepted technique of presenting the armed robber via a video (Steblay, 1992). We expect that the physical and psychological reaction that participants will have as a result of seeing a robbery on screen will not be the same intensity as would be experienced in a real life robbery. However, other researchers have found the weapon focus effect to be present when using video presentation of a crime (Cutler, 2004; Kramer, Buckhout, & Eugenio, 1990; Stanny & Johnson, 2000; Tooley, Brigham, Maass, & Bothwell, 1987). In addition, pilot testing that we have conducted has shown a pronounced weapon effect in both police and citizens. In addition, we expect, based on the theories described above, that additional stress would only cause our participants to be less accurate. And thus, although video presentation is not a perfect way to measure the weapon focus effect, it has proven to yield acceptable results to the degree necessary for the current study.

An additional factor that might influence participants' lineup decisions is that there are no consequences for choosing or making a mistake (Haber & Haber, 2001). Unlike real cases, a wrong choice will not send an innocent person to jail or set a guilty person free. In his discussion of methodological biases referenced above, Wells (1978) points out that no research existed where the participant was also the victim. Although it is not possible within the standards of ethical research to victimize a research participant, it is hoped that the first person point-of-view simulations, life size displays and realistic sound will better approximate the reality of a real crime event than had been previously possible using third-person observation on small computer screens.

Also among the list of limitations is that of memory retention over time (Brewer, Weber, Semmler, & Williams, 2005; Cutler, et al., 1987; Deffenbacher, Bornstein, McGorty, & Penrod, 2008). In the real world, it is not uncommon for an eyewitness to wait days or weeks before viewing a lineup. Because of practical limitations in scheduling police officers in the study, only fifteen minutes was allowed to elapse between viewing the crime simulation and the lineup. However, the effects of time delay on memory are well established in the literature and consistently show a negative effect on memory. Thus, if longer delay intervals had been used in the current study, the only possible effects it could have had would have been negative. In other words, identification accuracy would have been worse than in the current study.

The sample of police officers used in this research was a convenience sample which relies on available subjects in the agencies that agreed to participate in the research. As such, the sample may not have been entirely representative of the general population of police officers in the United States. As an example, women represented 11.3% of all American police officers in 2004 (Reaves, 2004) but the police departments that participated in this study had less than a 3% female representation. Because the availability of female officers was so low, even an oversampling strategy was not possible and thus the study was conducted with only male officers. In the eyewitness literature, however, there has not been a demonstrated interaction or main effect of witness gender on identification accuracy (Wells et al., 2006).

General Discussion and Conclusions

Looking at our hypotheses and what we took away from the current study was interesting. We found that police and citizens make about the same rate of accuracy errors in lineup identifications both in weapon conditions and multi-perpetrator conditions. This finding was contrary to conventional wisdom that police might make better witnesses. Indeed, as prior literature suggests, (Ainsworth, 1981; Christianson, et al., 1998; Yarmey, 1986; Yuille, 1984) police might, under some circumstances have better recall of events. Memory of events however, is not the same as memory for faces, especially when the encoding of those memories might be interrupted by factors that cause stress or divide attention like weapons and multiple people present during a crime. The point is that when a police officer identifies a suspect or testifies in court, the officer should be given the same weight that any other witness is given when it comes to identification accuracy.

We also learned that prior exposure to weapons does not remove whatever effect it is that causes them to reduce identification accuracy. Police were as subject to WFE (and the presence of multiple culprits in Experiment 2) as were citizens. However, the reason(s) that police produced fewer filler identifications in target present conditions is not understood. If further research can isolate the cause, it has the potential to reduce false positives in lineups in general. Wells (2005) very aptly noted that all false positives inhibit the search for the actual criminal.

Police indeed did experience lower arousal, measured by Spielberger's Y-1 inventory, in Experiments 1 and 2. This is important to recognize since they did not exhibit greater identification accuracy. This set of findings leaves open the question as to whether arousal is the only moderating factor in WFE. Why did a lower arousal rate result in a similar identification rate between police and citizens? Perhaps the answer to that question lies in the fact that the Y-1 inventory measures perceived stress. Since the

GSR and pulse measurements for police and citizens were similar, the study might possibly have illustrated that the cue utilization curve is more tied to psychological factors than physiological ones. Further research into these areas in order to understand this effect might help to disambiguate the causes of WFE.

Police officers handle guns on a daily basis yet they were as influenced by WFE as citizens who have much more limited exposure to weapons. Since guns are not novel to police officers, why was their presence the basis for the same level of diminishing accuracy as they were with citizens? Perhaps future research can illuminate, through techniques such as fMRI and digital eye-tracking, the reasons why identification accuracy becomes lower when implied or actual weapons are present as in Experiment 1, or when there is more than one person committing a crime, as in Experiment 2. As discussed previously, does context take precedence over novelty? Does the context in which a gun is viewed have more weight in explaining WFE than whether the witness is exposed to guns on a more frequent basis? Does a weapon become novel in a whole new way when its context is changed and it becomes a threat instead of a tool?

One of the major questions that we asked was if there was a difference in the way that police and citizens were affected, if at all, by weapons and multiple-perpetrators. The fact that there were no significant differences between police and citizens' overall identification rates in either Experiment 1 or Experiment 2, suggests that the processes that both police and citizens use to acquire and store information during crimes, similar to those used in the current study, are more alike than different.

