The Effects of Cognitive Bias, Examiner Experience, and Stimulus Material on Forensic Evidence Analysis

Michelle M. Pena
Florida International University, mpena034@fiu.edu

Follow this and additional works at: https://digitalcommons.fiu.edu/etd

Part of the Cognition and Perception Commons, Cognitive Psychology Commons, Other Psychology Commons, and the Social Psychology Commons

Recommended Citation
https://digitalcommons.fiu.edu/etd/4246

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.
FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

THE EFFECTS OF COGNITIVE BIAS, EXAMINER EXPERIENCE, AND
STIMULUS MATERIAL ON FORENSIC EVIDENCE ANALYSIS

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Michelle M. Pena

2019
To: Dean Michael R. Heithaus  
College of Arts, Sciences and Education  

This dissertation, written by Michelle M. Pena, and entitled The Effects of Cognitive Bias, Examiner Experience, and Stimulus Material on Forensic Evidence Analysis, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_______________________________________  
Jacqueline R. Evans  

_______________________________________  
Stephen Charman  

_______________________________________  
Rob T. Guerette  

_______________________________________  
Nadja Schreiber Compo, Major Professor  

Date of Defense: June 24, 2019  

The dissertation of Michelle M. Pena is approved.

_______________________________________  
Dean Michael R. Heithaus  
College of Arts, Sciences and Education  

_______________________________________  
Andrés G. Gil  
Vice President for Research and Economic Development  
and Dean of the University Graduate School  

Florida International University, 2019
DEDICATION

This dissertation is dedicated to my parents, Lourdes and Carlos and my brothers and sisters, Lourdes, Jorge, Chris, and Lupe. Gracias por su amor incondicional y apollo sin fin durante todo este proceso. Thank you for your unconditional love and never-ending support throughout this entire process.
ACKNOWLEDGMENTS

First and foremost, I would like to thank my mentor, Dr. Nadja Schreiber Compo, whose advice and guidance led me to find my passion in research. Her faith in me during my undergraduate and graduate career helped me believe in myself and become the researcher I am today.

I would also like to thank my committee members for all of their advice and support. Their constructive criticisms and feedback helped me develop a project I am proud of. Their words of support were also very much appreciated during the course of this project.

Finally, I would like to thank Stephanie Stoiloff, commander of the Miami-Dade Forensic Services Bureau, for her patience and understanding throughout all stages of this project. Her constant feedback and willingness to help with this project allowed me to truly understand the importance of collaborating with practitioners in the field.
ABSTRACT OF THE DISSERTATION
THE EFFECTS OF COGNITIVE BIAS, EXAMINER EXPERIENCE, AND STIMULUS MATERIAL ON FORENSIC EVIDENCE ANALYSIS

by

Michelle M. Pena

Florida International University, 2019
Miami, Florida

Professor Nadja Schreiber Compo, Major Professor

Forensic examiners have come under scrutiny in recent years because of high profile exoneration cases that have highlighted the negative impact contextual bias can have on investigations including forensic evidence analyses. This has led to several proposed solutions to reduce the effects of bias including blind testing and redacting task-irrelevant information. However, practitioners have not been receptive to such recommendations because of the limitations found in past research, such as the use of untrained undergraduate students to examine complex pieces of forensic evidence (e.g., fingerprints). The current study thus had the following aims: (a) examine the effect of contextual bias on examiners’ evaluation of forensic evidence by varying the amount of pre-comparison information available to participants, (b) compare novice and expert examiners’ performance when conducting forensic analyses and their vulnerabilities to contextual bias, and (c) investigate the effect of contextual bias on examiners’ evaluation of different types of forensic evidence. Expert forensic examiners and novice undergraduate students were recruited and provided with a lab analysis request form that either contained all case summary details about a mock crime, no case summary details
about the crime (i.e., blind testing) or had task-irrelevant case summary details redacted. Participants were asked to compare matching and non-matching fingerprint and footwear impression evidence. Results suggest no effect of blinding examiners from case information or redacting task-irrelevant information on examiners’ decisions. Findings also suggest that both examiner experience and the type of forensic evidence analyzed can have a significant effect on examiners’ judgments. Expert examiners were significantly more accurate than novices. However, expert examiners were only significantly more accurate and better able to discriminate between matching and non-matching pairs of evidence when analyzing fingerprint evidence and not footwear impression evidence. These findings suggest caution when using forensic stimuli with novice samples to investigate cognitive bias in forensic examination. Finally, the present study suggest that the proposed blinding and redaction procedures require additional research including expert examiners and a spectrum of forensic stimulus material to yield measurable effects of cognitive bias.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Cognitive Bias</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Cognitive Bias in Legal Contexts</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Cognitive Bias in Forensic Contexts</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Dangers of Cognitive Bias in Forensic Sciences</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Proposed Solutions</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Practitioners’ Concerns</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Present Study</td>
<td>.................................................................</td>
</tr>
<tr>
<td>II.</td>
<td></td>
</tr>
<tr>
<td>PILOT STUDY ONE METHODS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Participants</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Materials</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Procedure</td>
<td>.................................................................</td>
</tr>
<tr>
<td>III.</td>
<td></td>
</tr>
<tr>
<td>PILOT STUDY ONE RESULTS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Match Fingerprint Judgments</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Non-Match Fingerprint Judgments</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Selection of Match and Non-Match Fingerprint Pairs for Pilot Study Two</td>
<td>.................................................................</td>
</tr>
<tr>
<td>IV.</td>
<td></td>
</tr>
<tr>
<td>PILOT STUDY TWO METHODS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Design</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Participants</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Materials</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Procedure</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Coding of Outcome Variables</td>
<td>.................................................................</td>
</tr>
<tr>
<td>V.</td>
<td></td>
</tr>
<tr>
<td>PILOT STUDY TWO RESULTS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Match Judgments</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Non-Match Judgments</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Inconclusive Judgments</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Overall Accuracy Rate</td>
<td>.................................................................</td>
</tr>
<tr>
<td>VI.</td>
<td></td>
</tr>
<tr>
<td>MAIN STUDY METHODS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Design</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Participants</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Materials</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Procedure</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Coding of Outcome Variables</td>
<td>.................................................................</td>
</tr>
<tr>
<td>VII.</td>
<td></td>
</tr>
<tr>
<td>MAIN STUDY RESULTS</td>
<td>.................................................................</td>
</tr>
<tr>
<td>Time 1 Judgments</td>
<td>.................................................................</td>
</tr>
</tbody>
</table>
Signal Detection Measures .................................................................67
Confidence Ratings ........................................................................69
Time 2 Judgments .........................................................................71
Perceptions Questionnaire .............................................................72

VIII. DISCUSSION .............................................................................74
Contextual Information ..................................................................74
Examiner Experience ......................................................................78
Stimulus Material ..........................................................................81
Testing Linear Sequential Unmasking Procedures .........................82
Limitations and Future Directions ..................................................83
Implications and Conclusions .........................................................85

TABLES ..............................................................................................88

FIGURES ..........................................................................................104

LIST OF REFERENCES .....................................................................110

APPENDICES ....................................................................................117

VITA ..................................................................................................132
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables and Likelihood of Making a Correct Fingerprint ID</td>
<td>88</td>
</tr>
<tr>
<td>2. Time 1 Judgments by Examiner Experience</td>
<td>89</td>
</tr>
<tr>
<td>3. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Footwear Impression ID</td>
<td>90</td>
</tr>
<tr>
<td>4. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making an Incorrect Fingerprint Exclusion</td>
<td>91</td>
</tr>
<tr>
<td>5. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making an Incorrect Footwear Impression Exclusion</td>
<td>92</td>
</tr>
<tr>
<td>6. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Fingerprint Exclusion</td>
<td>93</td>
</tr>
<tr>
<td>7. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Footwear Impression Exclusion</td>
<td>94</td>
</tr>
<tr>
<td>8. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a False Fingerprint ID</td>
<td>95</td>
</tr>
<tr>
<td>9. Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a False Footwear Impression ID</td>
<td>96</td>
</tr>
<tr>
<td>10. Mixed ANOVA Analysis Examining Effects of Predictor Variables on Total Number of Inconclusive Judgments Made</td>
<td>97</td>
</tr>
<tr>
<td>11. Mixed ANOVA Analysis Examining Effects of Predictor Variables on Accuracy Rate</td>
<td>98</td>
</tr>
<tr>
<td>12. Mixed ANOVA Analysis Examining Effects of Predictor Variables on Discrimination Accuracy</td>
<td>99</td>
</tr>
<tr>
<td>13. Mixed ANOVA Analysis Examining Effects of Predictor Variables on Response Bias</td>
<td>100</td>
</tr>
</tbody>
</table>
14. Mixed ANOVA Analysis Examining Effects of Predictor Variables on Confidence Ratings ............................................................. 101

15. Confidence and Accuracy Bivariate Correlation by LAR and Experience Condition ................................................................. 102

16. Ordinal Logistic Regression Analyses Examining Effects of Contextual Information and Examiner Experience on Likelihood of Reviewing Time 1 Judgments .................................................................................. 103
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student Match Fingerprint Judgements</td>
<td>104</td>
</tr>
<tr>
<td>2. Student Non-Match Fingerprint Judgements</td>
<td>105</td>
</tr>
<tr>
<td>3. Average Number of Inconclusive Judgments by Examiner Experience and Stimulus Material</td>
<td>106</td>
</tr>
<tr>
<td>4. Average Accuracy Rate by Examiner Experience and Evidence Material</td>
<td>107</td>
</tr>
<tr>
<td>5. Average Discrimination Accuracy by Examiner Experience and Evidence Material</td>
<td>108</td>
</tr>
<tr>
<td>6. Average Confidence Ratings by Examiner Experience and Evidence Material</td>
<td>109</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Forensic sciences, including DNA and fingerprint analyses, were once believed to be error-free and consist of objective and purely scientific procedures, thus being regularly admitted as evidence at trial (Mnookin et al., 2011). Unfortunately, various high-profile cases, such as that of Brandon Mayfield (i.e., Madrid Bomber case), have brought forensic sciences under scrutiny in recent years as they highlight the subjective nature of such analyses and the subsequent consequences of human error. Along with these high-profile cases, analyses of hundreds of cases by the Innocence Project suggest that a large number of DNA exoneration are due, at least in part, to improper analysis or use of forensic evidence, casting doubt on the reliability of forensic sciences (Innocence Project, 2017).

In recent years, researchers have begun to explore the potential factors that may lead forensic examiners to make such consequential errors. Specifically, a significant amount of research has examined how cognitive biases can affect the analysis of forensic evidence including DNA samples, fingerprints, handwriting samples, and footwear impressions (Dror & Hampikian, 2011; Dror, Peron, Hind, & Charlton, 2005; Kerstholt, Paashuis & Sjerps, 2007; Kukucka & Kassin, 2014). Much of this research suggests that having contextual information indicating that two forensic evidence samples came from the same person (e.g., suspect confessed to the crime) can increase the likelihood that an examiner will arrive at the same conclusion (e.g., fingerprint found at the crime scene matches suspect’s fingerprint), regardless of whether or not the samples did originate from the same person (Dror, Charlton, & Peron, 2006; Kukucka & Kassin, 2014).
Together, wrongful conviction cases and research on the effect of cognitive bias on forensic evidence analyses have led to the development of various solutions aimed at reducing these negative effects of bias. Some of these recommendations include implementing blind testing, using evidence lineups, and hiring case managers to redact any task-irrelevant information from lab analysis request forms (LARs) – a method often suggested as part of the linear sequential unmasking procedure (Edmond et al., 2016; Kassin, Dror, & Kukucka, 2013; Krane et al., 2008; Mattijssen, Kerkoff, Berger, Dror & Stoel, 2016; Saks, Risinger, Rosenthal, & Thompson, 2003). However, some of the recommendations put forth by researchers, such as using the linear sequential unmasking procedure, which are believed to reduce bias, have not been experimentally tested in forensic contexts.

In addition, practitioners are often hesitant to incorporate recommendations that stem from research testing inexperienced examiners (often undergraduate students) who have no training in analyzing complex forensic evidence such as fingerprints. The present study thus aims to address the gap in the literature and the concerns voiced by forensic examiners (e.g., Butt, 2013) by directly comparing the use of blind versus redaction procedures on the analysis of different types of forensic evidence (i.e., fingerprints and footwear impressions) and the role examiner experience plays in explaining the relationship between cognitive bias and examiners’ judgments.

**Cognitive Bias**

Social and cognitive scientists often use the term cognitive bias to describe a variety of different biases in human cognition and behavior. For instance, confirmation bias, contextual bias, expectancy effects, observer effects, tunnel vision, and belief
perseverance are often used interchangeably within the literature. Although the exact definitions for each of these terms vary to some degree, they all describe a variation of cognitive bias. In its broadest sense, cognitive bias describes how individuals’ prior beliefs and preconceptions can bias their subsequent interpretation of information in the direction of those expectations (Kassin et al., 2013; Nickerson, 1998; Rosenthal, 1994; Rosenthal & Fode, 1963). Cognitive biases are processes that occur automatically without the individuals’ awareness (Kunda, 1990). As such, individuals are not purposefully seeking out information or behaving in ways that support their prior beliefs nor are they deliberately choosing to ignore information that does not fit with their hypothesis (Ask & Granhag, 2005; Nickerson, 1998).

Although cognitive bias occurs without conscious awareness, years of research have consistently shown that individuals’ expectations can produce dramatic behavioral consequences. For example, in one of the first studies examining bias, conducted by Rosenthal and Rubin (1978), grade-school teachers were arbitrarily told that some of their students had scored above average on an intelligence test and were thus expected to show a significant growth intellectually. Results demonstrated that students who were expected to produce this greater intellectual growth did in fact do so compared to those students who were not designated “high-scoring.” Similarly, Snyder and Swann (1978) examined how individuals’ prior beliefs about another individual could affect how they behaved and interacted with the other individual. The researchers led participants to believe that they would be interviewing a person who was either an extrovert or an introvert. Results revealed that interviewers selected questions and elicited responses that
were consistent with their expectations regarding the interviewee’s personality. The pattern of results was true regardless of whether interviewers’ expectations were correct.

Cognitive bias has not only been shown to affect individuals’ overt behaviors, but also their perceptions of presented stimuli. For example, one of the classic studies conducted by Brunner and Potter (1964) shows how creating an initial hypothesis regarding ambiguous material affects how one perceives the same material as it becomes less ambiguous. Specifically, they presented participants with images of objects that were initially out of focus and were then shown images of these same objects as they were slowly brought into focus. They found that individuals’ initial beliefs of what they believed the object was in the first blurred image interfered with its recognition later on, when the same image became less blurry.

In another study conducted by Bressan and Dan Martello (2002), participants were asked to rate the similarity between two faces and were told either a truthful or deceitful statement regarding the degree to which the two faces were related. Participants who had been told that the two faces were related rated the faces as more similar than participants who had been led to believe that the faces were unrelated regardless of whether the faces were truly related or not. In other words, what participants were told and thus their prior beliefs about the relatedness of the faces was a stronger predictor of how similar they rated the faces than the actual relatedness of the faces. Both of these studies demonstrate how cognitive biases, which are often considered to be mental shortcuts that ultimately simplify human decision making, can lead to errors in how one perceives subsequent stimuli and thus decisions made regarding such stimuli.
Cognitive Bias in Legal Contexts

Research on cognitive bias has expanded beyond basic research lab paradigms into applied legal contexts, with similar results (Cantlon, Payne, & Erbaugh, 1996; Hasel & Kassin, 2009; Kassin, Goldstein, & Savitsky, 2003). Specifically, researchers have examined the effects of cognitive bias in witness and suspect interviewing contexts as well as within lineup procedures and eyewitness identifications. Regardless of the legal context in which cognitive bias has been examined, researchers continue to find support for its effect on human behavior and decision-making. For example, Kassin, Goldstein and Savitsky (2003), investigated how manipulating interrogators’ a priori expectations of a suspect’s guilt affected the types of questions interrogators asked and the different tactics used when interviewing the suspect. Specifically, interrogators were led to believe that the person they would be interviewing was either likely to be guilty or likely to be innocent and were informed that either 80% of the suspects in the study committed the crime (guilty expectation condition) or that only 20% committed the crime (innocent expectation condition). Results suggested that pre-interview expectations of guilt led to changes in an interrogator’s behavior such that those in the guilty expectation condition used more guilt-presumptive questions and interrogation tactics (e.g., promises of leniency, presentation of false evidence) than participants in the innocent expectation condition.

Other studies examining the effect of pre-interview information have also found an effect on witness recall and victim disclosures (Cantlon et al., 1996; Rivard, Pena, & Schreiber Compo, 2015). For example, Rivard and colleagues (2015) conducted a study examining the effect of cognitive bias in adult witness interviewing contexts. Student-
interviewers were provided either with correct, incorrect, or no information (i.e., blind interviewers) about a crime prior to interviewing student-witnesses who had watched said mock-crime. They found that witnesses recalled more correct details when interviewed by blind interviewers than by correctly informed interviewers. Taken together, findings across different legal interviewing contexts suggest that having pre-interview information may yield behavior that is detrimental to interviewers’ questioning strategies and the quality and quantity of information elicited from witnesses.

Within the eyewitness literature, there is also evidence to suggest that having information regarding the suspect’s guilt can affect eyewitness identification decisions. Hasel and Kassin (2009) exposed participants to a staged theft and a week later asked them to make an identification from a lineup that did not include the perpetrator. Participants were then told that a certain lineup member had confessed to the crime before being provided with the opportunity to change their original identification decision. They found that more than half of the participants who had selected a lineup member, but were later informed that it had been a different lineup member who confessed to the crime, changed their initial identification. In fact, 50% of those participants who changed their initial identification, selected the lineup member who allegedly confessed. Here too, research suggests that having prior beliefs about an individuals’ guilt can have direct consequences on a witness’ decision-making process as well as their perceptions of information presented to them.

Relatedly, other studies examining the effects of cognitive bias in eyewitness identification contexts have found that having information regarding a suspect’s location in the lineup can affect lineup administrators’ verbal and nonverbal behaviors.
That is, administrators who are aware of the suspect’s location in a lineup may inadvertently leak cues of the specific location to the witness, which in turn can influence eyewitness decisions (Greathouse & Kovera, 2009; Haw & Fisher, 2004). It has therefore been recommended that lineups be administered in a double-blind fashion, with the lineup administrator “blind” to who the suspect is (IACP, 2010). Overall, findings across various legal contexts mirror those found in initial studies examining the effects of cognitive bias, such that individuals’ expectations can affect how they behave and interpret subsequent information.

**Cognitive Bias in Forensic Contexts**

More recently research on cognitive bias has been extended to the domain of forensic sciences. As a result of high-profile cases highlighting the negative consequences of cognitive bias on forensic examiners’ decision-making, research has more directly examined the intricacies of how and when cognitive bias can affect the analysis of forensic evidence. Similar to past research on the effects of cognitive bias in other legal contexts (e.g., suspect interviewing, eyewitness identifications), research in the forensic domain suggests that individuals assessing forensic evidence are not immune to the effects of cognitive bias (Dror & Cole, 2010; Dror & Rosenthal, 2008). Kassin, Dror, and Kukucka (2013) coined the term *forensic confirmation bias* to describe “the class of effect through which an individual’s preexisting beliefs, expectations, motives, and situation context influence the collection, perception, and interpretation of evidence during the course of a criminal case” (p. 45). Similarly, the term contextual bias is often used in research examining the effects of bias in forensic sciences to refer to the specific effect of having extraneous information about the case (e.g., suspect’s ethnicity,
additional incriminating evidence) on examiners’ evaluations of forensic evidence (Edmond, Tangen, Searston, & Dror, 2015; Lange, Thomas, Dana, & Dawes, 2011).

Research has examined the effect of contextual bias across a number of forensic domains including fingerprints, handwriting samples, shoeprints, bullets, auditory samples, and polygraph results (Dror et al., 2005; Elaad, Ginton, & Shakhar, 1994; Kerstholt, Paashuis, & Sjerps, 2007; Kerstholt et al., 2010; Kukucka & Kassin, 2014; Lange et al., 2011). Even DNA testing, which is often considered to be the “gold-standard” of forensic evidence, has been demonstrated to fall prey to cognitive bias (Dror & Hampikian, 2011; Lynch, 2003), as all forensic domains require, to some extent, an individual’s subjective opinion to determine whether two pieces of evidence are similar enough to be considered a match regardless of the type of stimulus material.