These experiments attempted to address two broad areas of the criminal justice system where misinterpretations might mean that innocent people are jailed and guilty individuals remain at large to possibly commit other crimes. First, there is a dichotomy in the working literature on eyewitness identification. Most academic eyewitness literature is based on empirical studies and recognizes the value of such methods as sequential lineups and double-blind administration (Wells, Memon, & Penrod, 2006; Wells, Seelau, et al., 1994; Wells, et al., 1998). Professional police literature, such as Police Chief Magazine, sometimes takes a viewpoint that differs from academic literature (Wells, et al., 2000). The academic journal Law and Human Behavior devoted an entire issue to discourse on the controversial Illinois Pilot Study (Cutler & Kovera, 2008; Mecklenburg, Bailey, & Larson, 2008; Steblay, 2008). Conversely, the magazine, Police Chief, read by law enforcement administrators and potentially a million police officers in the approximately 18,500 police departments in the United States, only printed one side of the controversial study. Police Chief is not peer reviewed and has run very few items on eyewitness identification. One recent article was by the director of the Illinois Pilot Study (Mecklenburg, Larson, & Bailey, 2008) presenting a viewpoint strongly favorable toward their findings, even though valid questions have been presented questioning the methodology of the study (Wells, 2008; Winzeler, 2008).

When the publishers of Police Chief Magazine were contacted to ask for an opportunity to present the countervailing viewpoint they had the author of the article contact the requesters to ask if there was an issue (DeCarlo & Dysart, 2008). Instead of being able to publically engage in discourse on the subject, only a private conversation ensued and the potential for law enforcement policy makers to hear both sides of the discussion was lost.

In addition to this disconnect between research and practice described above, certain issues of training based on anecdotal beliefs might also be informed by the current study. In spite of several research studies that have examined police v. citizen memory, courts, police officers and the public still give greater than normal credibility to the observational powers of police officers. This attribute is sometimes made because of the assertion that police officers have special training in observation and identification which increases their potential to be accurate eyewitnesses. Contrary to conventional wisdom, training for recruits in the area of eyewitness identification was not part of the curriculum in police academies in eight states that were surveyed. There is no reason to believe that the curriculum of academies in other states is substantially different in failing to include these topics in their training programs.

If eyewitnesses have been accorded great validity by the courts, police officers are given even more legitimacy in this role by both the legal system and by the public (Ainsworth, 1981; Correll, et al., 2007; Yarmey, 1986, 1998; Yuille, 1984). The findings of this study have the potential to both make a distinct contribution to the existing literature and to hopefully inform police training and practice. The number of studies that have access to populations of police officers is relatively small. To be able to work with practitioners and learn how the important job that they do might be improved is a valuable opportunity. For policy to change for the better, especially to reduce the numbers of errors that have made eyewitness evidence the largest cause of mistaken convictions (Ferrero, 2009; Scheck, 2008), practice must take part in and be aware of research results.

For substantive change to take place in this area the entire police paradigm must be involved, from officers in basic training to chiefs who set and affect policy made by legislators. In Connecticut, as an example, prosecutors in 2007 had been moving slowly toward reform in eyewitness identifications. When the Illinois Pilot Study, complete with its methodological idiosyncrasies, was released, state prosecutors advised an immediate return to older styles of eyewitness identifications. A bill dealing with eyewitness reform, then on the Connecticut legislature's docket ("An act concerning eyewitness identification," 2008) was lobbied against by Connecticut Police Chiefs Association (CPCA) and defeated on the floor. Had that association and the Connecticut State Legislature been aware of research in this study, it is possible that they would have had a broader, more varied assortment of literature on which to base their decisions. Instead, over the last few years, over two hundred inmates have requested the assistance of the Connecticut Innocence project. It is sobering that if only 10% of them are ever exonerated, than ten needless lifetimes might have been spent in prison based on poor eyewitness identifications, and other systemic issues that plague our criminal justice system. Brewer and Williams (2005) make the observation that although it intuitively seems like judges and police are the policy makers when it comes to eyewitness identification procedures, in reality, it is actually prosecutors who dictate policy in this arena.

Research such as this study is of vital importance to both understand the environment that eyewitness decisions are made in and to inform both court officers and police how to improve eyewitness methods, understand when they may not be reliable and deliver justice in a fair manner which is based on the best information that they can obtain. The most important policy recommendations that are a consequence of the results of this study have to do with training and wide dissemination and implementation of recommendations by the National Institute of Justice to make eyewitness evidence more accurate. Training on how to construct lineups that have less chance to produce incorrect identifications should be given at the basic recruit level in police academies. This training should consist of how to make sequential, double blind lineups and how to deliver admonitions to the witness that the suspect may or may not be present. In addition, senior police officers should also receive recurrent training in the same techniques. Prosecutors and other court officers also need to be educated on the findings of research and the recommendations of the NIJ. In private communications with prosecutors, there is worry that to adopt new procedures will be a tacit admission that the old methods were incorrect, thus opening up closed cases for re-inspection. Of course, this is a circular argument and the perpetuation of a status quo that has high potential for eyewitness error is the very real result.

Academics must make research accessible to practitioners who do not read peer reviewed technical journals. The National Institute of Justice, along with other federal agencies, must make funding available to researcher/practitioner teams to explore ways to increase eyewitness accuracy and at the same time be mindful of how investigations are done in the field and the limitations and restrictions that are sometime imposed by the many social factors that influence police investigations.

Appendix A. Spielberger's State Anxiety Form

For use by Psychological researchers.