Taken together, findings across domains suggest that individuals’ decisions regarding the extent to which two stimuli/pieces of evidence derive from the same source (i.e., are matches) can be influenced by outside information about the crime, the suspect, or other available evidence. Research has also begun to expand beyond simply examining what type of contextual information can bias examiners’ analyses by investigating what specific moderating factors, if any, can influence the strength of the cognitive bias effect. For example, Dror and colleagues (2005) conducted a study examining the effects of emotional contextual case information and stimuli ambiguity on examiners’ interpretation of fingerprint evidence. Students were presented with numerous pairs of fingerprints that varied in level of ambiguity (i.e., ambiguous versus unambiguous). Before deciding whether each pair of prints were matches or non-matches, participants were presented with background information and photographs of the crime (i.e., contextual information).
The background information for each pair of prints varied in the emotional state it invoked; participants were provided with less serious, low-in-emotion background information about the crime (e.g., a bicycle theft) or with serious, high-in-emotion background information (e.g., a murder). Participants were also provided with crime scene pictures consistent with their condition; the low emotional state participants viewed pictures of the stolen items whereas those in the high emotional state condition were presented with pictures of victims who had been badly hurt or killed. They found that participants in the high emotional state condition were more likely to state that two fingerprints were a match (regardless of whether or not they were indeed a match) compared to those in the low emotional state. However, the effect was only found for participants who were given ambiguous fingerprints and not for those who were given clear fingerprints. This suggests that although cognitive bias may be present, other variables such as evidence ambiguity, can influence the relationship between contextual bias and examiner accuracy.

Similarly, Kukucka and Kassin (2014) conducted a study to examine whether the similarity between two pieces of evidence could influence the degree to which having contextual information about a suspect could affect examiners’ decisions about the evidence. Participants were presented with a case in which an individual had robbed a bank and had been brought in for questioning. They then provided participants with evidence consisting of a handwritten note the perpetrator had given the bank teller demanding the money and a Miranda waiver the suspect in custody had signed. Before participants were asked to compare the two handwriting samples and decide whether they were a match or a non-match, participants were presented with information regarding the
innocence of the suspect. While some participants were told that the suspect had confessed during a prior interview, other participants were told that the suspect had maintained his innocence throughout the entire interview. The similarity between the two handwriting samples was also manipulated; samples were either categorized as very similar or not similar to one another. The effect of bias was hypothesized to have a stronger impact when participants were presented with a pair of highly similar non-matching handwriting samples (a more difficult comparison) compared to less similar non-matching handwriting samples. Although results suggested an effect of cognitive bias, such that participants who had been told the suspect confessed were more likely to incorrectly judge the non-matching handwriting samples as matching, handwriting similarity did not moderate the effect cognitive bias had on participants’ analyses of the handwriting samples. The lack of the moderating effect suggests that regardless of whether it was a “clear” non-match decision (i.e., handwriting samples were less similar to one another) or not, participants were just as likely to erroneously state the samples were a match, which differs from Dror and colleagues’ (2005) findings where stimulus ambiguity did have a moderating effect on examiners’ judgments.

It is also important to note that while some studies find an effect of cognitive bias on forensic analyses, others fail to demonstrate the effect at all (Dror & Cole, 2010; Kerstholt et al., 2007; Kerstholt et al., 2010). For example, Kerstholt and colleagues (2007) examined the effects of expectation and evidence complexity on trained police officers’ footwear impression examinations. Participants were presented with either neutral or guilt-presumptive information regarding the background of the case before being presented with four footwear impressions, two of which were simple and two of
which were considered complex. Kerstholt and colleagues (2007) found no main effect of expectation such that the guilt-presumptive information about the suspect given to participants did not alter how much evidentiary value was assigned to each footwear impression. These findings are in direct contrast to prior research suggesting that case details can affect individuals’ judgment and perception of evidence (Dror et al., 2005).

Differences in results across studies may be caused by differences in outcome measures, cognitive bias manipulations, study procedures, and participant samples used. For example, while some studies may provide participants with both matching and non-matching pieces of evidence, others may simply present non-matching prints or matching-prints. The problem lies in that one study may only measure the likelihood of participants making incorrect “match” decisions, while another study may only measure the likelihood of making incorrect “non-match” decisions; two different decisions, which arguably use different decision-making criterion. Similarly, some studies may ask participants to make a match/non-match judgment while others may simply ask them to rate how valuable the piece of evidence is. Again, differences in outcome measures make it difficult to pinpoint under what circumstances cognitive bias may be particularly harmful to individuals’ analyses of forensic evidence. In addition, a common limitation found in past research examining the effect of cognitive bias on forensic analyses is the lack of information provided regarding the methods used to categorize their stimuli as similar/not similar, ambiguous/unambiguous, or complex/non-complex. Without such information on how and if evidence materials were piloted, it is almost impossible to compare the effects of such moderating factors, likely further accounting for differences across studies.
Dangers of Cognitive Bias in Forensic Sciences

Despite these inconsistencies and a lack of empirical research disentangling possible reasons for them, researchers generally agree that the potential dangers of cognitive bias in forensic sciences need to be addressed, and have proposed reforms (Kassin, Dror, & Kukucka, 2013; National Academy of Sciences, 2009). One of the first calls for reform was made by the National Academy of Science, which published a report critiquing various forensic disciplines (e.g., firearms, handwriting, fingerprints) for their lack of standardization and the risk of contextual bias affecting analyses (National Academy of Sciences, 2009). The report strongly urged the discipline of forensic science to work together with cognitive psychologists to address the issues related to cognitive bias in the forensic science disciplines.

Many researchers have also argued that the dangers of cognitive bias go beyond simply affecting forensic examiners in the lab and have the potential to spill over to other pieces of evidence and other stages of the investigation, thus creating a bias snowball effect (Charman, Kavetski, & Mueller, 2017; Dror, Morgan, Rando, & Nakhaeizadeh, 2017). The bias snowball effect explains how the effect of bias can increase as additional pieces of irrelevant information from different sources are integrated, influencing one another. For example, a forensic examiner may be influenced by information regarding a suspect’s confession, which in turn affects how the examiner perceives a pair of fingerprints, which can then subsequently affect an eyewitness’ identification if they are made aware of the confession information. Thus, these “independent” pieces of evidence are no longer independent of one another and instead lead to “corroboration inflation” (Kassin, 2012), where an illusion of false support is produced.
Furthermore, a theoretical model has been proposed to better explain how cognitive bias affects the integration of evidence and how it affects the investigation of a crime as a whole, making it especially problematic (Charman, 2013). The cognitive coherence approach argues that research must move beyond simply understanding how one piece of evidence (e.g., confession) influences the evaluation of another piece of evidence (e.g., eyewitness identification), and highlights how the evaluation of evidence affects greater evolving conclusions that are formed by the investigative team (i.e., whether the suspect is guilty or not guilty). It also emphasizes the bidirectionality of these effects such that the evaluation of one piece of evidence affects the emerging conclusion, and this conclusion in turn can affect the evaluation of other evidence (Charman, 2013; Dror et al., 2017).

Overall, a significant amount of research has focused on factors that can bias forensic examiners such as knowing that the suspect confessed (Kukucka & Kassin, 2014), that the evidence is part of a high-profile case (Schiffer & Champod, 2007) or knowing the conclusions of a previous examiner (Dror, Charlton, & Peron, 2006). Regardless of the context, findings have often led researchers to warn practitioners against the potential dangers that stem from the cross-contamination of evidentiary information when examiners are informed of other, forensic or non-forensic pieces of information.

**Proposed Solutions**

Previous work examining the possible factors that can make examiners particularly susceptible to bias has led researchers to propose and advocate for procedural changes in forensic laboratories that minimize the effects of cognitive bias (see Kassin,
Dror, Kukucka, 2013). For example, researchers have proposed using the following, (1) blind testing, (2) evidence lineups, and (3) contextual information management procedures (Cole, 2013; Edmond et al., 2016; Kassin et al., 2013; Mattijssen et al., 2016; Risinger, Saks, Thompson, & Rosenthal, 2002; Saks et al., 2003; Thompson, 2011).

**Blind testing.** Blind testing is one of the most recommended and used methods in both the biomedical field (Kaptchuk, 1998) and in psychological research (Rosenthal, 1966), rendering both scientists and researchers unaware of the conditions participants have been assigned to and of study hypotheses, which both have the potential to provide biasing information. Similarly, the President’s Council of Advisors on Science and Technology (2016) has noted the importance of blinding forensic examiners from extraneous and potentially biasing information. In addition to the recent body of work on cognitive bias in forensic sciences, these recommendations are largely influenced by the years of research examining the effect of cognitive bias on individuals’ decision-making (Nickerson, 1998; Rosenthal & Fode, 1963).

Unfortunately, there is a large discrepancy between how strongly researchers advocate for the use of blind testing in forensic sciences and the amount of research that has actually tested this method in forensically-relevant domains. That is, only a few studies have directly examined how keeping examiners blind to extraneous information about a case influences their decisions. The majority of studies manipulate participants’ expectations regarding a piece of evidence by presenting some participants with incriminating evidence (e.g., suspect confessed) or potentially biasing case information (e.g., high profile case) and other participants with “neutral” information (e.g., suspect maintained his innocence throughout entire interview, low-profile case; Kukucka &
Kassin, 2014; Schiffer & Champod, 2007). However, presenting participants with “neutral” information is not a true test of blind testing, which would require participants to receive no information (or at least no extraneous information) about the case.

Those few studies that have directly tested a bona fide blind testing manipulation (i.e., examiners were not provided with any information about the case and were solely given pieces of evidence to analyze) have yielded mixed results regarding a possible advantage of using blind testing (Lange et al., 2011; Langenburg, Champod, & Wertheim, 2009). For example, Langenburg and colleagues (2009) assigned participants who were either experts or trained students who had taken a forensic fingerprint analyses training course to either a high bias, low bias, or no context condition. Participants in the no context condition were not provided with any information about the case and were simply asked to make a judgment regarding the fingerprints. Participants in the high and low bias group were provided with information regarding a prior examiner’s decision. Participants in the no context condition made a similar number of errors when judging the fingerprints compared to those in the high and low bias groups. However, in a different study conducted by Lange and colleagues (2011), a different finding emerged. Participants listened to and transcribed degraded pieces of audio and were either told that the audio they would be listening to was from a suspect interview or were not given any information (i.e., blind condition). They found that participants who listened to significantly degraded audio clips and were told it was a suspect interview were 7.23 times more likely to misinterpret the audio than those in the blind condition suggesting that blind testing can be used to increase accuracy.
Although past research on the effects of blind testing in other areas support the notion for blind testing in forensic contexts and the intuitively sensible notion that providing forensic examiners with no information about the crime will prevent bias or at least the notion of bias, research has yet to provide results on this issue that are directly relevant to forensic sciences in general and forensic examiners in particular. The present study thus aims to address the gap in the literature and provide a direct examination of blind testing procedures within a forensically-relevant context.

**Evidence lineups.** Another novel and innovative solution that has been proposed by researchers to reduce the effect of cognitive bias on forensic examinations is the use of a filler-control method, also known as evidence lineups (Kassin et al., 2013; Saks et al., 2003; Wells, Wilford, & Smalarz, 2013). Evidence lineups are likened to photo lineup identification procedures in principle as they present examiners with at least three samples of evidence – the crime scene sample, the suspect sample, and an innocent filler sample(s). The suspect sample and the filler sample are presented so that the examiner is unaware of which sample belongs to the suspect and which one belongs to the filler (Saks et al., 2003). Thus, it is similar to an eyewitness lineup such that the suspect sample is present among “foils.” Current examination of evidence in forensic labs are more similar to eyewitness “show-ups,” instead of ‘lineups,” which imply that the correct suspect sample is present (i.e., a match). Evidence lineups on the other hand are aimed at decreasing cognitive bias such that even if examiners are exposed to contextual information (e.g., suspect confessed), they still do not know which sample came from the suspect and which came from the fillers (Wells et al., 2013).
Research examining evidence lineup procedures suggest that it can increase the accuracy of examiners’ judgments (Miller, 1987; Quigley-Mcbride, 2017). For example, in a study conducted by Miller (1987), student participants who were enrolled in a crime laboratory college course were asked to compare hair samples for four different criminal investigations. All participants were given a hair sample that had been recovered from each of the four different crime scenes and were given background case information. However, for two of the four cases, participants were given only one other suspect hair sample to compare to. For the remaining two cases, participants were asked to make the comparison using an evidence lineup where they were presented with five suspect-known hair samples for each case. Examiners made significantly more incorrect judgments when using the standard procedure (comparing evidence found at the crime scene to only one suspect sample) compared to when using the evidence lineup. Although this initial study highlights the ability of evidence lineups to increase examiners’ accuracy, it did not manipulate the presence of contextual bias.

Quigley-McBride (2017), expanded upon the study conducted by Miller (1987) by examining whether using an evidence lineup could successfully eliminate the effect of cognitive bias. Undergraduate student participants were assigned to assess a set of fingerprints using either a standard procedure (comparing one crime scene fingerprint with one suspect fingerprint) or an evidence lineup procedure (comparing one crime scene fingerprint with six suspect fingerprints). Some participants received a police report containing additional evidence that strongly incriminated the suspect (i.e., DNA found under witness’ fingernails matched suspect’s DNA) while other participants did not receive additional incriminating evidence. The use of an evidence lineup procedure was
found to successfully decrease the effect of biasing information. Albeit promising, to date the Quigley-McBride (2017) study is the only empirical test of whether evidence lineups may be useful in protecting forensic examiners from the biasing effects of having contextual case information. As the Quigley-McBride (2017) study included only novice, inexperienced examiners, it is of crucial importance that future research examining the effects of evidence lineups include expert examiners and other stimuli in addition to hair samples and fingerprints. Furthermore, filler selection criteria and processes need to be transparent and clearly delineated to ensure that forensic experts agree with the methodology of these studies. This transparency will not only allow us to systematically investigate the importance of training/expertise and type of stimuli/evidence, but also contribute more directly to the translation of such research for forensic experts.

**Contextual information management procedures.** One of the recently proposed solutions for eliminating cognitive bias in forensic analyses is the use of contextual information management procedures (Kassin et al., 2013; Edmond et al., 2016; Cole, 2013; Mattijssen, et al., 2016; Risinger et al., 2002; Thompson, 2011). The goal of such procedures is to protect examiners from receiving potentially biasing and task-irrelevant information about a case while still providing them with case details that are considered necessary and relevant to conducting their analyses.

One specific approach discussed in the contextual information management literature involves the linear sequential unmasking procedure (Dror et al., 2015). It was originally introduced as the “sequential unmasking” protocol to be used with DNA evidence (Krane et al., 2008), but was later developed into a generalized step-wise procedure that could be used with other types of forensic evidence (e.g., fingerprints) and
was termed the ‘linear sequential unmasking’ procedure (Dror et al., 2015). The linear sequential unmasking procedure requires examiners to receive contextual information about the evidence and case in a linear, step-wise fashion. That is, initially examiners are only presented with information that is absolutely critical and directly relevant to the analyses requested (i.e., task-relevant information) before being asked to make an initial comparison. Only after an initial review and comparison of the evidence do examiners receive additional case information (i.e., task-irrelevant information). A critical element of the linear sequential unmasking procedure is that once examiners are presented with additional contextual information, they are allowed to revisit the evidence and make changes to their initial judgment. The idea is that examiners’ observations and conclusions will be documented at each step (Inman & Rudin, 2013). The linear sequential unmasking procedure is usually paired with the suggestion that a case manager be appointed to filter out the extraneous, task-irrelevant information and pass on only the task-relevant information to the forensic examiner (Archer & Wallman, 2016; Dror et al., 2015; Found & Ganas, 2013). Unfortunately, because of the lack of empirical research on the linear sequential unmasking procedure, there has been no clear explanation as to the exact steps that constitute this procedure or any detailed discussion of the challenges that may arise when applying it to real-world cases.

There have only been a few published case studies examining the use of the linear sequential unmasking procedure. For example, Archer and Wallman (2016) described a forensic entomologist’s experience with employing the sequential unmasking procedure in her lab and claimed that it increased the confidence in the validity of the examiner’s judgments. That is, by eliminating contextual influences, the authors anticipated the
examiner’s opinion about the evidence to increase in value to the court. However, the forensic entomologist who used the procedure served as both the “case manager” and the examiner, which led to the accidental disclosure of task-irrelevant and potentially biasing information and ultimately defeats the purpose of using sequential unmasking.

An alternative contextual information management approach is the context-manager model in which a context manager is responsible for deciding what information in a case file is passed along to the forensic examiner (Mattijssen et al., 2016). Thus, the context manager is the only individual who has complete access to the contextual information (both relevant and irrelevant information) and provides the examiner only with the information that is considered relevant to the examination. In this way, the forensic examiner conducts their analyses without any biasing information.

Found and Ganas (2013) described an implementation of the context-manager model at the Document Examination Unit of the Victoria Police Forensic Services Department. They stated that no negative outcomes had resulted from the implementation of the procedure and that the steps they had implemented were not overly time-consuming or expensive. However, no data were provided to support this claim. As such, it remains unknown exactly how much time and money were spent to implement such a procedure. In addition, no experimental research on the proposed case-manager model has been conducted to test the possible extent to which this method truly reduces the effect of cognitive bias when analyzing forensic evidence. Nevertheless, researchers continue to strongly advocate its use and, specifically, for the elimination of task-irrelevant information.
**Task-relevant versus task-irrelevant information.** In addition to the lack of information on the specific steps that should be taken by those who wish to use or examine different contextual information management procedures, there is also a lack of clarity surrounding what type of information is considered task-relevant and task-irrelevant (Inman & Rudin, 2013; Langenburg, 2017). Importantly, deciding what constitutes task-relevant and task-irrelevant information is a much more difficult process than what has been suggested by those advocating the use of a linear sequential unmasking procedure or context-manager model (Inman & Rudin, 2013). Task-relevant information has been generally defined as information that is directly related to the task at hand and includes things such as the physical pieces of evidence being analyzed (Dror et al., 2015). However, there is no exact definition or list of items that are considered “critical” enough to be labeled as task-relevant. This is mostly because what is considered to be task-relevant varies vastly on a case-by-case basis and depends largely on the type of evidence being examined (e.g., DNA versus fingerprints; Inman & Rudin, 2013; Langenburg, 2017). On the other hand, task-irrelevant information is loosely defined as information that is composed of “unstated assumptions, information that is collateral, and improper expectations and motivations” (Risinger et al., 2002, p.45), a rather ambiguous and unspecific definition.

**Practitioners’ Concerns**

Although research examining the effect of cognitive bias in forensic sciences has highlighted many potential dangers examiners face when presented with contextual information prior to examining evidence, practitioners in the field are often hesitant to accept these findings (Butt, 2013; Ostrom, 2009; Thornton, 2010). Specifically, expert
examiners have often critiqued this research for using inexperienced students as participants, having very small sample sizes and presenting participants with testing materials beyond their level of expertise. In addition, expert examiners have various concerns regarding the practicality of the recommendations researchers strongly advocate for, such as blind testing, evidence lineups and linear sequential unmasking procedures (Risinger et al., 2002; Thornton, 2010). Expert examiners argue that many of these recommendations would require significantly more resources, would require a host of decisions on an individual case level, and could in fact hinder their ability to make accurate interpretations of the evidence (Risinger et al., 2002; Thornton, 2010).

**Drawbacks to proposed solutions.** As previously discussed, blind testing, evidence lineups, and linear sequential unmasking procedures are all methods researchers recommend forensic examiners use to decrease the effects of cognitive bias when examining and interpreting evidence. However, each of these methods have their individual limitations and drawbacks that often deter real-world forensic examiners from incorporating them into their daily lab work. For example, practitioners continuously argue that blind testing may actually hinder their ability to accurately evaluate a piece of forensic evidence because certain contextual information is needed to properly analyze a piece of evidence (Risinger et al., 2002). For instance, fingerprint examiners may need to know the type of surface and where exactly the latent fingerprint was collected from. Such information (e.g., whether prints were lifted from the inside or outside of a car window) can inform examiners of extraneous factors (e.g., weather conditions) that can affect the quality of the print and subsequent analyses. Because of this common critique, researchers proposed the linear sequential unmasking procedure and the context-manager
model, which would allow examiners to have the necessary contextual information while remaining blind to task-irrelevant information that could bias their conclusions.

However, contextual information management procedures have not been fully accepted by forensic examiners either. Practitioners are often concerned with the additional resources that would be needed to implement a successful linear sequential unmasking procedure or a case-manager model in their lab (Charlton, 2013). For example, a case manager would need to be hired and extensively trained across forensic disciplines to decide what information could and should be made available to examiners, a difficult task that arguably escapes researchers outside the specific discipline. Hiring a case manager would also strain already existing financial resources and require that practitioners work within “leaner working models” (Charlton, 2013, p. 72). However, recent work on the implementation of a “case-manager model,” suggests it has minimal impact on the number of additional personnel required for the protocol, cost, or time (Found & Ganas, 2013).

Implementing evidence lineups would also necessitate more resources considering it would require an additional examiner to create the evidence lineup with similar-looking ‘foils’ (Risinger et al., 2002). Finding similar-looking foils is a particularly difficult task considering the variation of evidence (e.g., fingerprints) in terms of difficulty and similarity. For example, to properly create an evidence lineup for fingerprints, it would first require fingerprint foils to be piloted with a number of fingerprint experts who assess hundreds of prints to obtain average difficulty and similarity ratings for pairs of fingerprints. Thus, in order to properly select similar-looking ‘foils’ for evidence lineups, examiners would have to obtain from expert examiners an average similarity rating for
the target piece of evidence and each potential foil to ensure an “unbiased” lineup. If proper measures are not taken to obtain truly similar-looking foils, then the target piece of evidence may simply be chosen because it stands out from the rest of the samples in the evidence lineup.

Furthermore, a survey conducted by Kukucka, Kassin, Zapf, and Dror (2017), found that the majority of forensic examiners they questioned believed they were immune to bias or could reduce the effect of bias simply by willpower. As such, similar to other professions, forensic examiners may underestimate the extent to which bias may influence their work without a clear understanding of the automatic processes underlying this effect, which in turn may decrease their openness to modifying procedures. Commentaries written by forensic examiners about cognitive bias research often suggest practitioners feel insulted by such research and proposed solutions. For example, Thornton, an Emeritus Professor of Forensic Science at the University of California at Berkeley with 20 years of experience in operational crime laboratories wrote “I reject the insinuation that we do not have the wit or the intellectual capacity to deal with bias, of whatever sort” (Thornton, 2010, p. 1663). Despite the fact that research strongly suggests all individuals, regardless of profession, can fall prey to cognitive bias, concerns voiced by practitioners regarding the various drawbacks that come with many of the proposed solutions highlight the need for truly bi-directional research collaborations between researchers and practitioners to tackle cognitive bias in forensic examinations.

**Participant samples used in research.** Another concern practitioners have raised regarding the existing research examining the effect of cognitive bias in forensic contexts is the use of student participants who have no experience or training analyzing forensic
evidence (Ostrom, 2009). Although some research does suggest that contextual information, such as the decision of a prior examiner (Dror & Charlton, 2006; Dror et al., 2005), can also influence experts, sample sizes in studies using experts tend to be very small, mainly because of the difficulty of recruiting experts in the field to participate in research. For example, Dror and Charlton (2006) examined the effect of bias on fingerprint analyses using only six fingerprint examiners. Similarly, in a second separate study Dror, Charlton and Peron (2005), relied on data from five fingerprint examiners. It is important to point out, however, that both of these studies used a within-participant design, such that each pair of prints presented to the examiner-participants had been previously analyzed by the same examiner within the normal course of their work. Although using a within-participant design increases power, a sample of five to six expert examiners arguably does not provide a sufficient experimental power and thus basis to convincingly argue to practitioners that their own judgments can also be influenced by cognitive bias.

In addition, the concern of using students who have no experience analyzing forensic evidence is valid considering research has found that experience can moderate the effect of cognitive bias (Langenburg et al., 2009). For example, in a study conducted by Langenburg and colleagues (2009), both novice and expert examiners were recruited as participants. Novice participants were composed of students with no experience or training in fingerprint analyses, whereas expert participants were trained fingerprint examiners recruited at an international conference. All participants were assigned either to a control group, a high bias group or a low bias group. The control group received no contextual information whereas the two bias groups received information regarding a
previous examiner’s decision on the fingerprints that were to be analyzed. The low bias group was simply told that the previous examiner was trained to competency and the high bias group was told that the previous examiner was an internationally recognized expert. They found that novices were more likely to agree with the high bias prompt than were experts suggesting novices may be more influenced by bias than experts. Unfortunately, the majority of research examining the effects of cognitive bias on forensic evidence examination relies on collecting data from undergraduate students who have no experience analyzing forensic evidence such as fingerprints – a skill that requires a significant amount of training and expertise.

**Type of evidence.** Research has found that cognitive bias can affect the evaluation and interpretation of various types of evidence (e.g., Dror & Hampikian, 2011; Elaad et al., 1994; Kerstholt et al., 2010; Lange et al., 2011). However, different types of evidence require vastly different levels and training. For example, analyzing DNA evidence requires extensive academic training, analyzing fingerprint evidence requires extensive specialized training differing in content and time from other pattern evidence such as footwear impression evidence, which requires only basic visual-spatial abilities (International Association for Identification, 2018; Sirchie, 2018). Taken together, different pieces of forensic evidence require different levels of specialized training. Therefore, presenting participants, who have no forensic evaluation experience, with stimuli that require specialized training (e.g., fingerprints), may find an effect of bias not because of the presence of contextual information, but because of the difficulty of the task itself (Butt, 2013). The more ambiguous a decision-making task, the more likely a participant is to look for external cues to guide his/her decision, thus inflating the
estimated effect of cognitive bias (Dror et al., 2005). As such, results from studies using inexperienced examiners are likely overestimating the effects of cognitive bias. The use of stimuli that require specialized training coupled with the use of inexperienced examiners are therefore important and valid concerns practitioners have regarding the current body of work. The sampling and stimulus material issues present in the body of work thus far render it difficult for researchers to convince forensic experts of the likelihood and extent of cognitive bias found in these studies and to successfully advocate for the implementation of bias-reducing procedures.

**Present Study**

Decades of research on the effects of cognitive bias suggest that having expectations and prior beliefs can affect the way an individual interprets subsequent information (Rosenthal, 1994). Specifically, individuals tend to look for information that confirms their beliefs while ignoring information that disconfirms their beliefs (Ask & Granhag, 2005; Nickerson, 1998). Research examining the effect of cognitive bias within forensic contexts also finds a similar pattern of results, which has led to the development of various procedures including blind testing and contextual information management procedures, to decrease the effect of cognitive bias (Found & Ganas, 2013; Kassin et al., 2013; Saks et al., 2003). However, only few studies have examined blind testing methods and their ability to reduce the effect of bias on analyses of forensic evidence, with no study to date experimentally testing contextual information management procedures to eliminate task-irrelevant information with fingerprint and footwear impression evidence. Relatedly, this lack of research has resulted in only a vague description of the specific
steps that forensic examiners should take if implementing a linear sequential unmasking procedure or a case-manager model in their lab.

The current study directly compares the effects of blind testing, redaction of task-irrelevant information and informed testing on forensic evidence analyses. It also offers a first operationalization of eliminating task-irrelevant information within a field setting, in collaboration with real-world forensic experts from the Miami-Dade Forensic Sciences Bureau by using redaction procedures to eliminate any information from a lab analysis request form (LAR) that is not necessary for examination purposes. To fully test the potential for cognitive bias while investigating the strategy of contextual information management procedures, the proposed study assigned participants to receive either no information about the case summary (i.e., blind condition), a partially redacted LAR case summary (i.e., task-irrelevant details blacked-out), or a non-blind LAR summary (i.e., all case details available). In addition, after an initial judgment of the evidence was made, all participants received a full case summary (with all details available) and were asked if they would like to re-analyze the evidence, likened to real-world scenarios in which examiners receive additional case information after analyzing evidence when using the linear sequential unmasking procedure (Saks et al., 2003).

In line with past research on cognitive bias and forensic evidence analyses demonstrating that having contextual information about a case, including suggestive and guilt-presumptive details about a suspect, can result in more match judgments, than not having contextual case information (e.g., Dror et al., 2005; Kassin & Kukucka, 2014), it was hypothesized that redacting task-irrelevant information from LARs would reduce the effect of contextual bias. Specifically, it was hypothesized that participants provided with
a partially redacted LAR (i.e., no guilt-presumptive suspect information available) would be more accurate in their judgments compared to participants who received a non-blind, complete LAR (i.e., had guilt-presumptive suspect information available). That is, examiners who receive the complete LAR would be less likely to judge a pair of non-matching evidence samples as non-matches (i.e., correct exclusion) and more likely to judge a pair of non-matching evidence samples as matching (i.e., false ID).

It was also predicted that blind testing procedures would decrease the effect of bias such that participants would be more accurate at making match and non-match judgments of evidence (i.e., more correct exclusions and fewer false IDs) compared to participants who were provided with all case details about the case. However, it was also predicted that participants who received no information about the crime would be less likely to correctly judge a pair of matching samples of evidence as matching (i.e., correct ID) and more likely to incorrectly judge a pair of matching evidence samples as non-matches (i.e., incorrect exclusions) than those who received a partially redacted form, as participants who were not given any information about the case did not have available to them any task-relevant information compared to participants who were given a partially redacted LAR. Task-relevant information is believed to help examiners make more informed and accurate decisions (Inman & Rudin, 2013); participants who did not have this information are thus expected to be at a greater disadvantage to those who did have the task-relevant information on the partially redacted LAR. For example, knowing that the fingerprints being analyzed where found inside of a car allows examiners to discount any odd marks on the fingerprints as being a result from weather conditions.
To obtain a fine-grained understanding of how each of these LAR conditions affect examiners’ judgments, a signal detection approach was used. Specifically, examiners were presented with both matching and non-matching samples of evidence in order to calculate their discrimination accuracy, which measures the ability to accurately discriminate between matching and non-matching pairs of evidence. Discrimination accuracy takes into account the number of correct match (correct IDs) and false match decisions (false IDs) made. Thus, in line with the previous hypotheses, it was predicted that those provided with a partially redacted LAR would have significantly better discrimination accuracy than those who were provided with a blind or non-blind LAR as they would be more likely to render a correct match decision and less likely to make an incorrect match decision.

Furthermore, individuals’ response bias was calculated to measure examiners’ match and non-match judgement tendencies. A “match response bias” was defined as an individual’s tendency to judge a pair of evidence samples as matching regardless of whether they were matches or not, whereas a “non-match response bias” was defined as an individual’s tendency to make a non-match judgment regardless of whether they were non-matches or not. It was predicted that examiners who received the non-blind LAR would be more likely to show a match response bias compared to those who received the blind or partially redacted LAR because participants provided with the non-blind LAR would have information available to them that would bias them into making more match judgments (i.e., suspects are in custody) regardless of the evidence samples’ ground truth.

Finally, because of a lack of information, participants in the blind condition were predicted to be less confident in their judgements about the evidence compared to
participants who received either all the information about the case or a partially redacted case summary. Results from the present study thus served to inform which bias-reducing procedures allow examiners to be the most accurate and confident in their decisions.

A secondary aim of the present study was to compare novice and expert examiners’ performance when conducting forensic examinations as well as their vulnerabilities to contextual bias. Although some research suggests that experts can be influenced by task-irrelevant information (Dror et al., 2006), the majority of research examining the effect of cognitive bias on analysis of forensic evidence relies on collecting data from undergraduate students who have no experience analyzing forensic evidence that require at least some level of specialized training (Dror et al., 2005; Smalarz, Madon, Yang, Guyll, & Buck, 2016). Using participants who have no experience analyzing forensic evidence may inflate the effect cognitive bias may have on examiners’ interpretation of evidence as a consequence of its increased level of difficulty and thus ambiguity when assessed by novices. The present study therefore systematically investigated the differences in susceptibility to bias between novice and expert examiners. It was hypothesized that novices would be more inaccurate (i.e., make more incorrect match and non-match decisions) and have poorer discrimination accuracy than expert forensic examiners, because of their lack of experience in analyzing the stimuli used in the proposed study. For similar reasons, it was predicted that novice examiners would be less confident in their judgments compared to expert examiners.

In addition, it was predicted that the main effect of experience would be qualified by an interaction between contextual information available on the LAR and experience. Specifically, because of the lack of experience and thus increased level of the difficulty
of the task, novices were expected to be more influenced than experts by the task-
irrelevant contextual information when it was available (i.e., non-blind LAR condition)
compared to when they were kept blind to case details or when task-irrelevant
information had been redacted.

Not only is it important to systematically investigate the possibility that research
conducted on undergraduate students may be inflating the presence of bias in the analysis
of forensic evidence, but it is also important to examine how the type of stimuli used in a
study can influence the likelihood of finding an effect of bias. Thus, the third aim of the
current study was to investigate the effect of contextual bias on examiners’ evaluation of
forensic evidence as a function of stimulus material. Specifically, the present study
compared the effect of cognitive bias when comparing fingerprints and footwear
impressions, with fingerprint examination requiring significantly more expertise and
training than footwear impression analyses (International Association for Identification,
2018; Sirchie, 2018). This comparison could help inform the specific contexts in which
cognitive bias may play a reduced versus enhanced role in forensic examinations. In
addition, as past research has failed to manipulate the type of stimuli presented to
participants, it is difficult to compare results across studies that differ in the type of
stimuli presented (e.g., footwear impressions versus handwriting samples versus
fingerprints). Thus, the present study was the first to directly compare the effect of
cognitively biasing conditions across different forensic stimuli.

It was hypothesized that contextual information would have a weaker impact on
examiners’ evaluations when the to-be-examined forensic evidence required less
specialized training (i.e., footwear impressions) compared to more difficult evidence
samples (i.e., fingerprints). Finally, it was predicted that there would be an interaction between examiner experience and stimulus material such that differences found between novice examiners and expert forensic examiners would be greatest when evaluating fingerprint evidence compared to footwear impression evidence because novice, undergraduate examiners would be at a significantly larger disadvantage when analyzing fingerprints. Thus, novice examiners would make a larger number of mistakes, have poorer discrimination accuracy, and be less confident in their judgments compared to expert examiners, but more so when judging fingerprint evidence compared to footwear impression evidence.

In conclusion, the present study had the following primary objectives: (a) examine the effect of contextual bias on examiners’ evaluation of forensic evidence by varying the amount of pre-comparison information available to participants, (b) compare novice and expert examiners’ performance when conducting forensic analyses and their vulnerabilities to contextual bias, and (c) investigate the effect of contextual bias on examiners’ evaluation of different types of forensic evidence. Together, these objectives can inform both researchers and practitioners of the impact and feasibility of several proposed solutions in reducing the effect of contextual bias across different forensic domains. In addition, the current study directly addresses concerns voiced by forensic scientists and practitioners including the use of undergraduate students as participants and using stimuli that are considered to require more specialized training.

Two pilot studies were first conducted for the purpose of stimulus material selection and to minimize the potential of confounding variables in the Main Study. Pilot Study One was conducted to select the fingerprint evidence stimuli to be used in Pilot
Study Two and the Main Study. The primary aim of Pilot Study One was to obtain ratings for fingerprint stimuli pairs from undergraduate students who had no experience analyzing fingerprint evidence in order to select fingerprint pairs that were of comparable difficulty level to latent fingerprint experts’ ratings. In order to minimize the presence of a confounding variable (i.e., difference in perceived task difficulty) fingerprints of comparable difficulty for experts and novices were selected from Pilot Study One so that examiner experience, and not task difficulty, would yield any potential differences between novice and expert examiners the Main Study. Pilot Study Two was conducted to test the bias in the central manipulations of the Main Study. Specifically, the main aim of Pilot Study Two was to ensure that the LAR containing all the case information resulted in more incorrect match judgements than the LAR with no case information (i.e., blind testing procedure).

CHAPTER II
PILOT STUDY ONE METHODS

Participants

A total of 82 students ($N = 82$) at a large southeastern university were recruited via the university’s psychology research participation system in exchange for course credit. The final sample of 82 students was 78% female ($n = 64$) and representative of the community from which they were drawn (67.1% Hispanic, 14.6% African American, 9.8% Caucasian, 3.7% Asian American, and 4.9% Other). The sample ranged in age from 18 to 34 years ($M = 20.93, SD = 3.49$).
Materials

**Match fingerprint pair stimuli.** Fingerprints from a previous study examining the efficacy of the ACE-V fingerprint analyses procedure conducted by the Miami-Dade Forensic Services Bureau (MDFSB) (Pacheco, Cerchiai, & Stoiloff, 2014) were used for the current study. In the previous study, fingerprint standards and multiple latent prints were collected from nine different individuals. Latent prints are prints lifted from objects and surfaces whereas standard prints are those collected directly from the individual. A total of 80 final photographs of latent prints collected from flat and curved surfaces (i.e., plastic, tile, metal and glass) were provided by each of the nine volunteers (out of 720 total). Latent prints were lifted using black powder, tape, and white backing cards. These cards were then scanned to create individual photographs. Each individual’s prints were then rated by three certified latent print examiners on several dimensions including strength of value of latent print, whether the latent was in agreement with the standard, and difficulty of the comparison. The strength of value of the latent print referred to the clarity of the latent print and the agreement referred to whether the latent print and the standard matched in clarity and had the same minutiae present. Minutiae were defined as those characteristics of a fingerprint that make it unique compared to other fingerprints (e.g., bifurcations, ending ridges, dots).

The difficulty level of the comparisons was calculated by averaging the scores the print received on the strength of value and agreement with standard ratings. Each print was then categorized into one of the three following categories: (a) insufficient to difficult, (b) difficult to moderate, or (c) moderate to easy. On the basis of the study conducted by Pacheco and colleagues (2014), a total of 15 latent prints along with their
corresponding standard print were classified into the ‘moderate to easy’ category by three forensic examiners, nine were classified as ‘difficult to moderate,’ and 12 were classified as ‘insufficient to difficult.’

In order to address a critical limitation and confound present in previous research examining the effect of cognitive bias in forensic analyses, ‘moderate to easy’ fingerprints shown to expert examiners in Pachecho and colleagues’ (2014) initial study were pilot-tested with student examiners to ensure that the fingerprints chosen for the Main Study were rated as ‘moderate to easy’ by both forensic experts and students, who would constitute the participant pool in the proposed Main Study. The decision to include only the ‘moderate to easy’ difficulty level in the proposed final study as opposed to fingerprints categorized into one of the other two more difficult categories was determined because forensic examiners typically use specialized fingerprint marking software programs (i.e., CSIpix and PiAnoS) to analyze fingerprints which was unable to be included in this study. That is, for logistic purposes, the proposed study used Qualtrics, an online research system that allows the distribution of all research materials to participating students and experts across the country. As Qualtrics does not allow the incorporation of the fingerprint marking software commonly used by forensic examiners and student examiners are unfamiliar with this software, the researcher instead opted for the inclusion of only ‘easy to moderate’ fingerprints. This allowed for both student and expert examiners to compare prints without software while simultaneously eliminating the alternative explanation for differences across expertise being due to familiarity with the software.
On the basis of the results of Pachecho and colleagues’ (2014) initial study, the final stimulus material for matching fingerprint pairs for this pilot study consisted of a total of 15 standard and 15 latent ‘moderate to easy’ difficulty level fingerprints. In addition, in order to increase the range of ratings participants used, two ‘difficult to moderate’ and two ‘insufficient to difficult’ matching fingerprint pairs were also included into the final stimulus material of Pilot Study One, amounting to a total of 19 matching fingerprint pairs.

**Non-match fingerprint pair stimuli.** Each of the 19 standard prints were combined with each of the latent prints of the same difficulty category that did not originate from the same person in order to develop the non-matching print pairs. Thus, a total of 200 non-matching fingerprint pairs were created for piloting.

**Procedure**

Student participants were scheduled to come into the lab to participate in the study. Once participants arrived at the lab, they were seated in front of a computer. All study procedures were administered via Qualtrics. After online consent was obtained, each student participant was randomly presented with 13 of the 15 ‘moderate to easy’ matching fingerprint pairs, one of the two ‘difficult to moderate’ matching fingerprint pairs, one of the two ‘insufficient to difficult’ matching fingerprint pairs, and 15 of the non-matching fingerprint pairs, totaling 30 fingerprint pairs.

Fingerprint pairs were shown one at a time and participants were asked to judge each pair of prints on three different rating scales that were similar to those used by the expert latent print examiners in Pachecho and colleagues’ (2014) study. To account for undergraduate students’ lack of expertise, the rating scale was modified. Specifically,
students were not asked to count the number of minutiae on each fingerprint and were instead simply asked to rate the clarity of the latent print on a 7-point scale with higher numbers indicating better clarity. Additionally, after viewing each pair of prints, students were asked to rate the similarity between the fingerprint pairs on a 7-point scale with higher numbers indicating higher similarity. Finally, students were asked to determine whether the prints were a match, non-match or inconclusive and rate the difficulty of the comparison on a 7-point scale with higher numbers indicating higher difficulty.

CHAPTER III

PILOT STUDY ONE RESULTS

The clarity of the latent print (reverse coded), similarity of prints, and difficulty of comparison ratings were averaged across participants for each pair of fingerprints. Those pairs receiving an average difficulty rating between five and seven points were categorized into the ‘insufficient to difficult category,’ those between three and five points were categorized as ‘difficult to moderate,’ and those between one and three points were categorized as ‘moderate to easy.’ Once all fingerprint pairs had been categorized into one of the three categories, only those fingerprint pairs that both experts and students had labeled as ‘moderate to easy’ were selected to be used in Pilot Study Two and the Main Study.