SELF-EVALUATION QUESTIONNAIRE **STAI Form Y-1**

Please provide the following information:

Name____Date______________S_____

Age____________________ Gender (*Circle*) **M F** T_____

DIRECTIONS:

A number of statements which people have used to describe themselves are given below. Read each statement and then blacken the appropriate circle to the right of the statement to indicate how you feel *right* now, that is, *at this moment*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

	NOT AT ALL	SOMEWHAT	MODERATELY SO	VERY MUCH SO
1. I feel calm	**1**	**2**	**3**	**4**
2. I feel secure	**1**	**2**	**3**	**4**
3. I am tense	**1**	**2**	**3**	**4**
4. I feel strained	**1**	**2**	**3**	**4**
5. I feel at ease	**1**	**2**	**3**	**4**
6. I feel upset	**1**	**2**	**3**	**4**
7. I am presently worrying over possible misfortunes	**1**	**2**	**3**	**4**
8. I feel satisfied	**1**	**2**	**3**	**4**
9. I feel frightened	**1**	**2**	**3**	**4**
10. I feel comfortable	**1**	**2**	**3**	**4**
11. I feel self-confident	**1**	**2**	**3**	**4**
12. I feel nervous	**1**	**2**	**3**	**4**
13. I am jittery	**1**	**2**	**3**	**4**
14. I feel indecisive	**1**	**2**	**3**	**4**
15. I am relaxed	**1**	**2**	**3**	**4**
16. I feel content	**1**	**2**	**3**	**4**
17. I am worried	**1**	**2**	**3**	**4**
18. I feel confused	**1**	**2**	**3**	**4**
19. I feel steady	**1**	**2**	**3**	**4**
20. I feel pleasant	**1**	**2**	**3**	**4**

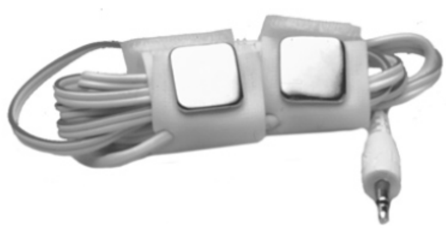
-72-

Appendix B. Concord Medical Instruments Fingertip pulse Oximeter used to measure pulse rate

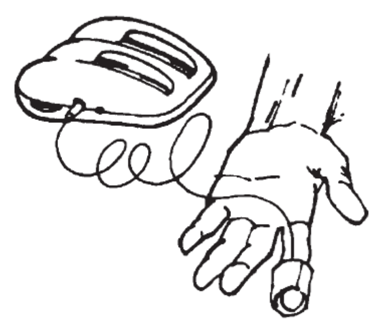
Appendix C. Biomedical Instruments GSR-2 Galvanic Skin Response System used to measure Galvanic Skin Response.

Galvanic Skin Response (GSR) Collection device

GSR meter to measure skin resistance



Electrodes to pick up electro
dermal response from fingers



How finger is placed in
electrode

Specifications

Skin resistance range 1,000 ohms - 3,000,000 ohms.

Variable frequency range 0 to 40,000 Hz.

9 volt battery. Battery current in use < 2.0 mA.

2.5mm output jack for 8 ohm earphone and output meter.

Protected Pin input jack for thermistor and electrodes.

Appendix D. Consent form.

Informed Consent Form

You are invited to participate in a research study entitled "An Experiment in Eyewitness Memory." The purpose of this research is to test memory under different conditions. We plan to enroll 500 participants into this study. If you decide to participate, you will be asked to take part in routine police video training and to attempt a photo line-up identification. Participation should take about1 hour for duration of 1 day.

There are no foreseeable risks of participation in this study. The possible benefits to you are by possibly contributing to the improvement of police training procedures. The potential benefits to society are the possibility of decreasing inaccuracy in eyewitness identifications.

Your participation in this study is completely voluntary. You have a right to refuse to participate without consequences. If you decide not to participate your decision will not affect your relationship with John Jay College or the Branford Police Department.

If you decide to participate you may discontinue participation at any time. You may refuse to answer any specific questions or refuse to engage in any task at any time during the study. Withdrawal or refusing to answer specific questions or engage in specific tasks will not result in any consequences to you and will not affect your relationship with John Jay College or the Branford Police Department.

Information gathered from you will be coded and anonymous. All data will be kept on file but will not include names or other identifiers that would connect any individual to the data.

Your signature below means that you have read this consent form, that you fully understand the nature and consequences of participation and that you have had all questions regarding participation in this study answered satisfactorily. If you have further questions about this research please feel free to contact the Principle Investigator, John DeCarlo, at 203-481-4241 or Dr. Jennifer Dysart at jdysart@jjay.cuny.edu.

If you have any questions regarding your rights as a research participant please feel free to contact the John Jay Institutional Review Board Office at
jj-irb@jjay.cuny.edu, or (212) 237-8961.

____________________________ ____________________________
Participant Name (print) Participant Signature

____________________________ ____________________
Principle Investigator/Research Staff Signature Date

IRB Approved
Date: 04/27/09

Appendix E. Institutional Review Board (IRB) approval.

DeCarlo JJ-09-069

frequency or degree of severity greater than that anticipated. In addition, the principal investigator must report any event or series of events that prompt the temporary or permanent suspension of a research project involving human subjects or any deviations from the approved protocol.

Amendments/Modifications: All amendments/modifications of protocols involving human subjects must have prior IRB approval, except those involving the prevention of immediate harm to a subject. Amendments/modifications for the prevention of immediate harm to a subject must be reported within 24 hours to the IRB.

Stipulations: NONE

If you have any questions, please do not hesitate to contact Carina Quintian in the IRB Office at 212.237.8961.

Good luck on your project.

cc: Dr. Dysart

Sign the Verification Statement below. Return the original signed copy of this memo to the IRB Office and retain a copy for your records. The IRB Office must receive a copy of the signed verification statement before research may begin.