Descriptive statistics were used to analyze the average difficulty rating for each of the matching and non-matching fingerprint pairs. Once the average difficulty rating was calculated for each pair of fingerprints, they were classified into one of the three difficulty level categories (moderate to easy, difficult to moderate, and insufficient to difficult). Frequencies for the total number of fingerprints classified as ‘moderate to easy’ were selected to be used in Pilot Study Two and the Main Study.
easy,’ ‘difficult to moderate,’ and ‘insufficient to difficult’ were computed in addition to
the percentage of correct, incorrect, and inconclusive decisions made for each match and
non-match fingerprint across the three difficulty levels.

**Match Fingerprint Judgments**

Across all three difficulty category levels, none (0%, n = 0) of the matching
fingerprint pairs were perceived by students as ‘moderate to easy’ or ‘insufficient to
difficult.’ Instead, all (100%, n = 19) were classified as ‘difficult to moderate.’ In
addition, for ‘moderate to easy’ match fingerprint pairs, the majority of participants made
accurate match decisions (47%), followed by inconclusive decisions (27.91%) and
incorrect non-match decisions (24.59%). For ‘difficult to moderate’ match fingerprint
pairs, most participants made inconclusive decisions (40.1%) followed by incorrect non-
match decisions (36.2%) and correct match decisions (23.7%). For ‘insufficient to
difficult’ match fingerprint pairs, the majority of participants made inconclusive
decisions (78%) followed by incorrect non-match decisions (20.7%) and correct match
decisions (1.3%; see Figure 1).

**Non-Match Fingerprint Judgments**

Similar results were found for non-match fingerprint pairs such that across all
three difficulty categories, the majority of comparisons were rated as being ‘difficult to
moderate’ (97.62%). Specifically, for fingerprint pairs in the ‘moderate to easy’ category,
only 5.1% of the pairs were perceived by students as ‘moderate to easy,’ with the
majority of students (92.86%) rating them as ‘difficult to moderate’ and some even
perceiving them as ‘insufficient to difficult’ (1.5%). For fingerprint pairs in the ‘difficult
to moderate’ and ‘insufficient to difficult’ category, all pairs (100%) were perceived as being ‘difficult to moderate’ by student participants.

In addition, for ‘moderate to easy’ non-match fingerprint pairs, the majority of participants made accurate non-match decisions (50.32%), followed by inconclusive decisions (29.05%) and incorrect match decisions (20.64%). For ‘difficult to moderate’ non-match fingerprint pairs, most participants made inconclusive decisions (49%) followed by correct non-match decisions (33.80%), and incorrect match decisions (17.20%). For ‘insufficient to difficult’ match fingerprint pairs, the majority of participants made inconclusive decisions (71.05%) followed by correct non-match decisions (28.95%) and no participants made an incorrect match decision (0%; see Figure 2).

**Selection of Match and Non-Match Fingerprint Pairs for Pilot Study Two**

Unfortunately, albeit unsurprisingly, students perceived the fingerprint comparisons to be more difficult than the expert latent fingerprint examiners as evidenced by the very few pairs categorized by students as ‘moderate to easy.’ Since none of the match fingerprints were perceived to be ‘moderate to easy,’ the four match fingerprint pairs with the lowest average difficulty ratings were selected to be used for Pilot Study Two. For selection of the non-match fingerprint pairs to be used in Pilot Study Two, all non-match fingerprints that fell within the range of the difficulty rating of the four match fingerprint pairs selected were grouped together. From these, four non-matching fingerprint pairs were randomly selected to pair (in terms of difficulty level) with the matching fingerprint pairs.
CHAPTER IV
PILOT STUDY TWO METHODS

Design

A 2 (Contextual Information: blind vs. non-blind) X 2 (Stimulus Material: footwear impression vs. fingerprint) mixed factorial design was employed with stimulus material and stimulus pair manipulated within participants. The primary aim of Pilot Study Two was to test the bias of the central manipulation of the Main Study.

Participants

A total of 96 (N = 96) students at a large southeastern university were recruited via the University’s psychology research participation system in exchange for course credit. Of these participants, two participants were excluded because they did not consent, 36 participants because they did not pass the instructions quiz, and 10 participants because they did not pass the manipulation check question. The final sample of 48 (n = 48) students was 83.3% female (n = 40) and representative of the community from which they were drawn (68.8% Hispanic, 25% African American, 4.2% Caucasian, and 2.1% Asian American). They ranged in age from 19 to 61 years (M = 22.94, SD = 7.52).

Materials

Lab analysis request forms. Lab analysis request forms mirrored those used by the MDFSB (see Appendix A). They contained basic information such as the name of the lead investigator, type of offense, name of suspects, name of victims, description of the case, and description of the evidence. These pieces of information are commonly found on LARs (sometimes referred to as evidence submission forms) used across the United
States (Gardner, Kelley, Murrie, & Blaisdell, 2019). Specifically, in a recent study, Gardner and colleagues (2019) collected LARs from different forensic labs across the United States to gather information on the type of information that is typically included on such forms. They found that many request similar information compared to the LAR used in the proposed study. For example, 94.8% ask about the type of offense, 94.8% ask for the name of suspect(s), 92.8% ask for the name of victim(s), 45.5% ask for a description of the case, and 88.7% ask for a description of the evidence. As such, the LAR used in the proposed study was developed in accordance with this information to be representative of information typically gathered in other U.S. forensic lab’s LARs.

**LAR case summaries.** A total of two versions of the LAR were piloted in the present study to ensure differences between providing all case summary details (i.e., non-blind condition) and providing no case summary details (i.e., blind condition) on examiners’ judgements of fingerprint and footwear impression evidence. Specifically, participants in the ‘non-blind’ condition received all case details in the summary, as is currently the case in the MDFSB, which consisted of both task-relevant and irrelevant (e.g., potentially biasing) information. The case summary was developed in close collaboration with a forensic expert to allow for a realistic version and portrayed a real-world armed robbery in Miami, Florida that involved one female victim and two male perpetrators. In addition, the case summary included various pieces of information that are commonly found on LARs, but have the potential to bias examiners’ decisions when analyzing forensic evidence.

Initially, within the non-blind LAR, participants were made aware that both subjects were in custody, were given information about the suspects’ criminal history and
were also told that surveillance footage of the two suspects committing the crime was obtained. The objective was to bias participants into thinking that the suspects in custody were likely to be guilty, potentially leading participants to make more incorrect match judgements when analyzing non-matching pieces of forensic evidence compared to participants who did not receive such biasing information of the suspects. On the other hand, participants in the ‘blind’ condition were given a case summary containing only the word “routine” (see Appendix B). This served as the ‘blind’ condition because participants had neither case-specific nor biasing information that was provided to participants in the non-blind condition. After consulting with the commander of the MDFS, the word “Routine” was chosen to be entered in the case summary section of the blind LAR to increase external validity as this term is commonly found on LARs, despite the fact that it provides examiners with no substantive information about the case.

Preliminary analyses conducted during the initial stages of data collection revealed an unsuccessful manipulation of the LAR form such that we found a significant effect of LAR form on participant’s accuracy rate, $F(1, 82) = 4.34, p = .040$, in the opposite expected direction. Specifically, participants who received the biasing LAR had higher accuracy rates ($M = 0.46, SD = 0.28$) than participants who received the blind LAR ($M = 0.33, SD = 0.25$). It was believed that some of the information provided to participants in the biasing condition (i.e., surveillance footage of suspects committing the crime) led student participants to become particularly skeptical of the case summary information and pay more attention to the evidence when analyzing the footwear impressions and fingerprint evidence. As a result, we decided to decrease the level of “obvious” bias by eliminating the surveillance footage evidence found in the biasing
LAR case summary for the remainder of pilot study two. However, participants were still provided with information regarding the suspects’ custody status and criminal history along with the remaining pieces of information regarding the armed robbery (see Appendix C).

**Fingerprint stimuli pairs.** Fingerprints selected for this pilot study were chosen from the same pool of fingerprints used for Pilot Study One. However, for Pilot Study Two only the easiest non-matching fingerprint pairs were used. Due to the Pilot Study One results (no match fingerprint pairs were categorized as ‘moderate to easy’), it was necessary to choose non-matching fingerprint pairs that had comparable difficulty ratings to matching fingerprint pairs since both match and non-match fingerprint pairs would be used for the Main Study. Thus, four non-matching fingerprint pairs were randomly selected from a pool of non-matching fingerprint pairs that had similar difficulty ratings as the matching fingerprint pairs with the lowest difficulty ratings ($M_{\text{difficulty}} = 3.37 – 4.29$). Four fingerprint pairs were chosen for stimulus generalization purposes. The average difficulty rating for each of the four fingerprint pairs did not significantly differ from one another (all $p > .05$). Of the four randomly selected non-matching fingerprint pairs chosen for Pilot Study Two, each participant was presented with only two fingerprint pairs due to study material, design and procedures. That is, on the LAR provided to participants, there were only two fingerprints that were collected at the crime scene.

**Footwear impression stimuli pairs.** Footwear impressions were selected from a footwear impressions matching exam that was previously used by the MDFSB as part of their interview for entry level forensic examiners. The exam consisted of 10 matching
and 94 non-matching footwear impression pairs. These footwear impressions’ difficulty levels had been piloted as part of another ongoing study. A total of 145 undergraduate students were recruited as part of this previous pilot study and were asked to classify each matching and non-matching footwear impression pair as either a match, non-match, or made an inconclusive decision. They were also asked to rate their confidence in the decision and the level of difficulty on a 7-point Likert scale (1 = not at all difficult/confident to 7 = extremely difficult/confident). Footwear impression pairs were categorized into three levels based on these pilot data. Similar to the fingerprint evidence pairs in Pilot Study One, footwear impression pairs were categorized as either a) insufficient to difficult, b) difficult to moderate, or c) moderate to easy.

Similar to the fingerprint evidence, an average difficulty and confidence rating for each footwear impression pair was computed. Those comparisons that received between five to seven points were categorized into the ‘insufficient to difficult’ category, those with three to five points were categorized as ‘difficult to moderate,’ and those which received one to three points were categorized as ‘moderate to easy.’ For the current pilot study only ‘moderate to easy’ non-matching footwear impression pairs were used. Specifically, of the 94 non-matching pairs, 25 were labeled as ‘moderate to easy.’ Of these 25, four non-matching footwear impression pairs were randomly selected to be used for Pilot Study Two for stimulus generalization purposes. Each participant was randomly presented with two of the possible four footwear impression pairs due to study material, design and procedures. That is, on the LAR provided to participants, there were only two footwear impressions that were collected at the crime scene.
Procedure

Participants completed all study procedures via a Qualtrics online survey. After online consent was obtained, participants received basic task instructions. Specifically, they were told that they would be presented with an LAR and given information explaining what an LAR is and the type of information that can be found on the form. Participants were also informed that they would be asked to analyze various pieces of evidence from a crime scene including two footwear impressions and two fingerprints and make a match, non-match, or inconclusive decision. A matching pair was defined as a pair of prints originating from the same source (i.e., they came from the same person) and a non-matching pair was defined as prints originating from different sources (i.e., they came from different people). An inconclusive decision was defined as prints having insufficient detail for a conclusive determination to be made (e.g., “I don’t know” response). Participants were then instructed to carefully read the LAR before making any decisions and were told that they could refer back to the LAR whenever they felt necessary.

Once participants read and understood the initial instructions, they were randomly presented with one of the two possible LARs (i.e., blind or non-blind LAR). After reading the LAR, they were randomly shown one of the four pieces of evidence. Specifically, they were shown either one of the four possible non-matching footwear impression pairs or one of the four non-matching fingerprint pairs. After analyzing the first pair of evidence, they were asked to make a match or non-match decision. The LAR was available to participants during each comparison. After completing the ratings for the first pair of evidence (either fingerprint or footwear impression), they were shown the
second piece of evidence. The same procedure was employed for the remaining pieces of evidence. Once participants had completed comparing all four pieces of evidence, they were provided with a demographics questionnaire. After the completion of the demographics questionnaire, participants were debriefed and thanked for their time.

**Coding of Outcome Variables**

**Match and non-match judgments.** Participants’ match and non-match judgements were coded as correct or incorrect. As all pairs of evidence presented to participants were non-matches, all match judgements were coded as incorrect and all non-match judgments were coded as correct. Average accuracy scores for fingerprints and footwear impressions were calculated by dividing the total number of correct judgments made by the total number of judgments made.

**CHAPTER V**

**PILOT STUDY TWO RESULTS**

Only data from the participants recruited after the non-blind LAR was updated were included in the following analyses. Thus, participants who received the original non-blind LAR which included the surveillance footage detail were excluded from all subsequent analyses.

**Match Judgments**

A two-way mixed ANOVA was conducted to investigate the effect of LAR and evidence type on number of incorrect match judgments made. There was no significant effect of LAR, $F(1, 46) = 0.86, p = .358$ on number of incorrect match judgments. However, there was a significant effect of evidence type, $F(1, 46) = 39.09, p < .001$ such that participants made more incorrect match judgments when judging footwear.
impression evidence ($M = 1.11, SE = 0.12$) than when judging fingerprint evidence ($M = 0.36, SE = 0.08$). There was no significant interaction between LAR and evidence type, $F(1, 46) = 0.07, p = .788$.

**Non-Match Judgments**

A two-way mixed ANOVA was conducted to investigate the effect of LAR and evidence type on number of correct non-match judgments made. There was no significant effect of LAR, $F(1, 46) = 1.00, p = .753$ on number of correct non-match judgments. However, as with match judgments, there was a significant effect of evidence type on non-match judgments, $F(1, 46) = 18.06, p < .001$ such that participants made more correct non-match judgments when judging fingerprint evidence ($M = 1.07, SE = 0.10$) than when judging footwear impression evidence ($M = 0.57, SE = 0.12$). There was no significant interaction between LAR and evidence type, $F(1, 46) = 0.49, p = .489$.

**Inconclusive Judgments**

A two-way mixed ANOVA was conducted to investigate the effect of LAR and evidence type on number of inconclusive judgments made. There was no significant effect of LAR on the number of inconclusive decisions made, $F(1, 46) = 2.85, p = .089$ and no significant effect of evidence type, $F(1, 46) = 3.99, p = .052$. The interaction was also non-significant, $F(1, 46) = 0.83, p = .368$.

**Overall Accuracy Rate**

A two-way mixed ANOVA was conducted to investigate the effects of LAR and evidence type on participants’ overall accuracy rate. Although there was no significant effect of LAR on accuracy rate, $F(1, 46) = 0.10, p = .753$, it is important to note the successful shift in results compared to the preliminary findings reported above.
Specifically, changes to the LAR resulted in lower accuracy rates – albeit not significantly so - in participants who were presented with the biasing LAR ($M = 0.39, SE = 0.07$) than those presented with the blind LAR ($M = 0.42, SE = .06$), which was the opposite of what was found before changes were made. Unfortunately, due to time constraints and difficulties with participant recruitment, we were unable to collect more data, which may have increased power and thus rendered this difference statistically significant. However, because we did see a change in the hypothesized direction of the effect the same LARs were used to manipulate contextual information in the Main Study. There was also a significant effect of evidence type, $F(1, 46) = 18.06, p < .001$, such that participants were significantly more accurate when judging fingerprint evidence ($M = 0.53, SE = 0.05$) than when judging footwear impression evidence ($M = 0.28, SD = 0.06$). The interaction between LAR and evidence type was not significant, $F(1, 46) = 0.49, p = .489$.

**CHAPTER VI**

**MAIN STUDY METHODS**

**Design**

A 3 (Contextual Information: blind vs. non-blind vs. partially redacted) X 2 (Examiner Experience: expert vs. novice) X 2 (Stimulus Material: footwear impression vs. fingerprint) X 2 (Stimulus Pair: match vs. non-match) mixed factorial design was employed with stimulus material and stimulus pair manipulated within participants.

**Participants**

Novice examiners were composed of undergraduate students from a large southeastern university. A total of 442 novices ($N = 442$) were recruited via the
university’s psychology research participation system in exchange for course credit. Of these participants, 155 participants were excluded because they did not pass the instructions quiz, they did not pass the LAR Questionnaire, and/or they did not allow us to use their data after being debriefed. The final sample of 287 novices \((n = 287)\) was 83.3\% female \((n = 239)\) and representative of the community from which they were drawn (70.4\% Hispanic, 14.6\% African American, 11.5\% Caucasian, 1.4\% Asian or Pacific Islander, 0.3\% Native American or American Indian, and 1.4\% Other). They ranged in age from 18 to 63 years \((M = 22.85, SD = 5.89)\). One participant did not report their gender, age or ethnicity.

Expert examiners were composed of latent fingerprint expert examiners from forensic laboratories across the United States. A total of 59 experts \((N = 59)\) were recruited via multiple forensic laboratory email listservers, websites, and newsletters. Of these participants, four participants were excluded because they did not consent, did not pass the instructions quiz, did not pass the LAR Questionnaire, and/or they did not allow us to use their data after being debriefed. The final sample of 55 experts \((n = 55)\) was mainly female (72.7\%) and Caucasian (90.9\%), followed by Hispanic (3.6\%), Asian or Pacific Islander (1.8\%), and Native American or American Indian (1.8\%). One participant did not report their gender, age, or ethnicity. They ranged in age from 23 to 63 years \((M = 39.15, SD = 10.72)\) with an average number of 10.86 years of active casework experience \((M = 10.86, SD = 9.06)\).

**Materials**

**Fingerprint stimuli pairs.** Each participant was presented with two of the four non-matching fingerprint pairs used in Pilot Study Two. The difficulty ratings of these
two *non-matching* fingerprint pairs were then used to select two *matching* fingerprint pairs to ensure that matching and non-matching fingerprint pairs had comparable difficulty levels. Thus, a total of four fingerprint pairs with similar (low) difficulty ratings were selected for stimulus generalization purposes. Prior to data collection, analyses were conducted to ensure that matching and non-matching fingerprint pairs did not differ in average difficulty rating. Results revealed that the fingerprints did not significantly differ from one another (all \( p > .05 \)).

Although fingerprint pairs selected for the Main Study were rated by experts as ‘moderate to easy,’ it is important to note that, as shown by Pilot Study One results, students found these to be ‘difficult to moderate’ (no match fingerprint pairs were rated as ‘moderate to easy’ by students). Nonetheless, those chosen for the Main Study consisted of the pairs students rated as being the easiest of the ‘difficult to moderate’ category (\( M_{\text{Difficulty}} = 4.30, SD = .90 \)).

Each participant was randomly presented with one matching and one non-matching fingerprint pair from the final pool of fingerprints. Participants were only shown a total of two fingerprint pairs due to study material, design and procedures. That is, participants received an LAR describing the evidence collected at the crime scene and were told that only two fingerprints were collected and needed to be analyzed before being presented with the fingerprints.

**Footwear impression stimuli pairs.** Similar to the fingerprint stimuli pairs, two of the four non-matching footwear impression pairs used in Pilot Study Two were randomly selected. The difficulty ratings of these two footwear impressions pairs were then used to select the two matching footwear impression pairs to ensure that both
matching and non-matching pairs selected had comparable difficulty levels. Thus, a total of four footwear impression pairs categorized as ‘moderate to easy’ were selected. Analyses were conducted to ensure that matching and non-matching footwear impression pairs did not differ in average difficulty rating. Results revealed that the footwear impression pairs’ difficulty rating did not significantly differ from one another (all $p > .05$). Given that the ability to compare the footwear impressions used in the proposed study was not expected to differ across levels of expertise, these stimuli were not piloted with forensic experts.

Each participant was randomly presented with one matching and one non-matching footwear impression pair from the final pool of two matching and two non-matching footwear impressions. Participants were only presented with a total of two footwear impression pairs due to study material, design and procedures. That is, participants received an LAR describing the evidence collected at the crime scene and were told that only two footwear impressions had been collected at the crime scene and needed to be analyzed.

**General task instructions.** Participants were provided with general instructions that provided them with information regarding the tasks they would be asked to complete. First, participants were told that they would be presented with evidence obtained from a real-world case and their job was to accurately analyze various pieces of forensic evidence. Participants were also told that before analyzing the forensic evidence, they would be presented with an LAR containing information about the case and evidence. Before being presented with the LAR, participants received a description of the forensic evidence and were told that they would be shown various prints, some of which would be
footwear impressions and others fingerprints. Further clarification was provided to participants by stating that they would be presented with footprint impressions and latent fingerprints that were collected at a crime scene along with their standards. Basic definitions for standard and latent prints were also provided to all participants. Participants were then instructed to compare each latent print (i.e., print collected at the crime scene) with its standard (i.e., known print collected directly from individual) and judge whether they were a match, non-match, or inconclusive. Similar to Pilot Study Two, a match was defined as a pair of prints originating from the same source (i.e., they came from the same person), a non-match was defined as prints originating from different sources (i.e., they came from different people), and an inconclusive decision was defined as prints having insufficient detail for a conclusive determination to be made (e.g., “I don’t know” response).

**Lab analysis request forms.** The same blind and non-blind lab analysis request forms from Pilot Study Two were used for the present study in addition to a third version of the LAR. Thus, there were a total of three versions of the LAR in order to manipulate the amount of contextual information given to examiners and to examine the effect of redacting task-irrelevant information. Lab analysis request forms across all three experimental conditions were identical with the exception of the “Brief Case Summary” section. That is, the name of the lead investigator, the type of offense, information found in the routing table, etc. was included in all LARs. However, like in Pilot Study Two, the amount of available information about the case varied depending on each participants’ randomized condition. Specifically, participants in the ‘non-blind’ condition received all case details in the summary and those in the ‘blind’ condition were given a case summary
containing only the word “Routine.” The third version of the LAR was the ‘partially redacted’ condition in which participants received the same case summary as the non-blind participants, but with task-irrelevant details redacted (i.e., blacked-out; see Appendix D). As suggested by various contextual information management procedures including the linear sequential unmasking procedure, only the critical and task-relevant details were available to participants. To decide which details should be redacted, a forensic expert, the Senior Police Bureau Commander of the MDFS, was consulted to determine which details were absolutely necessary for examiners to know (see Appendix E for list of critical details). All other task-irrelevant details were redacted.

**Time 1 dependent measures.** After viewing each pair of prints, participants were asked to make a match, non-match or inconclusive judgment. A match judgment indicated that the two pieces of evidence shown originated from the same individual, whereas a non-match judgment indicated that the pair of evidence did not originate from the same individual. An inconclusive judgment was made if the participant felt that there was insufficient detail in the prints to make a conclusive decision. Participants were also asked to rate the confidence in their decision on a 7-point Likert scale with higher numbers indicating higher confidence.

**LAR questionnaire.** All participants received an LAR questionnaire to ensure they had understood the instructions provided to them and had read the LAR, thus also serving as a manipulation check questionnaire. Participants were asked questions regarding information they may or may not have been exposed to in the case summary (see Appendix F). All participants were first asked to write down everything they remembered about the information provided in the lab request form before answering a
series of multiple-choice questions. Participants were asked basic questions regarding the type of case summary they received and the evidence they were asked to analyze. Next, they were asked to answer questions regarding details of the crime that were consistent across the three versions of the LAR. Specifically, all participants received questions about the type of crime that occurred, how many suspects were involved, how many victims were involved, and what item was brought in as evidence. Answers across conditions were expected to be similar given that all LARs contained this information.

Next, participants were asked multiple-choice questions regarding specific information that may or may not have been available in the case summary depending on the condition they were assigned to. Specifically, they were asked questions about where the victim was coming from, where the crime took place, what the suspects demanded from the victim, what weapon was used, what item was taken from the victim, where the car was parked, and how many suspects were in custody. Participants were also presented with various options including an “Information not given” response, which would be the correct answer for the specific questions that asked about pieces of information that were unavailable to participants in the partially redacted and blind conditions.

**Time 2 dependent measures.** After being presented with an LAR containing all information at Time 2 (see procedure), participants were asked if they would like to change any of the previous decisions they made concerning the pieces of evidence they analyzed. If they indicated that they would like to change any of their initial decisions, they were asked to judge the same pieces of evidence a second time and make a match, non-match or inconclusive decision. If they decided to change their initial decision, they were also asked to provide an explanation as to why they decided to make the change.
**Perceptions questionnaire.** Participants were also presented with a questionnaire that was used to gauge their overall perception of the evidence, the LAR and their experience participating in the study (see Appendix G). They were first asked to rate how accurate they believed they were in analyzing the different pieces of evidence on a scale from 0 – 100% and were also asked to rate how guilty they believed each suspect was on a scale from 1 (definitely innocent) to 10 (definitely guilty). In addition, depending on what condition they were assigned to, they answered questions about whether they thought the case summary provided sufficient information to make accurate decisions and whether they thought they would have been influenced by having additional information about the case. Finally, they were asked what type of LAR they preferred to have when analyzing forensic evidence.

**Demographics questionnaire.** Participants were also asked basic demographic questions regarding their age, gender, ethnicity, education, and occupation (see Appendix H). Forensic experts were asked additional questions about their work experience as forensic latent fingerprint examiners. Specifically, experts were asked about the numbers of years of active casework experience they had completed, the type of law enforcement agency they were currently employed by, whether they examine other types of forensic evidence in addition to latent print examinations, and the type and duration of any structured latent fingerprint training program they had received.

**Debriefing script.** The debriefing script informed participants of the true nature of the experiment after participants completed the experiment and provided them with the contact information of the researchers in charge of the study. They were also given the
opportunity to deny the use of their data for final analyses after reading the debriefing script.

**Procedure**

For recruiting purposes, participants were told that the study consisted of examining various pieces of forensic evidence that were taken from a real-world criminal case. Participants were under the impression that the main purpose of the study was to examine the importance of piloting materials (e.g., case reports, lab analysis request forms, and pieces of forensic evidence) used for routine proficiency testing of forensic scientists. However, after the collection of the first 14 expert examiners, it was suggested that recruitment would not go well with the existing “cover story” because of examiners’ expected skepticism regarding the materials being part of existing proficiency exams. They were expected to be skeptical because of their experience with professionally distributed proficiency exams. Thus, the wording used for recruitment purposes was made more general such that participants were instead told that the purpose of the current study was simply to examine the importance of testing stimulus materials (e.g., police case reports, evidence submission forms, and pieces of forensic evidence). The only changes made were the deletion of the “proficiency testing” wording. This change was implemented after having collected 14 expert participants and 189 student participants. We do not believe that this change in recruitment wording had an effect on student participants performance, given that when asked what they believed the study was about, only one of the 189 students recalled that the purpose of the experiment had something to do with testing materials for proficiency exams suggesting students did not pay much attention to the cover story.
Once participants decided to participate in the study, they were provided with a Qualtrics survey link and brief instructions stating that they would have as much time as they need to complete the study. However, they were also told that the study was to be completed in one sitting. All study procedures were administered via the university’s online Qualtrics survey program. Novices (i.e., undergraduate students) were provided with the survey link via SONA systems in order to receive credit for participating, whereas experts were emailed the link to the survey or clicked on the survey link if it appeared on a laboratories’ website or newsletter. All participants completed the online study alone and on their own time.

After online consent was obtained, participants received the general task instructions regarding the specific tasks they would be asked to complete. Once participants finished reading the instructions, they were asked to take a quiz about the instructions they had just read. After the completion of the instructions quiz, participants were randomly presented with one of the three LARs, either containing all contextual information, partially redacted contextual information, or no contextual information. Following the LAR, participants were randomly presented with either the first pair of footwear impressions or fingerprints, randomly assigned to be either a match or non-match. Participants were allowed to refer back to the LAR as many times as needed while comparing the prints. Once they judged the first pair of prints as a match, non-match, or inconclusive, they were asked to indicate the confidence in their judgment. Next, participants were randomly presented with one of the other three remaining pieces of evidence (a total of two fingerprints and two footwear impressions).
After analyzing the two pairs of fingerprints (one match and one non-match) and two pairs of footwear impressions (one match and one non-match), all participants were presented with the LAR questionnaire. After completion of the LAR questionnaire, all participants received the non-blind LAR, regardless of the LAR condition they were originally assigned to. After reading the non-blind LAR, participants were asked if they would like to change any of their initial decisions. Participants who answered “Yes” to this question were presented with the four pieces of evidence for a second time along with a reminder of their prior decision. Every time they changed their initial decision, they were asked to provide an explanation as to why they decided to change their match, non-match or inconclusive decision. This was done to mirror proposed linear sequential unmasking procedures. After participants finished re-analyzing the evidence, they were provided with the perceptions questionnaire and finished the study by completing the demographics questionnaire. Participants who answered “No” when asked if they would like to change any of their initial decisions were instead taken directly to the perceptions questionnaire. Once participants finished answering the demographics questionnaire, they were debriefed and thanked for their time.

**Coding of Outcome Variables**

**Time 1 judgements.** Time 1 judgements are those participants made prior to being exposed to the second, complete LAR and are thus comprised of two fingerprint pair judgements and two footwear impression pair judgements. Fingerprint and footwear impression judgments were coded as correct, incorrect, or inconclusive. A decision was scored as correct when a participant judged a pair of matching stimuli as matches (i.e., correct identification) or a pair of non-matching stimuli as non-matches (i.e., correct
exclusion). A decision was scored as incorrect when a participant incorrectly judged a pair of matching stimuli as non-matches (i.e., incorrect exclusion) or a pair of non-matching stimuli as matches (i.e., false identification). In line with previous research, inconclusive decisions were not coded as correct or incorrect (Kerstholt et al., 2010; Langenburg, Champod, & Wertheim, 2009), but were instead coded separately. Overall accuracy was also calculated and was defined as the proportion of judgements made accurately. It was calculated by dividing the total number of correct judgments made (correct identifications and correct exclusions) by the total number of judgments made.

A signal detection approach was also used to determine how well examiners discriminated between matching and non-matching pairs of evidence (i.e., discrimination accuracy) and whether there were differences in examiners’ tendency to respond (i.e., response bias). Discrimination accuracy (d’) was calculated by subtracting the z-score of false IDs from the z-score of correct IDs. Response bias (c) was calculated by adding the z-score of correct IDs and the z-score of False IDs, multiplying the outcome by negative one, and dividing it by two.

**Time 2 judgments.** Participants were asked whether they would like to re-analyze the evidence after being exposed to the non-blind LAR case summary information, which contained a complete case summary. For individuals who decided to change their Time 1 judgments, their Time 2 judgements were coded for the type of change that was made. A correct change, for example, was one in which the participant either a) incorrectly judged a matching pair of stimuli as non-matching or inconclusive at Time 1 and changed their decision to matching or b) incorrectly judged a non-matching pair of stimuli as matching or inconclusive at Time 1 and changed their decision to non-
matching. An incorrect change, on the other hand, was one in which the participant either 
a) correctly judged a pair of matching stimuli as matches at Time 1 and changed their 
decision to non-matching or inconclusive or b) correctly judged a pair of non-matching 
stimuli as non-matches at Time 1 and changed their decision to matching or inconclusive.

CHAPTER VII

MAIN STUDY RESULTS

Time 1 Judgements

Separate logistic regressions were conducted to examine the effects of contextual 
information condition and examiner experience on the likelihood of making a correct 
fingerprint ID, correct footwear impression ID, incorrect fingerprint exclusion, incorrect 
footwear impression exclusion, correct fingerprint exclusion, correct footwear impression 
exclusion, false fingerprint ID, and false footwear impression ID at Time 1. In addition, a 
mixed analysis of variance (ANOVA) was conducted to examine the effect of contextual 
information condition, examiner experience, stimulus material, and stimulus pair on the 
number of inconclusive judgments made at Time 1. Contextual information was dummy 
coded for logistic regression analyses. First, the blind LAR was selected as the reference 
group in order to obtain the following comparisons: blind LAR versus non-blind LAR 
and blind LAR versus partially redacted LAR. Next, the partially redacted LAR was 
selected as the reference group in order to obtain the last comparison between partially 
redacted LAR and non-blind LAR.

Correct fingerprint ID. A logistic regression was conducted to examine the 
effect of contextual information condition and examiner experience on the likelihood of 
making a correct fingerprint ID (see Table 1). There was a significant effect of
experience, \( B = 2.50, \) Wald = 16.93, \( p < .001, \) OR = 12.15, 95% CI for OR = 3.70 to 39.90. Experts were more than 12 times more likely to make a correct fingerprint ID than novice examiners (see Table 2). However, there was no significant effect of contextual information on the likelihood of making a correct fingerprint ID such that those who received the blind LAR did not significantly differ from those who received the non-blind LAR or from examiners who received a redacted LAR. In addition, participants who received the non-blind LAR did not significantly differ from those who received the redacted LAR. There was also no significant interaction between experience and contextual information.

**Correct footwear impression ID.** A logistic regression was conducted to examine the effect of contextual information and examiner experience on the likelihood of making a correct footwear impression ID (see Table 3). There was a significant effect of experience, \( B = -1.88, \) Wald = 19.51, \( p < .001, \) OR = 0.15, 95% CI for OR = 0.07 to 0.35 such that, interestingly, novices were approximately 7 times more likely to make a correct footwear impression ID than expert examiners (see Table 2). There was no significant effect of contextual information on the likelihood of making a correct footwear impression ID such that the different LAR conditions did not significantly differ from one another. There was also no significant interaction between experience and contextual information.

**Incorrect fingerprint exclusion.** A logistic regression was conducted to examine the effect of contextual information and examiner experience on the likelihood of making an incorrect fingerprint exclusion (see Table 4). There was a significant effect of experience, \( B = -2.82, \) Wald = 7.65, \( p = .006, \) OR = 0.06, 95% CI for OR = 0.01 to 0.44.
Novices were approximately 17 times more likely than expert examiners to make an incorrect fingerprint exclusion (see Table 2). Again, there was no significant effect of contextual information such that regardless of the LAR the examiners received, there were no differences in their likelihood of making an incorrect fingerprint exclusion. There was also no significant interaction between experience and contextual information.

**Incorrect footwear impression exclusion.** A logistic regression was conducted to examine the effects of contextual information and examiner experience on the likelihood of making an incorrect footwear impression exclusion (see Table 5). There was no significant effect of experience meaning experts and novices were equally likely to make an incorrect footwear impression exclusion. There was also no significant effect of contextual information such that examiners in the different LAR conditions did not significantly differ from one another. There was also no significant interaction between experience and contextual information.

**Correct fingerprint exclusion.** A logistic regression was conducted to examine the effect of contextual information and examiner experience on the likelihood of making a correct fingerprint exclusion (see Table 6). There was a significant effect of experience, \( B = 2.65, \text{Wald} = 19.17, p < .001, OR = 14.19, 95\% \text{ CI for } OR = 4.33 \text{ to } 46.55 \) such that expert examiners were approximately 14 times more likely to make a correct fingerprint exclusion than novice examiners (see Table 2). There was no significant effect of contextual information on the likelihood of making a correct fingerprint exclusion such that examiners in the different LAR conditions did not significantly differ from one another. There was also no significant interaction between experience and contextual information.
Correct footwear impression exclusion. A logistic regression was conducted to examine the effect of contextual information and examiner experience on the likelihood of making a correct footwear impression exclusion (see Table 7). There was no significant effect of experience or of contextual information. Specifically, experts did not differ from novices and those who received the blind LAR did not significantly differ from those who received the non-blind LAR or from the redacted LAR. Those who received the non-blind LAR also did not significantly differ from those who received the redacted LAR. However, there was a significant interaction between experience and contextual information, $B = -0.78$, Wald = 4.12, $p = .043$, $OR = 0.46$, 95% CI for $OR = 0.22$ to 0.98. Follow up analyses found that only for expert examiners, did receiving a blind LAR result in a marginally significant increase in the likelihood of making a correct footwear impression exclusion, $B = 1.31$, Wald = 3.67, $p = .056$, $OR = 3.71$, 95% CI for $OR = 0.97$ to 14.23. Specifically, experts who were presented with a blind LAR were 3.71 times more likely to make a correct footwear impression exclusion than experts who were presented with a redacted LAR. However, this finding should be interpreted with extreme caution, as it may be a type 1 error.

False fingerprint ID. A logistic regression was conducted to examine the effects of contextual information and examiner experience on the likelihood of making a false fingerprint ID (see Table 8). There was no significant effect of experience or of contextual information on the likelihood of making a false fingerprint ID. That is, experts and novices did not differ from one another and examiners in the different LAR conditions also did not significantly differ from one another. There was also no significant interaction between experience and contextual information.
**False footwear impression ID.** A logistic regression was conducted to examine the effect of contextual information and examiner experience on the likelihood of making a false footwear impression ID (see Table 9). There was a significant effect of experience, $B = -2.87$, Wald = 7.92, $p = .005$, $OR = 0.06$, 95% CI for $OR = 0.01$ to 0.42. Novice examiners were approximately 17 times more likely to make a false footwear impression ID than expert examiners (see Table 2). However, there was no significant effect of contextual information on the likelihood of making a false footwear impression ID. There was also no significant interaction between experience and contextual information.

**Inconclusive judgments.** Regardless of the stimulus material individuals judged (i.e., fingerprints or footwear impressions) or whether each pair was a match or non-match, participants’ were able to make an “inconclusive” judgement. Thus, a participant could have made a total of one, two, three or four inconclusive judgments. A mixed ANOVA was conducted to examine the effect of contextual information, examiner experience, stimulus material, and stimulus pair on the total number of inconclusive judgments made (see Table 10).

There was a significant effect of stimulus material, $F(1, 336) = 9.50, p = .002, \eta_p^2 = .027$. Examiners made more inconclusive judgments when judging footwear impressions ($M = 0.50$, $SE = 0.04$) than when judging fingerprints ($M = 0.33$, $SE = 0.04$). However, this main effect was qualified by a significant interaction between examiner experience and stimulus material, $F(1, 336) = 60.71, p < .001, \eta_p^2 = .153$, for the number of inconclusive decisions made. Follow up independent sample t-tests revealed that experience had a different effect on the number of inconclusive decisions made
depending on whether the evidence analyzed was a fingerprint or a footwear impression (see Figure 3). Specifically, experts made significantly more inconclusive decisions \((M = 0.71, SD = 0.63)\) than novices \((M = 0.30, SD = 0.49)\) when judging footwear impression evidence, \(t(67.04) = -4.57, p < .001, d = 0.73, 95\% \text{ CI} (-0.59, -0.23)\), but experts made significantly fewer inconclusive decisions \((M = 0.09, SD = 0.29)\) than novice examiners \((M = 0.56, SD = 0.63)\) when judging fingerprint evidence, \(t(168.57) = 8.79, p < .001, d = 0.96, 95\% \text{ CI} (0.37, 0.58)\).

There was also a significant effect of stimulus pair, \(F(1, 336) = 48.72, p < .001, \eta^2_p = .127\), such that participants made more inconclusive decisions when judging non-matching pairs of stimuli \((M = 0.61, SE = 0.05)\) than when judging matching pairs of evidence \((M = 0.22, SE = 0.03)\). There was a significant interaction between stimulus pair and contextual information, \(F(1, 336) = 3.93, p = .048, \eta^2_p = .012\). However this two-way interaction was qualified by a three-way interaction between examiner experience, contextual information, and stimulus pair on the number of inconclusive judgements made, \(F(1, 336) = 4.02, p = .019, \eta^2_p = .023\). Follow-up independent sample t-tests revealed that only when examiners were provided with a Blind LAR, did novice examiners make marginally significantly more inconclusive judgements \((M = 0.74, SD = 0.74)\) compared to expert examiners \((M = 0.37, SD = 0.50)\), \(t(107) = 1.92, p = .056, d = 0.59, 95\% \text{ CI} (-0.01, 0.75)\) when judging non-matching pairs of evidence. There were no other significant interactions.

**Accuracy rate.** A mixed ANOVA was conducted to examine the effects of contextual information, examiner experience, stimulus material, and stimulus pair on examiners’ accuracy rate (see Table 11). As mentioned above, accuracy rate was
calculated by dividing the total number of correct decisions made, by the number of decisions made in total. There was a significant main effect of examiner experience, $F(1, 336) = 20.82, p < .001, \eta^2_p = .058$, such that expert examiners were significantly more accurate ($M = 0.79, SE = 0.03$) than novice examiners ($M = 0.64, SE = 0.01$). However, this main effect was qualified by a significant interaction between examiner experience and stimulus material, $F(1, 336) = 48.20, p < .001, \eta^2_p = 0.125$ (see Figure 4). Follow up independent sample t-tests revealed that examiner experience only had a significant effect on fingerprint accuracy rates, $t(182.88) = -12.44, p < .001, d = 1.36, 95\% CI (-0.43, -0.31)$ and not on footwear impression accuracy rates. Specifically, experts were significantly more accurate ($M = 0.95, SD = 0.16$) when judging fingerprint pairs than novice examiners ($M = 0.57, SD = 0.36$). There was also a significant effect of stimulus pair, $F(1, 336) = 35.28, p < .001, \eta^2_p = .095$. Both expert and novice examiners were significantly more accurate when judging matching pairs of stimuli ($M = 0.81, SE = 0.02$) than when judging non-matching pairs of stimuli ($M = 0.62, SE = 0.03$). There were no other significant main effects or interactions.