VERIFICATION:

BY SIGNING BELOW, I ACKNOWLEDGE THAT I HAVE RECEIVED THIS APPROVAL AND AM AWARE OF, AND AGREE TO ABIDE BY, ALL OF ITS STIPULATIONS IN ORDER TO MAINTAIN ACTIVE APPROVAL STATUS, INCLUDING TIMELY SUBMISSION OF CONTINUING REVIEW APPLICATIONS AND PROPOSED PROTOCOL MODIFICATIONS, AS WELL AS PROMPT REPORTING OF ADVERSE EVENTS, SERIOUS UNANTICIPATED PROBLEMS, AND PROTOCOL DEVIATIONS. I AM AWARE THAT IT IS MY RESPONSIBILITY TO BE KNOWLEDGEABLE OF ALL FEDERAL, STATE AND UNIVERSITY REGULATIONS REGARDING HUMAN SUBJECTS RESEARCH INCLUDING CUNY'S FEDERALWIDE ASSURANCE (FWA) WITH THE DEPARTMENT OF HEALTH AND HUMAN SERVICES OFFICE OF HUMAN RESEARCH PROTECTIONS.

Mark DeCarlo 5/11/09

Signature of Principal Investigator Date

Jennifer Dysart 5/11/09

Signature of Faculty Advisor for Student Research Date

References

An act concerning eyewitness identification, 5832, State of Connecticut General Assembly, 02765 2675 JUD Cong. Rec. (2008).

Ainsworth, P. B. (1981). Incident Perception by British Police Officers. *Law and Human Behavior, 5*, 231-236.

Aldrich, J., H., & Nelson, F., D. (2007). *Linear Probability, Logit, and Probit Models*. Thousand Oaks, CA: Sage Publishers.

Anderson, J. L. (1999). *Determinants of accuracy in cross-racial identification.* Ph.D., City University of New York, United States -- New York. Retrieved from http://proquest.umi.com/pqdweb?did=730187751&Fmt=7&clientId=31967&RQT=309&VName=PQD

Associated Press. (2008). Police sorry for fingering wrong person, *Star Telegram*.

Awad, A. (2001). Different Responses of Ear and Finger Pulse Oximeter Wave Form to Cold Pressor Test. *Anesthesia & Analgesia, 92*, 1483-1486.

Baker, L. M., & Taylor, W. M. (1954). The relationship under stress between changes in skin temperature, electrical skin resistance, and pulse rate. *Journal of Experimental Psychology, 48*, 361-366. doi: 10.1037/h0057145

Baker, L. M., & Taylor, W. M. (1955). "The Relationship Under Stress Between Changes in Skin Temperature, Electrical Skin Resistance, and Pulse Rate": Erratum. *Journal of Experimental Psychology, 49*, 300. doi: 10.1037/h0038809

Bartlett, J. C., Memon, A., Seipel, A., Hulse, L., & Searcy, J. (2003). *Crime characteristics affect lineup choices by young and older adult*. Paper presented at the Fifth Biennial Meeting of the Society for Applied Research in Memory and Cognition, Aberdeen, Scotland.

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior, 25*, 475-491.

Boggs, L. P. (1904). An experimental study of the physiological accompaniments of feeling. *Psychological Review, 11*, 223-248. doi: 10.1037/h0072046

Brandon, R., & Davies, C. (1973). *Wrongful imprisonment*. London: Allen & Unwin.

Brear, M. (2002). *The effects of violence and multiple perpetrators on eyewitness identification with visual and audio-visual lineups.* Thesis (M A), University of Guelph, 2002.

Brewer, N., Weber, N., Semmler, C., & Williams, K. D. (2005). Eyewitness Identification *Psychology and law: An empirical perspective.* (pp. 177-221): Guilford Press: New York.

Brewer, N., & Williams, K., D. (2005). *Psychology and Law: An empirical perspective*. New York: The Guilford Press.

Brigham, J. C., & Cairns, D. L. (1988). The effect of mugshot inspections on eyewitness identification accuracy. *Journal of Applied Social Psychology, 18*, 1394-1410. doi: 10.1111/j.1559-1816.1988.tb01214.x

British House of Commons. (1974). *In the matter of Mr. Laszlo Virag*. London, UK: Sitting of 8 April 1974 Series 5 Vol. 872

Bruce, V. (1988). *Recognising faces*. London, U.K.: Lawrence Erlbaum Associates Publishers.

Bruning, N. S., & Frew, D. R. (1987). Effects of exercise, relaxation, and management skills training on physiological stress indicators: A field experiment. *Journal of Applied Psychology, 72*, 515-521. doi: 10.1037/0021-9010.72.4.515

Buckhout, R., & et al. (1974). Determinants of eyewitness performance on a lineup. *Bulletin of the Psychonomic Society, 4*, 191-192.

Bureau of Justice Statistics. (1998). Criminal Victimization in the United States, 1996 Statistical Tables: National Crime Victimization Survey. Washington, D.C.: U.S. Department of Justice.

Bureau of Justice Statistics. (2007). Homicide trends in the United States Retrieved 10/5/2009, 2009, from http://www.ojp.usdoj.gov/bjs/homicide/homtmd.htm.

Bureau of Justice Statistics. (2008). Criminal Victimization in the United States, 2006 Statistical Tables: National Crime Victimization Survey. Washington, D.C.: U.S. Department of Justice.

Campbell, D., T, & Stanley, J., C. (1963). Experimental and Quasi-Experimental Designs for Research. In N. L. Gage (Ed.), *Handbook of Research on Teaching*. Boston: Houghton Mifflin.

Carden, N. A. (1989). *Effects of training and experience on the expression of own-race bias by law enforcement lineup constructors.* 49, ProQuest Information & Learning: US.

Carpenter, R., H. (1988). *Movements of the eyes* (2nd ed.). London: Pion, Ltd.

Christianson, S.-Å. (1997). On emotional stress and memory: We need to recognize threatning situations and we need to forget unpleasant experiences. In D. G.

Payne & F. G. Conrad (Eds.), *Intersections in Basic and Applied Memory Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Christianson, S.-Å., Karlsson, I., & Persson, L. G. W. (1998). Police personnel as eyewitnesses to a violent crime. *Legal and Criminological Psychology, 3*, 59-72.

Clark, S. E., & Wells, G. L. (2008). On the diagnosticity of multiple-witness identifications. *Law and Human Behavior, 32*, 406-422. doi: 10.1007/s10979-007-9115-7

Clifford, B. R., & Hollin, C. R. (1981). Effects of the Type of Incident and the Number of Perpetrators on Eyewitness Memory. *Journal of Applied Psychology, 66*, 364-370.

Cook, T., D., & Campbell, D., T. (1979). *Quasi-Experimentation, Design & Analysis Issues for Fielf Settings*. Boston: Houghton Mifflin Co. .

Coppock, H. W. (1955). Responses of subjects to their own galvanic skin responses. *The Journal of Abnormal and Social Psychology, 50*, 25-28. doi: 10.1037/h0047355

Correll, J., Park, B., Judd, C., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*, 1314–1329.

Correll, J., Park, B., Judd, C., Wittenbrook, B., Sadler, M., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology, 92*, 1006-1023.

Cross, J. F., Cross, J., & Daly, J. (1971). Sex, race, age and beauty as factors in recognition of faces. *Perception and Psychophysics, 10*, 393-396.

Cutler, B. L. (2004). *Overview of Estimator Variables: Findings from Research on the Effects of Witness, Crime and Perpetrator Characteristics on Eyewitness Accuracy*. Paper presented at the Reforming Eyewitness Identification: Convicting the Guilty, Protecting the Innocent Conference, Benjamin N. Cardozo School of Law, Yeshiva University New York, NY.

Cutler, B. L., Berman, G. L., Penrod, S., Fisher, R. P., Ross, D. F., Read, J. D., et al. (1994). Conceptual, practical, and empirical issues associated with eyewitness identification test media *Adult eyewitness testimony: Current trends and developments.* (pp. 163-181): Cambridge University Press: New York.

Cutler, B. L., & Kovera, M. B. (2008). Introduction to commentaries on the Illinois Pilot Study of lineup reforms. *Law and Human Behavior, 32*, 1-2. doi: 10.1007/s10979-007-9120-x

Cutler, B. L., & Penrod, S. (1995). *Mistaken identification : the eyewitness, psychology, and the law*. Cambridge ; New York: Cambridge University Press.

Cutler, B. L., Penrod, S. D., & Martens, T. K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72*, 629-637. doi: 10.1037/0021-9010.72.4.629

DeCarlo, J., & Dysart, J. E. (2008). [Personal communication with Charles Higginbotham].

Deffenbacher, K., Bornstein, B. H., Penrod, S., & McGorty, E. K. (2004). A Meta-Analytic Review of the Effects of High Stress on Eyewitness Memory. *Law and Human Behavior, 28*, 687-706. doi: 10.1007/s10979-004-0565-x

Deffenbacher, K., & Loftus, E. (1982). Do jurors share a common understanding concerning eyewitness behavior? *Law and Human Behavior, 6*, 15-30.

Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14*, 139-150. doi: 10.1037/1076-898x.14.2.139

Dempsey, J., L., & Pozzulo, J. D. (2008a). Identification accuracy of eyewitnesses for a multiple perpetrator crime: ex-amining the simultaneous and elimination lineup procedures. *American Journal of Forensic Psychology, 26*.

Dempsey, J. L., & Pozzulo, J. D. (2008b). Identification accuracy of eyewitnesses for a multiple perpetrator crime: Examining the simultaneous and elimination lineup procedures. *American Journal of Forensic Psychology, 26*, 67-81.

Devlin, L. P. (1976). Report to the secretary of state for the Home Department of the departmental committee on evidence of identification in criminal cases. London: Her Majesty's Staionery Office.

DiCosmo, B. (2009). Police identify wrong suspect in high speed chase, *Southeast Missourian*. Retrieved from http://www.semissourian.com/story/1566532.html

Du Toit, J. M. (1956). Skin conductance and peripheral vascular reactions. *Journal of Experimental Psychology, 52*, 392-397. doi: 10.1037/h0043055

Dysart, J. E., Lindsay, R. C. L., Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. (2007). The effects of delay on eyewitness identification accuracy: Should we be concerned? *The handbook of eyewitness psychology, Vol II: Memory for people.* (pp. 361-376): Lawrence Erlbaum Associates Publishers: Mahwah.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Ebbinghaus , H. (1885). *Memory: A Contribution to Experimental Psychology* (H. Ruger & C. Bussenius, Trans.). New York: New York by Teachers College, Columbia University.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review, 57*, 94-107.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2008). *Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses*. Mannheim, Germany.

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

Federal Bureau of Investigation Criminal Justice Information Services Division. (2008). Crime in the United States. Washington, DC: United States Department of Justice.

Ferrero, E. (2009). Eyewitness Identification Reform. New York: The Innocence Project.

Garrett, B., L. (2008). Judging Innocence. *Columbia Law Review, 108*, 101-190.