**Signal Detection Measures**

**Discrimination accuracy.** A mixed ANOVA was conducted to examine the effects of contextual information, examiner experience, and stimulus material on $d'$ (i.e., discrimination accuracy) (see Table 12). A discrimination accuracy $d'$ value may range from $+\infty$ to $-\infty$. A score of zero indicates an inability to distinguish between matches and non-matches, whereas larger values indicate a greater ability to distinguish matches from non-matches.
There was a significant main effect of examiner experience, $F(1, 335) = 8.14$, $p = .005$, $\eta^2_p = .024$, such that experts were better able to discriminate between matching and non-matching pairs of stimuli ($M = 0.34$, $SE = 0.13$) compared to novices ($M = -0.06$, $SE = 0.06$). There was also a significant main effect of evidence material, $F(1, 335) = 7.59$, $p = .006$, $\eta^2_p = .022$. Both expert and novice examiners were better able to discriminate between matching and non-matching fingerprints ($M = 0.34$, $SE = 0.11$) than matching and non-matching footwear impressions ($M = -0.07$, $SE = 0.10$). There was also a significant interaction between experience and stimulus material, $F(1, 335) = 16.18$, $p < .001$, $\eta^2_p = .046$. Follow up independent t-tests revealed that experience only had a significant effect on fingerprint (and not footwear impression evidence) discrimination accuracy, $t(261.74) = -9.04$, $d = 0.89$, $p < .001$, 95% CI (-1.22, -0.78) such that experts were better able to discriminate between matching and non-matching fingerprints ($M = 0.84$, $SD = 0.48$) compared to novices ($M = -0.16$, $SD = 1.51$; see Figure 5). There were no other significant main effects or interactions.

**Response bias.** A mixed ANOVA was conducted to examine the effects of contextual information, examiner experience, and stimulus material on $c$ (i.e., response bias) (see Table 13). Negative values of $c$ signify a bias towards responding “match”, whereas positive values signify a bias towards the “non-match” response. The neutral point of $c$ is 0, which indicates a neutral judgment tendency such that there is no greater tendency to respond with a match or non-match response.

There was a significant main effect of examiner experience, $F(1, 335) = 7.30$, $p = .007$, $\eta^2_p = .021$. Expert examiners were more likely to make a non-match decision regardless of whether the prints were matches or non-matches ($M = 0.18$, $SE = 0.07$)
compared to novices ($M = -0.03, SE = 0.03$). There was also a significant main effect of evidence material, $F(1, 335) = 19.55, p < .001, \eta^2_p = .055$. Examiners were more likely to make a non-match decision when judging footwear impressions ($M = 0.23, SE = 0.05$) compared to when judging fingerprints ($M = -0.08, SE = 0.05$). There was also a significant interaction between examiner experience and stimulus material, $F(1, 335) = 41.50, p < .001, \eta^2_p = .110$. Follow up independent t-tests revealed that examiner experience had a significant effect on response bias when judging fingerprint evidence, $t(257.65) = 4.32, p < .001, d = 0.42, 95\% CI (0.13, 0.34)$ such that experts appeared to have a “match” bias when judging fingerprints ($M = -0.19, SD = 0.24$) whereas novices demonstrated more of a “non-match” bias ($M = 0.04, SD = 0.74$). However, opposite findings were found when judging footwear impression evidence, $t(68.86) = -5.63, p < .001, d = 0.87, 95\% CI (-0.87, -0.42)$, such that experts tended to have a “non-match” bias ($M = 0.54, SD = 0.79$) whereas novices showed more of a “match” bias ($M = -0.10, SD = 0.68$). There were no other significant main effects or interactions.

Confidence Ratings

A mixed ANOVA was conducted to examine the effects of LAR condition, examiner experience, stimulus material, and stimulus pair on examiners’ confidence ratings (see Table 14). There was a significant effect of examiner experience, $F(1, 336) = 5.35, p = .021, \eta^2_p = 0.016$, such that experts reported feeling more confident in their judgements ($M = 5.75, SE = 0.13$) than novices ($M = 5.43, SE = 0.06$). There was also a significant effect of stimulus material, $F(1, 336) = 16.15, p < .001, \eta^2_p = .046$. Examiners reported feeling more confident in their fingerprint judgements ($M = 5.78, SE = 0.09$) than in their footwear impression judgments ($M = 5.40, SE = 0.08$). However, these main
effects were qualified by a significant interaction between examiner experience and stimulus material, $F(1, 336) = 152.05, p < .001, \eta^2_p = 0.312$ (see Figure 6). Specifically, experts reported feeling significantly more confident ($M = 6.53, SD = 0.75$) than novices ($M = 5.03, SD = 1.28$) when judging fingerprints $t(123.39) = -11.81, p < .001, d = 1.43, 95\% \text{ CI} (-1.75, -1.25)$, but reported feeling significantly less confident ($M = 4.96, SD = 1.38$) than novices ($M = 5.83, SD = 0.97$) when judging footwear impression evidence, $t(64.57) = 4.42, p < .001, d = 0.73, 95\% \text{ CI} = 0.47, 1.25$.

In addition, there was a significant main effect of stimulus pair, $F(1, 336) = 27.17, p < .001, \eta^2_p = 0.075$, such that examiners reported feeling more confident when judging matches ($M = 5.80, SE = 0.07$) than when judging non-matches ($M = 5.38, SE = 0.09$). This main effect was qualified by a significant interaction between examiner experience and stimulus pair, $F(1, 336) = 3.96, p = .047, \eta^2_p = 0.012$. Specifically, examiner experience only had a significant effect on examiners’ confidence ratings when analyzing non-matching pairs of evidence, $t(340) = -2.75, p = .006, d = 0.41, 95\% \text{ CI} (-0.40, -0.07)$, such that expert examiners felt significantly more confident when judging non-matching pairs of evidence ($M = 5.61, SD = 1.06$) than novices ($M = 5.14, SD = 1.19$). There were no other significant main effects or interactions.

Finally, bivariate correlations were conducted to examine the relationship between individuals’ confidence ratings and their accuracy rate. Analyses revealed a significant and positive correlation between accuracy rate and confidence, $r = .150, p = .005$, such that the more confident individuals felt in their judgments, the more likely they were to be accurate. Additional exploratory analyses revealed that the positive correlation between accuracy and confidence was present regardless of the type of stimuli material.
and stimuli pair examiners analyzed (see Table 15). However, there was only a significant accuracy-confidence correlation present for participants who viewed the non-blind LAR, \( r = .246, p = .007 \), and not for those who were given a blind LAR, \( r = .088, p = .364 \), or a redacted LAR, \( r = .093, p = .322 \). In addition, there was a marginally positive correlation between accuracy and confidence for novice examiners, \( r = .114, p = .053 \), but not for expert examiners, \( r = .200, p = .143 \).

**Time 2 Judgments**

**Decision to review time 1 judgments.** A logistic regression was conducted to examine the effect contextual information and examiner experience had on examiner’s decision to go back and review their Time 1 judgments after being presented with an LAR that had all case summary details (see Table 16). There was only a significant effect of examiner experience, \( B = -2.18, \text{ Wald } = 12.95, p < .001, OR = 0.11, 95\% \text{ CI for } OR = 0.03 \text{ to } 0.37 \). Novices were more than 9 times more likely than expert examiners to decide that they wanted to go back to their original responses and make changes. There were no other significant main effects or interactions. Specifically, only three forensic experts decided to go back to review their Time 1 judgments, whereas 97 novice examiners decided to go back to their Time 1 judgments.

**Correct and incorrect changes made.** Of those participants who decided to go back to their original responses and make changes (\( n = 100, 29.2\% \)), a total of 17% made at least one correct change to their initial response. However, 14% of participants who decided to go back to their responses after being presented with the non-blind LAR were incorrect in at least one of the changes they made. The remaining participants (69%), which included the three expert examiners who decided to review their Time 1
judgments, did not decide to change any of their initial decisions after being reminded of their initial judgments.

Incorrect match decisions made. Of particular importance to the current study are the incorrect changes examiners made when they changed their initial correct non-match decision or inconclusive decision to an incorrect match decision after being exposed to new biasing information (i.e., suspects are in custody) as it suggests they were influenced by such information. Overall, 11% of participants who decided to change their initial decision made an incorrect match decision after being exposed to additional case summary details.

Perceptions Questionnaire

Examiners were provided with various questions regarding how well they performed on the given tasks and their opinions regarding how guilty they found each suspect to be and whether they felt the initial LAR they received provided them with enough information to accurately analyze the evidence. In order to better understand how contextual information and experience can affect such beliefs, separate univariate ANOVAs were conducted on participants’ self-reported accuracy and each suspect’s guilt ratings. Furthermore, a logistic regression was conducted on examiners’ belief on whether or not the initial LAR they received provided them with sufficient information.

Self-reported accuracy. Overall, there was a significant effect of examiner experience on participants’ self-reported accuracy, $F(1, 336) = 60.72, p < .001, \eta_p^2 = .153$. Experts reported being more accurate in their judgments ($M = 91.12, SE = 2.94$) than novice examiners ($M = 66.10, SD = 1.28$). However, there was no significant effect of contextual information, nor was the interaction significant.
**Guilt ratings.** A similar pattern of findings was found for examiners’ guilt ratings for both suspects. Specifically, there was only a significant effect of examiner experience, such that expert examiners reported lower guilt ratings for suspect #1 ($M = 5.32, SE = 0.29$) than novices ($M = 7.03, SE = 0.13$) and for suspect #2 ($M = 5.21, SE = 0.27, M = 7.39, SE = 0.16$, respectively). This was supported by expert examiners’ comments regarding their beliefs about the suspects’ guilt. Many expert examiners explicitly stated it was not their place to determine whether a suspect was guilty or innocent.

**Perceptions of information available.** There was a significant main effect of examiner experience on examiners’ belief regarding whether they had sufficient information on their initial LAR to make accurate decisions about the various pieces of evidence, $B = 2.19$, Wald $= 16.77$, $p < .001$, $OR = 8.97$, 95% CI for $OR = 3.14$ to 25.64. Experts were approximately nine times more likely to believe they had sufficient information on their LAR compared to novice examiners. There was also a significant effect of contextual information, such that those examiners who received the non-blind LAR were more likely to believe they had sufficient information to make accurate decisions, compared to examiners who received a blind LAR, $B = 0.64$, Wald $= 4.68$, $p = .030$, $OR = 1.91$, 95% CI for $OR = 1.06$ to 3.41 and also compared to those who received a partially redacted LAR, $B = 0.82$, Wald $= 7.80$, $p = .005$, $OR = 2.27$, 95% CI for $OR = 1.28$ to 4.04. However, there was no significant differences between those who received a blind LAR and a partially redacted LAR, $B = 0.18$, Wald $= 0.38$, $p = .537$, $OR = 1.19$, 95% CI for $OR = 0.68$ to 2.09. There was no significant interaction between examiner experience and contextual information.
CHAPTER VIII
DISCUSSION

The present study was the first to examine the effects of contextual bias, examiner experience, and stimulus material on examiners’ analyses of forensic evidence. Specifically, the study examined how differential exposure to case details about a crime can affect examiners’ analyses of forensic evidence. In doing so, we were able to test the efficacy of blind testing and linear sequential unmasking methods (i.e., redacting task-irrelevant information), which are two of the most commonly recommended bias-reducing procedures (Kassin et al., 2013; Saks et al., 2003). In addition, the effect of examiner experience and stimulus material on examiners’ analyses were investigated to address some of the concerns voiced by practitioners regarding the limitations of existing research, including inexperienced student participants analyzing forensic evidence samples that require specialized training.

Contextual Information

The first aim of the present study was to examine the effects of contextual bias on examiners’ evaluation of forensic evidence by varying the amount of pre-comparison information available to participants while simultaneously testing the efficacy of various bias-reducing procedures. We hypothesized there would be an effect of contextual information such that examiners who received a complete, non-blind LAR prior to conducting analyses and thus had biasing information (e.g., suspects’ criminal history and custody status) would be more likely to make a match decision (correct and incorrect) compared to examiners who received the blind or partially redacted LAR. Furthermore, it was predicted that examiners who received the partially redacted LAR would be more
likely to make a correct ID and less likely to make a false ID and thus have better
discrimination accuracy than examiners who received the blind LAR. However, results
did not support our hypotheses regarding contextual bias. Instead, findings revealed no
differences between the three contextual information conditions on examiners’ judgments
and thus no differences in their overall accuracy, discrimination accuracy, or response
tendencies. As such, the current findings do not support that recommended blinding and
redaction procedures (Kassin, Dror, Kukucka, 2013) can reduce the broad effect of bias
on examiners’ forensic analyses under the conditions measured, but this may be because
there was no effect of bias found.

Nonetheless, the present findings mirror those found by Kerstholt and colleagues
(2007 & 2010) who also did not find an effect of contextual bias. Kerstholt and
colleagues (2007) included real-world police officers whom had been trained on footwear
impression analyses whereas Kertsholt and colleagues (2010) included forensic firearm
 examiners as participants. Specifically, Kerstholet and colleagues’ (2007) found that
providing police officers with additional guilt presumptive contextual information about
suspects in a hypothetical case did not affect their analyses of footwear impression
evidence. Kerstholt and colleagues’ (2010) also found that providing firearm examiners
with either neutral or biasing information about the number of individuals implicated in
the crime did not affect examiners’ decision regarding whether a set of bullets came from
the same or different firearm. Similarly, present results found that telling real-world
forensic examiners and student participants that the suspects were in custody and
providing them with information about the suspects’ criminal history did not bias their
analyses.
Differences in contextual bias across studies may be explained by differences in the types of biasing details available to examiners. Examiners may be affected less by peripheral information including a suspect’s criminal history or custody status and more influenced by information that more directly incriminates the suspect. For example, as previous research has found, telling forensic examiners that the suspect confessed, more directly associates a particular suspect with the crime being investigated and may thus more easily affect a forensic examiner’s analyses (Kukucka & Kassin, 2014). In this case, using a blinding procedure or redaction methods may be useful in decreasing the negative effect of contextual bias – an important notion for research moving forward.

Another possible explanation why the contextual information, regardless of its availability and content, did not impact examiners’ judgments of the evidence is that examiners felt particularly inclined to perform well and treat the analyses as they would in a real-world case because of the manner in which the information was relayed to them. Specifically, in the current study examiners were presented with a realistic LAR that mimicked the form used by the Miami Dade Forensic Services Bureau. Using these evidence submission forms to manipulate the availability of contextual information differs from past research, in which examiners are generally provided with a case summary in the form of a simple paragraph – outside the context of a LAR. Providing examiners with an official looking document may have led them to assume the role of a forensic examiner working a real case and thus pay closer attention to study materials and procedures, decreasing any potential effect of bias when the task-irrelevant information was available.
Moreover, it was hypothesized that due to the lack of information, participants in the blind condition would be less confident in their analyses compared to participants who received either all the information about the case or a partially redacted LAR. However, there were no differences in examiners’ confidence ratings depending on the contextual information condition they were in. This may be because examiners did not base their confidence ratings on the amount and type of information available to them on the LAR, but instead on other factors such as their experience with analyzing different types of forensic evidence. Indeed, the present results found a significant interaction between examiner experience and stimulus material, such that expert fingerprints examiners felt significantly more confident in their fingerprint judgments than novices. Interestingly, novices felt significantly more confident in their footwear impression judgments compared to experts. This may be because expert examiners recruited for this study were primarily expert latent fingerprint experts who may have had only a small amount of experience analyzing footwear impressions and were aware of this differential experience (unlike student participants who were inexperienced in examining both type of stimuli). It appears that expert fingerprint examiners adjusted their confidence judgments accordingly depending on the evidence they were analyzing such that they were more confident conducting analyses they had more experience with (i.e., fingerprints) and less confident when judging evidence they were not too familiar with (i.e., footwear impressions).

Exploratory analyses also revealed that examiners who received the non-blind LAR were more likely to believe they had sufficient information to conduct accurate analyses compared to examiners who received a blind or partially redacted LAR.
Interestingly, there were no differences in accuracy rates among the various conditions, which indicates that although examiners may feel they do not have sufficient information about the case when given blind or partially redacted LARs to perform well, they are still capable of performing just as well as those who do have all case details.

**Examiner Experience**

The majority of past research examining the effects of contextual bias on forensic analyses have used undergraduate, untrained student participants (Dror et al., 2005; Smalarz et al., 2016). The use of undergraduate students is one of the most common critiques practitioners in the field have of past research and they often struggle to accept findings from such studies, stating they cannot be generalized to experts in the field (Ostrom, 2009). For this reason, the second primary aim of the present study was to directly compare novice and expert examiners’ vulnerabilities to contextual bias and their ability to conduct forensic analyses. Overall, it was hypothesized that novices would make more mistakes than expert examiners (i.e., less likely to make correct IDs and correct exclusions and more likely to make false IDs and incorrect exclusions) and have poorer discrimination accuracy because of their lack of experience in analyzing fingerprint and footwear impression evidence. Findings of the present study supported this hypothesis such that novices tended to perform worse than expert forensic examiners. For example, novices were less likely to make correct fingerprint and footwear impression IDs, more likely to make false footwear impression IDs, and had worse overall accuracy and discrimination accuracy compared to expert examiners.

Furthermore, an interaction between examiner experience and contextual information was predicted such that task-irrelevant information was believed to influence
novice examiners more so than expert examiners. On the basis of past research suggesting that differences in examiners’ susceptibility to bias can depend on the level of experience with conducting fingerprint evidence analyses (Langenburg et al., 2009), examiners with no experience analyzing forensic evidence were expected to rely more on the contextual information when available (i.e., non-blind LAR condition) compared to expert examiners. That is, novice examiners would be more influenced than experts by task-irrelevant information because of the increased unfamiliarity and difficulty of the task at hand. However, an interaction between examiner experience and contextual information was not found. This lack of predicted finding may also be explained in line with earlier discussing of the role of stimulus material in this effect: The type of task-irrelevant and biasing information provided to examiners may not have been sufficiently biasing to allow examiners to rely on such information when it was available even when the examiner had no experience in conducting forensic analyses. In the context of real-world forensic examiners’ criticism of prior research however, this lack of interaction cautiously suggests that using lay examiners may not necessarily overestimate the potential for cognitive bias per se as student examiners did not showcase a heightened vulnerability to prior case information compared to expert examiners. However, given that there was no main effect of prior case information contained in LARs, future research needs to urgently address potential differences in vulnerability to cognitive bias in forensic examination between experts and novices when such bias does exist.

Although an interaction between examiner experience and contextual information was not supported by the present findings, there was a significant interaction between examiner experience and stimulus material, as predicted. Experts were expected to
perform significantly better than novices, but only when analyzing fingerprint evidence because it requires more specialized training than footwear impression evidence. This hypothesis was supported in that expert examiners performed significantly better than novice examiners, but only when analyzing fingerprint evidence and not footwear impression evidence. These findings suggest past research using forensic stimulus materials that require specialized training and recruiting untrained participant examiners (e.g., Dror et al., 2005) may be underestimating the accuracy of forensic examiners in the field and future research should only use undergraduate student samples if the stimulus material being analyzed has been extensively pilot-tested and does not require on-the-job specialized training (e.g., footwear impression evidence). In addition, an interaction between examiner experience and stimulus material on participants’ response bias was found such that experts were more likely to make a “match” decision when analyzing fingerprint evidence compared to novices. Interestingly however, expert examiners were more likely to make a “non-match” decision when analyzing footwear impressions compared to novices. It appears that differences in fingerprint analyses experience between our expert and novice participants led to differences in reporting tendencies. It may be that although footwear impression analyses requires less specialized training than fingerprint analyses and both are a form of pattern analyses, simply being trained in one does not make you proficient in the analyses of the other and may in fact hinder your abilities to analyze other forms of evidence. As such, future work should more finely examine the similarities and differences between forensic evidence trainings and investigate how these differences affect examiners’ judgments across a variety of forensic domains.
Stimulus Material

The third and final aim of the present study was to investigate the effect of contextual bias on examiners’ evaluation of fingerprint and footwear impression evidence, with fingerprint examination requiring significantly more expertise and training than footwear impression analyses (International Association for Identification, 2018; Sirchie, 2018). It was hypothesized that contextual bias would have an effect on examiners’ judgments when the evidence being analyzed required significantly more specialized training (fingerprints). However, present findings did not support the interaction between contextual bias and evidence type. Regardless of whether examiners were provided with a blind, non-blind or partially redacted LAR there were no differences in examiners’ performance when analyzing fingerprint evidence.

The present findings do not suggest fingerprint evidence analyses are particularly susceptible to the effects of contextual information compared to analyses of footwear impression evidence, which is at odds with the general belief that evidence difficulty and ambiguity moderate the effect of contextual bias (Dror et al., 2005). However, it is important to note that not all past research has found a moderating effect of evidence difficulty and ambiguity (with ambiguous materials being more difficult to analyze) (Kerstholt, et al., 2007). That is, some research has failed to find an effect of contextual information even with particularly difficult analyses. For example, Kerstholt and colleagues (2007) did not find an effect of bias on footwear impression analyses even when the footwear impressions were made to be ambiguous. It was initially hypothesized that these discrepant results from past research were due to differences in the evidence material used such that Dror and colleagues (2005) used fingerprints while Kerstholt and
colleagues (2007) used footwear impressions. Surprisingly, our results did not provide support for this hypothesis.