Geer, J. H., & Klein, K. (1969). Effects of two independent stresses upon autonomic responding. *Journal of Abnormal Psychology, 74*, 237-241. doi: 10.1037/h0027148

Geiselman, R. E., Haghighi, D., & Stown, R. (1996). Unconscious transference and characteristics of accurate and inaccurate eyewitnesses. *Psychology, Crime & Law, 2*, 197-209.

Geiselman, R. E., MacArthur, A., & Meerovitch, S. (1993). Transference of perpetrator roles in eyewitness identifications from photoarrays. *American Journal of Forensic Psychology, 11*, 5-15.

Gilmartin, K. (1990). The brotherhood of biochemistry: Its implications for a police career. In H. E. Russell & A. Beigel (Eds.), *Understanding Human Behavior for Effective Police Work* (3rd ed., pp. 397-418). New York: Basic Books.

Haber, R. N., & Haber, L. (2001). *Eyewitness Accuracy When Making an Identification in a Lineup*. Paper presented at the 2001 Annual Convention of the Psychonomics Society, Santa Cruz, CA.

Hagan, S. (2009). *An analysis of eyewitness non-identifications in actual criminal cases.* Psy.D., Alliant International University, Fresno, United States -- California. Retrieved from http://proquest.umi.com/pqdweb?did=1694937751&Fmt=7&clientId=31967&RQT=309&VName=PQD

Hardy, L. (2007). *Regression with dummy variables*. Thousand Oaks, CA: Sage Publishers.

Hasel, L. E., & Kassin, S. M. (2009). On the Presumption of Evidentiary Independence: Can Confessions Corrupt Eyewitness Identifications? *Psychological Science, 20*, 122-126.

Hope, L., & Wright, D. (2007). Beyond unusual? Examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology, 21*, 951-961.

Horvath, F. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception. *Journal of Applied Psychology, 63*, 338-344. doi: 10.1037/0021-9010.63.3.338

Horvath, F. (1979). Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response. *Journal of Applied Psychology, 64*, 323-330. doi: 10.1037/0021-9010.64.3.323

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.

Huff, C. J. (2004). Wrongful convictions: The American experience. *Canadian Journal of Criminology and Criminal Justice,, 46*, 107-120.

Hulse, L. M., & Memon, A. (2006). Fatal impact? The effects of emotional arousal and weapon presence on police officers' memories for a simulated crime. *Legal and Criminological Psychology, 11*, 313-325.

Innocence Project. (2010a). Eyewitness Identification Reform Retrieved 3/17/2010, 2010, from http://www.innocenceproject.org/Content/165PRINT.php

Innocence Project. (2010b). Innocence project. Retrieved 10/20/2010, 2010, from http://www.innocenceproject.org/causes/mistakenid.php

Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research: A new survey of the experts. *American Psychologist, 56*, 405-416.

Kimmel, H. D., & Hill, F. A. (1961). A comparison of two electrodermal measures of response to stress. *Journal of Comparative and Physiological Psychology, 54*, 395-397. doi: 10.1037/h0047767

Knoke, D., & Burke, P., J. (2007). *Log-Linear Models*. Thousand Oaks, CA: Sage Publications.

Kramer, T. H., Buckhout, R., & Eugenio, P. (1990). Weapon Focus, Arousal, and Eyewitness Memory: Attention Must Be Paid. *Law and Human Behavior, 14*, 167-184.

Kramer, T. H., Buckhout, R., Fox, P., Widman, E., & Tusche, B. (1991). Effects of stress on recall. *Applied Cognitive Psychology, 5*, 483-488. doi: 10.1002/acp.2350050603

Krouse, F. L. (1981). Effects of pose, pose change and decay on face recognition performance. *Journal of Applied Psychology, 66*, 651-654.

Leippe, M. R. (1994). The appraisal of eyewitness testmony. In D. F. Ross, J. D. Read & M. P. Toglia (Eds.), *Adult eyewitness testimony*. London: Cambridge University Press.

Levine, F. J., & Tapp, J. L. (1973). The psychology of criminal identification: The gap from Wade to Kirby. *University of Pennsylvania Law Review, 121*, 1079-1133.

Lindholm, T., Christianson, S.-Å., & Karlsson, I. (1997). Police Officers and Civilians as Witnesses: Intergroup Biases and Memory Performance. *Applied Cognitive Psychology, 11*, 431-444.

Lindsay, R. C. L., & Wells, G. L. (1980). What price justice? Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior, 4*, 303-313. doi: 10.1007/bf01040622

Loftus, E., F., Loftus, G., R., & Messo, J. (1987). Some facts about "weapon focus". *Law and Human Behavior, 11*, 55-62.

Loftus, E. F. (2005a). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361-366.

Loftus, E. F. (2005b). Searching for the neurobiology of the misinformation effect. *Learning & Memory, 12*, 1-2.

Maass, A., & Köhnken, G. (1989). Eyewitness identification: Simulating the "weapon effect.". *Law and Human Behavior, 13*, 397-408. doi: 10.1007/bf01056411

MacLin, O. H., Zimmerman, L. A., & Malpass, R. S. (2005). PC_Eyewitness and the Sequential Superiority Effect: Computer-Based Lineup Administration. *Law and Human Behavior, 29*, 303-321.

Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482-489. doi: 10.1037/0021-9010.66.4.482

McCloskey, M., & Egeth, H. E. (1983). Eyewitness identification: What can a psychologist tell a jury? *American Psychologist* 550-563.

Mecklenburg, S. H., Bailey, P. J., & Larson, M. R. (2008). The Illinois field study: A significant contribution to understanding real world eyewitness identification issues. *Law and Human Behavior, 32*, 22-27. doi: 10.1007/s10979-007-9108-6

Mecklenburg, S. H., Larson, M. R., & Bailey, P. J. (2008). Eyewitness Identification: What Chiefs Need to Know Now. *Police Chief, LXXV*.

Megreya, A. M., & Burton, A. M. (2006). Recognizing faces seen alone or with others: when two heads are worse than one. *Applied Cognitive Psychology, 20*, 957-972.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta analytic review. *Psychology, Public Policy, and Law, 7*, 3-35.

Menard, S. (2001). *Applied Logistic Regression* (2nd ed.). Thousand Oaks, CA: Sage Publishers.

Navon, D. (1990). How critical is the accuracy of an eyewitness's memory? Another look at the issue of lineup diagnosticity. *Journal of Applied Psychology, 75*, 506-510. doi: 10.1037/0021-9010.75.5.506

Niemark, E. D., & Estes, W. K. (1967). *Stimulus Sampling Theory*. San Francisco: Holden-Day.

NIJ Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence : a guide for law enforcement*. Washington, D.C.: U.S. Dept. of Justice, Office of Justice Programs.

Office of Technology Assessment. (1983). Scientific Validity of Polygraph Testing: A Research Review and Evaluation. Washington, D. C.: U.S. Congress.

Owens, J. N. (2009). *Is three really a crowd? : the effects of multiple perpetrators on eyewitness identification accuracy and confidence.* Thesis (M A), John Jay College of Criminal Justice, 2009.

Owens, J. N., Rainey, A., & Dysart, J. E. (2008). *Is three really a crowd?: The effects of multiple perpetrators on eyewitness identification accuracy*. Paper presented at the The American Psychology and Law Society, San Antonio, TX.

Pallant, J. (2007). *SPSS Survival Manual* (3rd ed.). Berkshire, UK: McGraw-Hill.

Pampel, F., C. (2007). *Logistic Regression: A Primer*. Thousand Oaks CA: Sage Publishers.

Penrod, S. D., Fulero, S. M., & Cutler, B. L. (1995). Expert psychological testimony on eyewitness reliability before and after Daubert: The state of the law and the science. *Behavioral Sciences & the Law, 13*, 229-259.

Phillips, M. R. (2001). *The National Institute of Justice guide for eyewitness identification evidence: Can it improve juror decision-making?* Ph.D., Florida International University, United States -- Florida. Retrieved from http://proquest.umi.com/pqdweb?did=729111291&Fmt=7&clientId=31967&RQT=309&VName=PQD

Phillips, M. R., Geiselman, R. E., Haghighi, D., & Lin, C. (1997). Some Boundary Conditions for Bystander Misidentification. *Criminal Justice and Behavior, 24*, 370-390.

Phillips, M. R., McAuliff, B. D., Kovera, M. B., & Cutler, B. L. (1999). Double-blind photoarray administration as a safeguard against investigator bias. *Journal of Applied Psychology, 84*, 940-951. doi: 10.1037/0021-9010.84.6.940

Pickel, K. L. (1999). The influence of context on the "weapon focus" effect. *Law and Human Behavior, 23*, 299-311.

Pickel, K. L. (2007). Remembering and identifying menacing perpetrators: Exposure to violence and the weapon focus effect. In R. C. Lindsay, D. F. Ross, J. D. Read & M. P. Toglia (Eds.), *The Handbook of Eyewitness Psychology* (Vol. 2). Mahwaw, NJ: Lawrence Erlbaum Associates, Publishers.

Pickel, K. L. (2009). The weapon focus effect on memory for female versus male perpetrators. *Memory, 17*, 664-678.

Pickel, K. L., Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. (2007). Remembering and identifying menacing perpetrators: Exposure to violence and the weapon focus effect *The handbook of eyewitness psychology, Vol II: Memory for people.* (pp. 339-360): Lawrence Erlbaum Associates Publishers: Mahwah.

Pishkin, V., & Hershiser, D. (1963). Respiration and GSR as functions of white sound in schizophrenia. *Journal of Consulting Psychology, 27*, 330-337. doi: 10.1037/h0047843

Quek , K., Low, W., Razack, A., Loh, C., & Chua, C. (2004). Reliability and validity of the Spielberger State-Trait Anxiety Inventory (STAI) among urological patients: a Malaysian study. *Medical Journal of Malaysia, 59*, 258-267.

Reaves, B. (2004). Census of State and Local Law Enforcement Agencies. Washington, D.C.: Bureau of Justice Statistics.

Revelle, W., & Loftus, E. (1990). Individual differences and arousal: implications for the study of mood and memory. *Cognition and emotion, 4*, 209-237.

Scheck, B. (2008). Response to the Illinois Pilot Program Report. New York: Innocence Project, Benjamin N. Cardozo School of Law, Yeshiva University.

Screbo, M., W. (1999). *Automation Technology and Human Performance: Current Research and Trends*. New York: Lawrence Erlbaum Associates.

Smeets, G., Merckelbach, H., & Griez, E. (1996). Panic Disorder and Right-Hemisphere Reliance. *Anxiety, Stress, and Coping, 10*, 216-230.

Spielberger, C. D. (1983). State-Trait Anxiety Inventory (Form Y). Palo Alto, CA: Mind Garden.

SPSS Inc. (2010). *SPSS User Manual*. Chicago, IL: IBM, Inc.

Stanny, C. J., & Johnson, T. C. (2000). Effects of stress induced by a simulated shooting on recall by police and citizen witnesses *The American Journal of Psychology, 113*, 359-386.