We can only speculate as to why contextual information did not have an effect on neither examiners’ fingerprint nor footwear impression analyses. Contextual information may not have influenced examiners’ analyses when judging footwear impression evidence because the footwear impression pairs presented to participants were of comparable difficulty to the fingerprint pairs. For example, on average, the difficulty rating of the footwear impression stimuli used for the current study was 2.56, and the average difficulty for the fingerprint stimuli was 4.06 (rated on a scale from 1 being not at all difficult to 7 being extremely difficult). It is possible that if fingerprints with a higher difficulty level (i.e., had been selected from the ‘insufficient to difficult’ category) and footwear impressions from the ‘moderate to easy’ category had been chosen, an effect of contextual information would have been found for fingerprints and not footwear impressions. Nonetheless, discrepancies between current results and existing research warrant a more precise examination of how evidence type (e.g., evidence complexity, ambiguity, difficulty) plays a role in the relationship between contextual bias and examiners’ forensic analyses.

**Testing Linear Sequential Unmasking Procedures**

Exploratory analyses were also conducted to examine the effect of presenting participants with a non-blind LAR (regardless of the initial LAR they had received) and asking them whether they would like to change any of their initial judgments, mirroring recommended linear sequential unmasking procedures. Although only 100 participants decided to go back to their original responses, 17% of those who decided to change at
least one of their Time 1 judgments made a correct change. However, 14% of participants who decided to go back to their Time 1 judgments made an incorrect change. It appears that the linear sequential unmasking procedure may result in both increases in correct and incorrect changes at Time 2. However, only novice examiners actually made changes at their Time 2 judgments, whereas no expert examiners did, questioning further the actual usefulness of the proposed procedure and also the usefulness of even providing the task-relevant information in the first place.

It is important to note however that because the present study is the first to experimentally examine the linear sequential unmasking procedure, it provided a very simplified manner in which additional case details were sequentially presented to examiners. Specifically, examiners who were initially given a blind or partially-redacted LAR later received all case details that were previously unavailable all at once in the second LAR and not one detail at a time. It may be that presenting examiners with one detail at a time, may result in fewer errors. Future research should investigate systematically how the level of bias and importance in any additional detail affect the potential for bias.

**Limitations and Future Directions**

The current study sought to examine differences between experts and novices and in doing so required study methods to be employed easily to practitioners in the field. Thus, the study was distributed using Qualtrics survey system, which - similar to past research in the field - did not allow for the implementation of tools commonly used by practitioners to analyze forensic fingerprint evidence (e.g., CSIpix and PiAnoS). In order to address this limitation, only ‘moderate to easy’ difficulty level fingerprint comparisons
were presented to examiners. Considering fingerprint stimuli chosen for the current study were rated by experts to be relatively “easy” comparisons – unlike any other prior study - the degree to which bias could have impacted examiners’ judgments may have been reduced. However, undergraduate students rated the same fingerprints as being more difficult, yet an effect of bias was still not found. Nonetheless, future research should aim at using more difficult stimuli and allow examiners to use the programs and analysis software commonly found in the field. Additionally, due to the extreme difficulty in recruiting expert examiners, the current study may be underpowered as only 55 expert examiners participated versus 287 novice examiners. Thus, there is the possibility that an interaction between examiner experience and contextual bias may still be found if additional expert examiners were to be recruited. Nonetheless, the current study’s sample is significantly larger than previous studies that have also aimed to recruit experts in the field (Dror & Charlton, 2006; Dror, Charlton & Peron 2005) and provides a positive outlook for future research. Future research should continue to engage in collaborative research processes that involve practitioners in order to build trusting relationships, which subsequently results in more practitioners agreeing to participate in research.

Another potential limitation of the current study is that examiners were presented only with an LAR and no other case materials. Forensic examiners are sometimes provided with additional materials such as witness statements, pathology reports and other types of contextual information, which were not used in the current study. Nonetheless, the current study improves upon previous research such that examiners were presented with an evidence submission form that mirrors those frequently found in the field and used by forensic labs across the United States (Gardener et al., 2019) instead of
a case summary in the form of a written paragraph, which is unlikely to mirror actual lab practices. Future research should continue to expand on the generalizability of materials used in experimental research by providing examiners with additional case materials that are commonly provided to forensic examiners.

Furthermore, research examining the effects of contextual bias and testing the efficacy of proposed procedures used to reduce the effect of bias should move beyond basic laboratory paradigms and into field settings. Although it is difficult to maintain the same experimental control found in laboratory studies, it is crucial to test contextual information management procedures (e.g., context-manager method) with practitioners in order to properly understand how they can be implemented in a forensic laboratory’s standard operating procedures without disrupting the existing workflow. In addition, testing bias-reducing procedures in the field allows researchers and practitioners to determine what additional resources, if any, are necessary to successfully implement such procedures. For example, Found and Ganas (2013) tested a contextual information management procedure in a field setting and were able to determine that no additional personnel, time, or money was necessary to implement such a procedure for a laboratory-based bloodstain pattern analysis. However, additional field studies, developed in collaboration with forensic experts must be conducted to discover if these same promising findings carry over to other forensic domains such as fingerprints, firearms, and DNA analyses.

**Implications and Conclusions**

Cognitive bias, specifically contextual bias, has become a primary area of concern within the forensic sciences due to high profile exoneration cases highlighting the
detrimental effects that having details about a case can have on a forensic examiner’s analyses (Innocence Project, 2017; National Academy of Sciences, 2009). Fortunately, researchers have proposed several solutions to decrease the negative effects of contextual bias such as the use of blind testing and redacting task-irrelevant information from case reports, both of which were examined in the current study (Kassin et al., 2013; Mattijssen et al., 2016). Although an effect of contextual information was not found, present findings still provide a promising outlook for the use of bias reducing procedures. That is, examiners often express concern about the use of such bias reducing procedures because it limits the amount of information they receive, which they believe may hinder their ability to accurately analyze forensic evidence. However, current results suggest there were no differences in examiners’ overall accuracy and discrimination accuracy rates suggesting that regardless of whether an examiner has all case details, no case details, or only task-relevant case details, an examiner is able to maintain the same level of accuracy.

The present study also sought to address another common critique made by forensic examiners in the field regarding past research - having untrained undergraduate students analyze complex pieces of forensic evidence. By directly comparing expert and novice examiners’ performance in analyzing two types of forensic evidence that differ in level of difficulty (i.e., fingerprint and footwear impression evidence), the current study was able to confirm the importance of using stimuli materials that require less specialized training when using untrained examiners in research. Expert examiners were found to be significantly more accurate than novice examiners. However, expert examiners were only significantly more accurate and better able to discriminate between matching and non-
matching pairs of evidence when analyzing fingerprint evidence and not footwear impression evidence. Thus, previous research using inexperienced student examiners must be interpreted with caution; especially when such novice examiners are asked to analyze evidence that requires a significant amount of training, such as fingerprints.

To summarize, the present work provides promising results for the use of several bias-reducing methods. In legal proceedings, using blind testing and other contextual information management procedures, such as redacting task-irrelevant information, allow forensic expert witnesses to demonstrate to the court that proper steps were taken to reduce the potential for bias in their analyses. In addition, findings of the current study provide an important insight as to what occurs when undergraduate students versus trained forensic examiners are asked to analyze complex pieces of forensic evidence. Mainly, researchers should take careful consideration and work closely with actual forensic experts when developing stimulus materials and deciding what type of participant sample to use when conducting psychological research within forensic science domains.
Table 1.

Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Fingerprint ID

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Fingerprint ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>2.50</td>
<td>0.61</td>
<td>.000*</td>
<td>12.15</td>
<td>3.70 - 39.90</td>
</tr>
<tr>
<td>Blind* v. partially redacted</td>
<td>-0.38</td>
<td>0.29</td>
<td>.195</td>
<td>0.69</td>
<td>0.39 - 1.21</td>
</tr>
<tr>
<td>Blind* v. non-blind</td>
<td>0.02</td>
<td>0.29</td>
<td>.948</td>
<td>1.02</td>
<td>0.57 - 1.81</td>
</tr>
<tr>
<td>Partially redacted* v. non-blind</td>
<td>0.40</td>
<td>0.27</td>
<td>.166</td>
<td>1.49</td>
<td>0.85 - 2.60</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>0.29</td>
<td>0.74</td>
<td>.693</td>
<td>1.34</td>
<td>0.32 - 5.67</td>
</tr>
</tbody>
</table>

Note. * = reference group.
* = significant.
## Table 2.

*Time 1 Judgments by Examiner Experience.*

<table>
<thead>
<tr>
<th>Time 1 Judgments</th>
<th>Examiner Experience</th>
<th>Frequency</th>
<th>Percentage</th>
<th>$M$ ($SD$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fingerprint Correct ID</strong></td>
<td>Expert</td>
<td>52</td>
<td>94.5%</td>
<td>0.95 (0.23)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>170</td>
<td>59.2%</td>
<td>0.59 (0.49)</td>
</tr>
<tr>
<td><strong>Footwear Impression Correct ID</strong></td>
<td>Expert</td>
<td>42</td>
<td>76.4%</td>
<td>0.76 (0.43)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>274</td>
<td>95.5%</td>
<td>0.95 (0.21)</td>
</tr>
<tr>
<td><strong>Incorrect Fingerprint Exclusion</strong></td>
<td>Expert</td>
<td>1</td>
<td>1.8%</td>
<td>0.02 (0.13)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>68</td>
<td>23.7%</td>
<td>0.23 (0.43)</td>
</tr>
<tr>
<td><strong>Incorrect Footwear Impression Exclusion</strong></td>
<td>Expert</td>
<td>1</td>
<td>1.8%</td>
<td>0.02 (0.13)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>9</td>
<td>3.1%</td>
<td>0.03 (0.17)</td>
</tr>
<tr>
<td><strong>Correct Fingerprint Exclusion</strong></td>
<td>Expert</td>
<td>52</td>
<td>94.5%</td>
<td>0.95 (0.23)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>158</td>
<td>55.1%</td>
<td>0.55 (0.50)</td>
</tr>
<tr>
<td><strong>Correct Footwear Impression Exclusion</strong></td>
<td>Expert</td>
<td>27</td>
<td>49.1%</td>
<td>0.49 (0.50)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>135</td>
<td>47%</td>
<td>0.47 (0.50)</td>
</tr>
<tr>
<td><strong>False Fingerprint ID</strong></td>
<td>Expert</td>
<td>0</td>
<td>0%</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>16</td>
<td>5.6%</td>
<td>0.06 (0.23)</td>
</tr>
<tr>
<td><strong>False Footwear Impression Exclusion</strong></td>
<td>Expert</td>
<td>1</td>
<td>1.8%</td>
<td>0.02 (0.13)</td>
</tr>
<tr>
<td></td>
<td>Novice</td>
<td>70</td>
<td>24.4%</td>
<td>0.24 (0.43)</td>
</tr>
</tbody>
</table>
Table 3.

*Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on*  
*Likelihood of Making a Correct Footwear Impression ID*

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>$B$</th>
<th>$SE$</th>
<th>$p$</th>
<th>Exp $(B)$</th>
<th>95% CI for Exp$(B)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Footwear Impression ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>-1.89</td>
<td>0.43</td>
<td>.000*</td>
<td>0.15</td>
<td>0.07 - 0.35</td>
</tr>
<tr>
<td>Blind* v. partially redacted</td>
<td>0.40</td>
<td>0.54</td>
<td>.459</td>
<td>1.49</td>
<td>0.52 - 4.32</td>
</tr>
<tr>
<td>Blind* v. non-blind</td>
<td>0.04</td>
<td>0.50</td>
<td>.933</td>
<td>1.04</td>
<td>0.39 - 2.78</td>
</tr>
<tr>
<td>Partially redacted* v. non-blind</td>
<td>-0.36</td>
<td>0.53</td>
<td>.498</td>
<td>0.70</td>
<td>0.25 - 1.97</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>0.74</td>
<td>0.56</td>
<td>.187</td>
<td>2.10</td>
<td>0.70 - 6.31</td>
</tr>
</tbody>
</table>

*Note.* * = reference group.  
* = significant.
Table 4.

*Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making an Incorrect Fingerprint Exclusion*

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp (B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect Fingerprint Exclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>-2.82</td>
<td>1.02</td>
<td>0.06*</td>
<td>0.06</td>
<td>0.01 - 0.44</td>
</tr>
<tr>
<td>Blind vs. partially redacted</td>
<td>0.03</td>
<td>0.34</td>
<td>0.940</td>
<td>1.03</td>
<td>0.53 - 1.98</td>
</tr>
<tr>
<td>Blind vs. non-blind</td>
<td>-0.14</td>
<td>0.34</td>
<td>0.677</td>
<td>0.87</td>
<td>0.45 - 1.69</td>
</tr>
<tr>
<td>Partially redacted vs. non-blind</td>
<td>-0.17</td>
<td>0.34</td>
<td>0.620</td>
<td>0.85</td>
<td>0.44 - 1.64</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>-0.10</td>
<td>1.25</td>
<td>0.935</td>
<td>0.90</td>
<td>0.08 - 10.37</td>
</tr>
</tbody>
</table>

*Note.*  
* = reference group.  
* = significant.
### Table 5.

**Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making an Incorrect Footwear Impression Exclusion**

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp (B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect Footwear Impression Exclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>-0.56</td>
<td>1.07</td>
<td>.598</td>
<td>0.57</td>
<td>0.07 - 4.61</td>
</tr>
<tr>
<td>Blind v. partially redacted</td>
<td>0.49</td>
<td>0.74</td>
<td>.508</td>
<td>1.64</td>
<td>0.38 - 7.03</td>
</tr>
<tr>
<td>Blind v. non-blind</td>
<td>-0.50</td>
<td>0.92</td>
<td>.592</td>
<td>0.61</td>
<td>0.10 - 3.72</td>
</tr>
<tr>
<td>Partially redacted v. non-blind</td>
<td>-0.99</td>
<td>0.85</td>
<td>.244</td>
<td>0.37</td>
<td>0.07 - 1.96</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>1.38</td>
<td>0.03</td>
<td>.997</td>
<td>0.90</td>
<td>0.00 - 0.01</td>
</tr>
</tbody>
</table>

*Note.* † = reference group.

* = significant.
Table 6.

**Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Fingerprint Exclusion**

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Fingerprint Exclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>2.65</td>
<td>0.61</td>
<td>.000*</td>
<td>14.19</td>
<td>4.33 - 46.55</td>
</tr>
<tr>
<td>Blind\textsuperscript{r} v. partially redacted</td>
<td>0.29</td>
<td>0.29</td>
<td>.311</td>
<td>1.34</td>
<td>0.76 - 2.37</td>
</tr>
<tr>
<td>Blind\textsuperscript{r} v. non-blind</td>
<td>-0.29</td>
<td>0.29</td>
<td>.308</td>
<td>0.75</td>
<td>0.43 - 1.31</td>
</tr>
<tr>
<td>Partially redacted\textsuperscript{r} v. non-blind</td>
<td>-0.99</td>
<td>0.85</td>
<td>.244</td>
<td>0.37</td>
<td>0.07 - 1.96</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>-0.06</td>
<td>0.78</td>
<td>.940</td>
<td>0.94</td>
<td>0.21 - 4.33</td>
</tr>
</tbody>
</table>

*Note.* \textsuperscript{r} = reference group.  
* = significant.
Table 7.

*Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a Correct Footwear Impression Exclusion*

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp (B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Footwear Impression Exclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>0.08</td>
<td>0.30</td>
<td>.778</td>
<td>1.09</td>
<td>0.61 - 1.94</td>
</tr>
<tr>
<td>Blind$r$ v. partially redacted</td>
<td>0.02</td>
<td>0.27</td>
<td>.941</td>
<td>1.02</td>
<td>0.60 - 1.73</td>
</tr>
<tr>
<td>Blind$r$ v. non-blind</td>
<td>-0.06</td>
<td>0.27</td>
<td>.817</td>
<td>0.94</td>
<td>0.56 - 1.58</td>
</tr>
<tr>
<td>Partially redacted$r$ v. non-blind</td>
<td>-0.08</td>
<td>0.26</td>
<td>.756</td>
<td>0.92</td>
<td>0.55 - 1.54</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>-0.78</td>
<td>0.39</td>
<td>.043*</td>
<td>0.46</td>
<td>0.22 - 0.98</td>
</tr>
</tbody>
</table>

*Note.* $^r$ = reference group.

* = significant.
Table 8.

*Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a False Fingerprint ID*

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Fingerprint ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>18.36</td>
<td>5377.50</td>
<td>.997</td>
<td>0.00</td>
<td>0.00 - --</td>
</tr>
<tr>
<td>Blind v. partially redacted</td>
<td>-0.75</td>
<td>0.72</td>
<td>.300</td>
<td>0.47</td>
<td>0.12 - 1.95</td>
</tr>
<tr>
<td>Blind v. non-blind</td>
<td>0.10</td>
<td>0.03</td>
<td>.865</td>
<td>1.10</td>
<td>0.36 - 3.41</td>
</tr>
<tr>
<td>Partially redacted v. non-blind</td>
<td>0.85</td>
<td>0.71</td>
<td>.230</td>
<td>2.33</td>
<td>0.59 - 9.30</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>0.39</td>
<td>7496.57</td>
<td>1.00</td>
<td>1.48</td>
<td>0.00 - --</td>
</tr>
</tbody>
</table>

*Note.* † = reference group.
   * = significant.
Table 9.

**Ordinal Logistic Regression Analyses Examining Effects of Predictor Variables on Likelihood of Making a False Footwear Impression ID**

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Footwear Impression ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>-2.87</td>
<td>1.02</td>
<td>.005*</td>
<td>0.06</td>
<td>0.01 - 0.42</td>
</tr>
<tr>
<td>Blind* v. partially redacted</td>
<td>0.03</td>
<td>0.35</td>
<td>.931</td>
<td>1.03</td>
<td>0.53 - 2.03</td>
</tr>
<tr>
<td>Blind* v. non-blind</td>
<td>0.30</td>
<td>0.33</td>
<td>.373</td>
<td>1.35</td>
<td>0.70 - 2.58</td>
</tr>
<tr>
<td>Partially redacted* v. non-blind</td>
<td>0.27</td>
<td>0.33</td>
<td>.418</td>
<td>1.31</td>
<td>0.69 - 2.48</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>-0.10</td>
<td>1.35</td>
<td>.938</td>
<td>0.90</td>
<td>0.06 - 12.64</td>
</tr>
</tbody>
</table>

*Note.* * = reference group.
* = significant.
Table 10.

*Mixed ANOVA Analysis Examining Effects of Predictor Variables on Total Number of Inconclusive Judgments Made*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$M$</th>
<th>$SE$</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual Information</td>
<td>0.03</td>
<td>.973</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind</td>
<td>0.41</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-blind</td>
<td>0.43</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially redacted</td>
<td>0.41</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner Experience</td>
<td>0.33</td>
<td>.569</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>0.40</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>0.43</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Material</td>
<td>9.50</td>
<td>.002*</td>
<td>.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fingerprints</td>
<td>0.33</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footwear impressions</td>
<td>0.50</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Pair</td>
<td>48.73</td>
<td>.000*</td>
<td>.127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>0.22</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-match</td>
<td>0.61</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* * = significant.
### Table 11.

**Mixed ANOVA Analysis Examining Effects of Predictor Variables on Accuracy Rate**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$M$</th>
<th>$SE$</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual Information</td>
<td>0.16</td>
<td>.856</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind</td>
<td>0.73</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-blind</td>
<td>0.70</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially redacted</td>
<td>0.72</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner Experience</td>
<td>20.82</td>
<td>.000*</td>
<td>.058</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>0.79</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>0.64</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Material</td>
<td>6.94</td>
<td>.009*</td>
<td>.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fingerprints</td>
<td>0.76</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footwear impressions</td>
<td>0.67</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Pair</td>
<td>35.28</td>
<td>.000*</td>
<td>.095</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>0.81</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-match</td>
<td>0.61</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* * = significant.
Table 12.

*Mixed ANOVA Analysis Examining Effects of Predictor Variables on Discrimination Accuracy*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>M</th>
<th>SE</th>
<th>F</th>
<th>p</th>
<th>$\eta_p^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contextual Information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind</td>
<td>0.10</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-blind</td>
<td>0.04</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially redacted</td>
<td>0.28</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Examiner Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>0.34</td>
<td>0.06</td>
<td></td>
<td></td>
<td>0.024</td>
</tr>
<tr>
<td>Novice</td>
<td>-0.06</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stimulus Material</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fingerprints</td>
<td>0.34</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footwear impressions</td>
<td>-0.67</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* * = significant.
**Table 13.**

*Mixed ANOVA Analysis Examining Effects of Predictor Variables on Response Bias*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>M</th>
<th>SE</th>
<th>F</th>
<th>p</th>
<th>( \eta_p^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contextual Information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind</td>
<td>0.10</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-blind</td>
<td>0.07</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially redacted</td>
<td>0.05</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Examiner Experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>0.18</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>-0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stimulus Material</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fingerprints</td>
<td>-0.08</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footwear impressions</td>
<td>0.23</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* *= significant.*
Table 14.