Steblay, N. (1997). Social Influence in Eyewitness Recall: A Meta-Analytic Review of Lineup Instruction Effects. *Law and Human Behavior, 21*, 283-297.

Steblay, N. (2008). Commentary on studying eyewitness investigations in the field : A look forward. [Commentary]. *Law and Human Behavior, 32*.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 27*, 523-540. doi: 10.1023/a:1025438223608

Steblay, N., M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 16*, 413-424.

Tabachnick, B., C., & Fidell, L., S. (2007). *Using Multivariate Statistics* (5th ed.). New York: Pearson Education.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness Evidence:A Guide for Law Enforcement*. (NCJ 178240). Washington, D.C.: National Institute of Justice.

Terry, W. S. (2000). *Learning and memory*. Needham Heights, MA: Allyn & Bacon.

Tickner, A. H., & Poulton, E. C. (1975). Watching for people and actions. *Ergonomics, 18*, 35-51.

Tilton, S. R. (2008). Review of the State-Trait Anxiety Inventory (STAI). *NewsNotes, 48*, 245-255.

Tooley, V., Brigham, J. C., Maass, A., & Bothwell, R. K. (1987). Facial Recognition: Weapon Effect and Attentional Focus. *Journal of Applied Social Psychology, 17*, 845 - 859.

Tredoux, C. (1998). Statistical Inference on Measures of Lineup Fairness. *Law and Human Behavior, 22*.

Trochim, W., M.K. (2001). *The Research Methods Knowledge Base* (2nd. ed.). Cincinnati, OH: Atomic Dog Publishing Co. .

Wall, P. M. (1965). *Eyewitness Identification in Criminal Cases*. Springfield, Ill: Charles C.Thomas.

Wells, E. C., & Pozzulo, J. D. (2006). Accuracy of eyewitnesses with a two-culprit crime: Testing a new identification procedure. *Psychology, Crime & Law, 12*, 417-427.

Wells, G. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology, 36*, 1546-1557. doi: 10.1037/0022-3514.36.12.1546

Wells, G. (1993). What do we know about eyewitness identification? *American Psychologist, 48*, 553-571. doi: 10.1037/0003-066x.48.5.553

Wells, G. (2001). Police lineups: Data, theory, and policy. *Psychology, Public Policy, and Law, 7*, 791-801. doi: 10.1037/1076-8971.7.4.791

Wells, G. (2005). Helping experimental psychology affect legal policy. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: an empirical perspective*. New York: Guilford Press.

Wells, G. (2008). Field experiments on eyewitness identification: Towards a better understanding of pitfalls and prospects. [Commentary]. *Law and Human Behavior, 32*, 6-10. doi: 10.1007/s10979-007-9098-4

Wells, G., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88*, 776-784. doi: 10.1037/0033-2909.88.3.776

Wells, G., Luus, C. A. E., & Windschitl, P. D. (1994). Maximizing the utility of eyewitness identification evidence. *Current Directions in Psychological Science, 3*, 194-197. doi: 10.1111/1467-8721.ep10770833

Wells, G., & Luus, C. E. (1990). The diagnosticity of a lineup should not be confused with the diagnostic value of nonlineup evidence. *Journal of Applied Psychology, 75*, 511-516. doi: 10.1037/0021-9010.75.5.511

Wells, G., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J., & Fulero, S. (2000). From the Lab to the Police Station: A Successful Application of Eyewitness Research. *American Psychologist, 55*, 581-598.

Wells, G., Memon, A., & Penrod, S. D. (2006). Eyewitness Evidence: Improving Its Probative Value. *Psychological Science in the Public Interest, 7*, 45-75.

Wells, G., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*.

Wells, G., Seelau, E. P., Rydell, S. M., Luus, C. A. E., Ross, D. F., Read, J. D., et al. (1994). Recommendations for properly conducted lineup identification tasks *Adult eyewitness testimony: Current trends and developments.* (pp. 223-244): Cambridge University Press: New York.

Wells, G., Small, M., Penrod, S., Malpass, R. S., Fulero, S., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 1-39.

Wells, G., & Turtle, J. (1988). What is the best way to encode faces? In M. Gruenberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (pp. 163-168). New York: Wiely.

Wells, G., & Windschitl, P. D. (1999). Stimulus Sampling and Social Psychological Experimentation. *Personality and Social Psychology Bulletin, 25*, 1115-1125.

Winzeler, Z., L. (2008). Whoa, whoa, whoa . . . One at a time: examining the responses to the Illinois study on double-blind sequential lineup procedures. *Utah Law Review, 2008*, 1595-1612.

Yarmey, A. D. (1986). Perceived Expertness and Credibility of Police Officers as Eyewitnesses *Canadian Police College Journal, 10*, 31-52.

Yarmey, A. D. (1998). Police as Witnesses [Review]. *Expert Evidence, 6*, 237-239.

Yarmey, A. D. (2004). Eyewitness recall and photo identification: a field experiment. *Psychology, Crime & Law, 10*, 53-68. doi: 10.1080/10683160021000058379

Yarmey, A. D., & Jones, H. P. (1983). Accuracy of memory of male and female eyewitnesses to a criminal assault and rape. *Bulletin of the Psychonomic Society, 21*, 89-92.

Yuille, J. C. (1984). Research and teaching with police: A Canadian example. . *International Review of Applied Psychology, 33*, 5-23.

Yuille, J. C. (1993). We must study forensic eyewitnesses to know about them. *American Psychologist, 48*, 572-573.