*Mixed ANOVA Analysis Examining Effects of Predictor Variables on Confidence Ratings*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>M</th>
<th>SE</th>
<th>F</th>
<th>p</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual Information</td>
<td>0.76</td>
<td>.469</td>
<td>.005</td>
<td></td>
<td>.005</td>
</tr>
<tr>
<td>Blind</td>
<td>5.64</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-blind</td>
<td>5.65</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partially redacted</td>
<td>5.47</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examiner Experience</td>
<td>5.35</td>
<td>.021*</td>
<td></td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>5.75</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>5.43</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Material</td>
<td>16.15</td>
<td>.000*</td>
<td></td>
<td>.046</td>
<td></td>
</tr>
<tr>
<td>Fingerprints</td>
<td>5.78</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Footwear impressions</td>
<td>5.40</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus Pair</td>
<td>27.17</td>
<td>.000*</td>
<td></td>
<td>.075</td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>2.90</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-match</td>
<td>2.69</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * = significant.
Table 15.

*Confidence and Accuracy Bivariate Correlation by Predictor Variables*

<table>
<thead>
<tr>
<th>Condition</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contextual Information</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blind LAR</td>
<td>.088</td>
<td>.364</td>
</tr>
<tr>
<td>Non-blind LAR</td>
<td>.246</td>
<td>.007*</td>
</tr>
<tr>
<td>Redacted LAR</td>
<td>.093</td>
<td>.322</td>
</tr>
<tr>
<td><strong>Examiner Experience</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experts</td>
<td>.200</td>
<td>.143</td>
</tr>
<tr>
<td>Novices</td>
<td>.114</td>
<td>.053</td>
</tr>
<tr>
<td><strong>Stimulus Material</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fingerprint</td>
<td>.214</td>
<td>.000*</td>
</tr>
<tr>
<td>Footwear impression</td>
<td>.210</td>
<td>.000*</td>
</tr>
<tr>
<td><strong>Stimulus Pair</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>.200</td>
<td>.000*</td>
</tr>
<tr>
<td>Non-match</td>
<td>.219</td>
<td>.000*</td>
</tr>
</tbody>
</table>

*Note.* * = significant.
Table 16.

**Ordinal Logistic Regression Analyses Examining Effects of Contextual Information and Examiner Experience on Likelihood of Reviewing Time 1 Judgments**

<table>
<thead>
<tr>
<th>Time 1 Judgment</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>Exp(B)</th>
<th>95% CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Fingerprint ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>2.50</td>
<td>0.61</td>
<td>.000*</td>
<td>12.15</td>
<td>3.70 - 39.90</td>
</tr>
<tr>
<td>Blind v. partially redacted</td>
<td>-0.38</td>
<td>0.29</td>
<td>.195</td>
<td>0.69</td>
<td>0.39 - 1.21</td>
</tr>
<tr>
<td>Blind v. non-blind</td>
<td>0.02</td>
<td>0.29</td>
<td>.948</td>
<td>1.02</td>
<td>0.57 - 1.81</td>
</tr>
<tr>
<td>Partially redacted v. non-blind</td>
<td>0.40</td>
<td>0.27</td>
<td>.166</td>
<td>1.49</td>
<td>0.85 - 2.60</td>
</tr>
<tr>
<td>Contextual information x experience</td>
<td>0.29</td>
<td>0.74</td>
<td>.693</td>
<td>1.34</td>
<td>0.32 - 5.67</td>
</tr>
</tbody>
</table>

*Note.*  

* = reference group.  
* = significant.
Figure 1. Student Match Fingerprint Judgements
Figure 2. Student Non-Match Fingerprint Judgements
Figure 3. Average Number of Inconclusive Judgments by Examiner Experience and Stimulus Material.
Figure 4. Average Accuracy Rate by Examiner Experience and Evidence Material
Figure 5. Average Discrimination Accuracy by Examiner Experience and Evidence Material
Figure 6. Average Confidence Ratings by Examiner Experience and Evidence Material
LIST OF REFERENCES


IACP National Law Enforcement Policy Center, Eyewitness Identification, Model Policy, September 2010.


Langenburg, G., Champod, C., & Wertheim, P. (2009). Testing for potential contextual bias effects during the verification stage of the ACE-V methodology when conducting


President’s Council of Advisors on Science and Technology (2016). *Forensic science in criminal courts: Ensuring scientific validity of feature-comparison methods.*


Appendix A

Miami-Dade Police Department Lab Analysis Request Form

- **To:** Crime Laboratory □ Toxicology Laboratory □
- **From:** MDPD □ Submitting Agency:
- **MDPD Case Number**
- **Agency Case Number**
- **Category**
- **Type of Investigation**
- **Number of Packages**
- **Lead Investigator (Last, First Initial)**
- **Badge**
- **Title**
- **Unit**
- **Phone Number**
- **E-mail**
- **Fax Number**
- **Supervisor (Last, First Initial)**
- **Badge**
- **Title**
- **Phone Number**
- **E-mail**
- **Victim(s) (Last, First)**
- **Subject(s) (Last, First)**
- **Brief Case Summary**

**Routing Request Table**

<table>
<thead>
<tr>
<th>Item #</th>
<th>Biology</th>
<th>Chemistry</th>
<th>CODIS</th>
<th>Firearms</th>
<th>Serology</th>
<th>Latents</th>
<th>NIBIN</th>
<th>Toxicology</th>
<th>Trace</th>
<th>Additional Notes</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Item #</th>
<th>Biology</th>
<th>Chemistry</th>
<th>CODIS</th>
<th>Firearms</th>
<th>Serology</th>
<th>Latents</th>
<th>NIBIN</th>
<th>Toxicology</th>
<th>Trace</th>
<th>Additional Notes</th>
</tr>
</thead>
</table>

- **Submitter (Last, First Initial)**
- **Badge**
- **Unit**
- **Phone Number**
- **Signature**
- **Date**

FSB use only
Appendix B

Blind Condition Lab Analysis Request Form

Laboratory Analysis Request
Miami-Dade Police Department

1. To: Crime Laboratory ☒ Toxicology Laboratory ☐

2. From: MDPD ☒ Submitting Agency:

3. MDPD Case Number
   A089WMP

4. Agency Case Number
   PD146209851169

5. Category
   ATT. STRONG ARM ROBBERY

6. Type of Investigation

7. Number of Packages
   2

8a. Lead Investigator (Last, First Initial)
    GONZALEZ, R.

8b. Badge
    5898

8c. Title
    DET.

8d. Unit
    R319

8e. Phone Number
    305-471-9602

8f. E-mail
    RGONZALEZ@MDPD.COM

8g. Fax number
    305-471-2019

9a. Supervisor (Last, First Initial)
    BALLARD, A.

9b. Badge
    4441

9c. Title
    SGT.

9d. Phone Number
    305-471-8000

9e. E-mail
    ABALLARD@MDPD.COM

10. Victim(s) (Last, First)
    MENDEZ, JULIA

11. Subject(s) (Last, First)
    1. RODRIGUEZ, ERICK  2. ENGLISH, MICHAEL

12. Brief Case Summary
    Routine.

13. Routing Request Table

<table>
<thead>
<tr>
<th>Item #</th>
<th>BIOLOGY</th>
<th>CHEMISTRY</th>
<th>CODIS</th>
<th>FIREARMS</th>
<th>GSR KIT</th>
<th>LATENTS</th>
<th>MIBIN</th>
<th>TOXICOLOGY</th>
<th>TRACE</th>
<th>14. Additional notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LIGHTER</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FINGERPRINTS FROM PASSENGER DOOR WINDOW</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FOOTWEAR IMPRESSION</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHOES SUBJECT #1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FOOTWEAR IMPRESSION</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHOES SUBJECT #2</td>
</tr>
</tbody>
</table>

15a. Submitter (Last, First Initial)
    GONZALEZ, R.

15b. Badge
    5898

15d. Unit
    R319

15c. Phone Number
    305-471-9602

15d. Signature

15e. Date
    02/14/17

FSB use only

118
Appendix C

Non-Blind Condition Lab Analysis Request Form

Laboratory Analysis Request
Miami-Dade Police Department

<table>
<thead>
<tr>
<th>1. To:</th>
<th>Crime Laboratory ☒</th>
<th>Toxicology Laboratory ☐</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. From:</td>
<td>MDPD ☒</td>
<td>Submitting Agency:</td>
</tr>
<tr>
<td>3. MDPD Case Number</td>
<td>A089WMP</td>
<td></td>
</tr>
<tr>
<td>4. Agency Case Number</td>
<td>PD148209851169</td>
<td></td>
</tr>
<tr>
<td>5. Category</td>
<td>ATT. STRONG ARM ROBBERY</td>
<td></td>
</tr>
<tr>
<td>6. Type of Investigation</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7. Number of Packages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8a. Lead Investigator (Last, First Initial)</td>
<td>GONZALEZ, R.</td>
<td></td>
</tr>
<tr>
<td>8b. Badge</td>
<td>5898</td>
<td></td>
</tr>
<tr>
<td>8c. Title</td>
<td>DET.</td>
<td></td>
</tr>
<tr>
<td>8d. Unit</td>
<td>R319</td>
<td></td>
</tr>
<tr>
<td>8e. Phone Number</td>
<td>305-471-9602</td>
<td></td>
</tr>
<tr>
<td>8f. E-mail</td>
<td><a href="mailto:R.GONZALEZ@MDPD.COM">R.GONZALEZ@MDPD.COM</a></td>
<td></td>
</tr>
<tr>
<td>8g. Fax number</td>
<td>305-471-2019</td>
<td></td>
</tr>
<tr>
<td>9a. Supervisor (Last, First Initial)</td>
<td>BALLARD, A.</td>
<td></td>
</tr>
<tr>
<td>9b. Badge</td>
<td>4441</td>
<td></td>
</tr>
<tr>
<td>9c. Title</td>
<td>SGT.</td>
<td></td>
</tr>
<tr>
<td>9d. Phone Number</td>
<td>305-471-8000</td>
<td></td>
</tr>
<tr>
<td>9e. E-mail</td>
<td><a href="mailto:A.BALLARD@MDPD.COM">A.BALLARD@MDPD.COM</a></td>
<td></td>
</tr>
<tr>
<td>10. Victim(s) (Last, First)</td>
<td>MENDEZ, JULIA</td>
<td></td>
</tr>
<tr>
<td>11. Subject(s) (Last, First)</td>
<td>1. RODRIGUEZ, ERICK 2. ENGLISH, MICHAEL</td>
<td></td>
</tr>
</tbody>
</table>

12. Brief Case Summary
Victim was getting into her car after working at the Sedano's supermarket on 10795 NW 58th St. Car was parked facing northbound on dirt lot. Two subjects forced entry to car at gun point from both sides of the car. Subject #1 demanded victim's purse and jewelry from passenger side door. Subject #2 fled with victim's purse. Both subjects fled. Victim believes to have seen Subject #2 drop an object when running. A lighter was found approximately 20ft from car. Latent lifts were collected from inside of car. Shoeprint impressions were collected 30ft from car. Subject #1 & Subject #2 are in custody. Subject #1 is believed to have assaulted multiple women and Subject #2 has been recently released from prison.

13. Routing Request Table

<table>
<thead>
<tr>
<th>Item #</th>
<th>BIOLOGY</th>
<th>CHEMISTRY</th>
<th>CODIS</th>
<th>FIREARMS</th>
<th>GSR INT</th>
<th>LATENTS</th>
<th>NIBIN</th>
<th>TOXICOLOGY</th>
<th>TRACE</th>
<th>14. Additional notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td><em>X</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LIGHTER</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td><em>X</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FINGERPRINTS FROM PASSENGER DOOR WINDOW</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FOOTWEAR IMPRESSION</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td><em>X</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHOES SUBJECT #1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td><em>X</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FOOTWEAR IMPRESSION</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td><em>X</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SHOES SUBJECT #2</td>
</tr>
</tbody>
</table>

15a. Submitter (Last, First Initial) | GONZALEZ, R. |
15b. Badge | 5898 |
15d. Unit | R319 |
15c. Phone Number | 305-471-9602 |
15e. Date | 02/14/17 |

FSS use only
Appendix D

Partially Redacted Condition Lab Analysis Request Form
Appendix E

Critical Pieces of Information on Non-Blind and Partially Redacted Condition Lab Analysis Request Forms

1. Two subjects forced entry to car
   a. Needs to be known for collection and examination purposes. Examiners need to know that more than one subject entered the car because if they are not given this information, then they may believe that they can stop analyzing the evidence after they analyze the first pair of prints. If there is more than one subject, then examiners are aware that more evidence needs to be analyzed and tested.

2. A lighter was found
   a. Needs to be known for examination purposes. Examiners need to know that a lighter was found at the crime scene and thus needs to be analyzed for prints and DNA.

3. Latent lifts were collected from inside of car
   a. Needs to be known for examination and database entry purposes. Examiners need to know where the fingerprints were lifted from. If they were collected from inside the care, then the examiner knowns that it can be either from the victim, either one of the subjects who entered the car, or any other person who has been inside the vehicle. Also, in order for the evidence to be entered into certain databases (e.g., AFIS) items must be directly associated to the crime and so knowing that the prints were lifted
form inside of the car and the subjects forced entry into the car, then the prints become associated with the crime.
Appendix F

Lab Analysis Request Form Questionnaire

Please answer the following open-ended and multiple-choice questions regarding the lab analysis request form you were given.

1. Please write down everything you remember about the information provided in the lab request form.

__________________________________________________________________
__________________________________________________________________
__________________________________________________________________

2. What type of case summary did you receive as part of the Lab Analysis Request Form?
   a. Complete case summary
   b. Partially redacted (i.e., blacked-out) case summary
   c. No case summary (routine)

3. How many fingerprint pairs were you asked to analyze?
   a. 1 pair of fingerprints
   b. 2 pairs of fingerprints
   c. 3 pairs of fingerprints
   d. I don’t know

4. How many footwear impression pairs were you asked to analyze?
   a. 1 pair of footwear impressions
   b. 2 pairs of footwear impressions
   c. 3 pairs of footwear impressions
   d. I don’t know

5. What type of crime occurred?
   a. Attempted murder
   b. Attempted strong armed robbery
   c. Carjacking
   d. Information not given
   e. I don’t know

6. How many suspects were involved?
   a. 1 suspect
   b. 2 suspects
   c. 3 suspects
7. How many victims were involved?
   a. 1 victim
   b. 2 victims
   c. 3 victims
   d. Information not given
   e. I don’t know

8. What evidence was submitted to the crime lab?
   a. A gun
   b. A lighter
   c. A hat
   d. Information not given
   e. I don’t know

9. Where were fingerprints lifted from the car?
   a. Steering wheel
   b. Passenger door window
   c. Trunk of the car
   d. Information not given
   e. I don’t know

10. Where was the victim coming from?
    a. She was leaving work
    b. She was leaving school
    c. She was leaving a concert
    d. Information not given
    e. I don’t know

11. Where did the crime take place?
    a. At a hotel parking lot
    b. Near a supermarket
    c. On a school campus
    d. Information not given
    e. I don’t know

12. What did the suspects demand from the victim?
    a. Her purse and jewelry
    b. Her cellphone
    c. Nothing
d. Information not given
e. I don’t know

13. What weapon was used?
a. A gun
b. A knife
c. There was no weapon
d. Information not given
e. I don’t know

14. What item was taken from the victim?
a. A cellphone
b. A purse
c. Nothing
d. Information not given
e. I don’t know

15. Who dropped an item?
a. Subject #1
b. Subject #2
c. The victim
d. Information not given
e. I don’t know

16. Where was the car parked?
a. In a parking garage
b. On a dirt lot
c. At a meter
d. Information not given
e. I don’t know

17. How many suspects were in custody?
a. None
b. Only one
c. Both
d. Information not given
e. I don’t know
Appendix G
Perceptions Questionnaire

1. Across all judgments you made, how accurate do you think you were?

   0% (not at all accurate) – 100% (extremely accurate)

2. How guilty do you think suspect #1, Erick Rodriguez is?

   1 (definitely innocent) – 10 (definitely guilty)

3. How guilty do you think suspect #2, Michael English is?

   1 (definitely innocent) – 10 (definitely guilty)

For participants in non-blind condition:

1. Do you feel that sufficient information was provided in the lab analysis request forms to make accurate judgements?
   a. Yes
   b. No
      If not, why? ______________

2. Do you feel that your analyses were biased because you had information about the case?
   c. Yes
   d. No

3. Which type of lab request form would you prefer to have when analyzing forensic evidence?
   e. A lab analysis request form with all the possible case information
   f. A lab analysis request form with all irrelevant case information redacted
   g. A lab analysis request form with no case information

For participants in partially redacted condition:

1. Did you feel that sufficient information was provided in the first lab analysis request form to make accurate decisions about the forensic evidence?
   h. Yes
   i. No
      If not, why? ______________

2. Did you feel more comfortable in your decision after receiving all the information about the case in your second lab request form?
j. Yes  
k. No  

3. Do you feel that you may have been unduly influenced if provided with the redacted information (blacked-out) in your first lab request form?  
   l. Yes  
   m. No  

4. Which type of lab request form would you prefer to have when analyzing forensic evidence?  
   n. A lab analysis request form with all possible case information  
   o. A lab analysis request form with all irrelevant case information redacted  
   p. A lab analysis request form with no case information  

For those who had a case summary with no information:  

1. Did you feel that sufficient information was provided in the first lab analysis request form to make accurate decisions about the forensic evidence?  
   a. Yes  
   b. No  
   
   If not, why? ________________  

2. Did you feel more comfortable in your decision after receiving all the information about the case in your second lab request form?  
   a. Yes  
   b. No  

3. What was your interpretation of the word “Routine” found in the first lab request form?  

________________________________________________________________________  
________________________________________________________________________  
________________________________________________________________________  
________________________________________________________________________  

4. Do you feel that you may have been unduly influenced if provided with case information in your first lab request form?  
   a. Yes  
   b. No  

5. Which type of lab request form would you prefer to have when analyzing forensic evidence?  
   a. A lab analysis request form with all possible case information  
   b. A lab analysis request form with irrelevant case information redacted  
   c. A lab analysis request form with no case information
Appendix H
Demographics Questionnaire

1. What is your age? _______

2. What is your gender?
   a. Male
   b. Female

3. What is your ethnicity
   a. Caucasian
   b. Hispanic or Latino
   c. Black or African American
   d. Asian or Pacific Islander
   e. Native American or American Indian
   f. Other (Please Specify: _______________)

4. What is your highest level of education?
   a. High school degree
   b. Associates Degree
   c. Bachelors’ Degree
   d. Master’s Degree
   e. Doctoral Degree
   f. Other (Please Specify: _______________)

5. What is your occupation? _______________

6. Have you had any experience analyzing forensic evidence?
   a. No
b. Yes

   i. If yes, please briefly explain your experience below:

       ______________________________________________________
       ______________________________________________________
       ______________________________________________________

7. Please select one of the following:
   
a. I am an undergraduate student at FIU

   b. I am a trained forensic examiner (or have experience analyzing forensic evidence)

The following questions will be asked only to expert forensic examiners.

8. How many years of active casework experience have you completed?

    ____________

9. Are you employed by a law enforcement agency?
   
a. No
   
   i. If no, please explain

       ______________________________________________________
       ______________________________________________________
       ______________________________________________________

b. Yes

   i. If yes, what type of Law Enforcement Agency are you employed by?

       1. Local
       2. State
3. Federal

10. Are latent print examinations your primary duty?
   a. Yes
   b. No

11. Do you examine other types of evidence in addition to latent fingerprints?
   a. No
   b. Yes (If yes, please specify: _______________________________)

12. Are you an IAI certified latent print examiner?
   a. Yes
   b. No
   c. Does not apply

13. Do you usually apply ACE-V methodology in latent print examinations?
   a. Yes
   b. No
   c. Does not apply

14. Have you completed any structured latent fingerprint training program?
   a. Yes
      i. If yes, how long was the training program?
         1. 0 to 1 years
         2. 1 to 2 years
         3. More than 2 years
      ii. If yes, how long ago did you complete this training?
         1. Less than a year ago
2. Between 1 to 2 years ago

3. More than 2 years ago

iii. If yes, please describe the type of structured fingerprint training you have completed.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
____

b. No
VITA

MICHELLE PENA

2010-2013  B.A., Psychology
Florida International University
Miami, Florida

2014-2017  M.S., Psychology
Florida International University
Miami, Florida

2017-Present  Doctoral Candidate
Florida International University
Miami, Florida

Teaching Assistant
Florida International University
Miami, Florida

PUBLICATIONS AND PRESENTATIONS


Pena, M.M. (2019, May). *Bridging the gap between research and practice: How practitioners can enable the psychological study of forensic science.* Presented at the 8th Annual Forensic Science Symposium, Miami, FL.


