

Stereotypes and the Familiar-Stranger: The Role of Previous Contact and Gender Stereotypes on
Eyewitness Recall and Recognition Accuracy

by

Lauren E. Thompson

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Psychology

Carleton University

Ottawa, Ontario

© 2018

Lauren E. Thompson

Abstract

The purpose of the present study was to examine the influence of prior familiarity with a perpetrator and gender stereotypes on eyewitness recall and identification accuracy. Participants ($N = 257$) watched a crime video where the perpetrator was either someone they had never met before (unfamiliar/stranger condition), or was someone with whom they had a 1 minute exposure to prior to the crime (familiar condition). In the familiar conditions the 1 minute exposure included the target talking about their occupation; this occupation was either consistent or inconsistent with gender role stereotypes. There were no significant differences in recall accuracy or identification accuracy between the familiar or stereotypes conditions. However, when participants were asked to rate the degree to which they viewed the target as stereotype consistent versus inconsistent, higher stereotype consistency ratings were associated with reporting more total descriptors and higher proportions of correct descriptors (although this relationship was only marginally significant). On the other hand, lower stereotype consistency ratings predicted more correct identification decisions. It was hypothesized that cognitive load would influence familiar eyewitnesses memory such that eyewitnesses introduced to an inconsistent target would have better recall and identification accuracy. These hypotheses were based on previous research suggesting that memory is enhanced for stereotype inconsistent information when under low cognitive load. These hypotheses were not supported by the results as there were no group differences in ratings on the cognitive load measures. Overall, the results from the current study suggest that a 1 minute exposure to the target prior to the crime may not be enough exposure to elicit a feeling of familiarity and enhance eyewitness memory. These results also suggest that subjective gender role stereotypes influence memory for a previously met person. Conclusions and future directions are discussed.

Acknowledgements

First and foremost I would like to extend many thanks to my supervisor, Dr. Joanna Pozzulo. Thank you for your ongoing support and assistance throughout this degree. I am very grateful to have had your guidance throughout the last two years and appreciate all the time you have put into my progress!

I would also like to extend my gratitude to my committee members Dr. Craig Bennell and Dr. Andrew Smith. Thank you for your feedback and assistance throughout this project. I very much appreciate all the added knowledge I acquired from having the both of you sit on my committee. In addition, thank you to my internal examiner Dr. Diana Young for your interesting and thought provoking feedback during my defence. I also owe a big thank you to Dr. Chelsea Sheahan for reading over my drafts and being a mentor to me in our lab. Of course, a big thanks to my “actors” Alex Skutovich, Lindsay Fulham, and Grayden Drew for your participation and patience with the recordings for this project! I am very grateful for the three of you volunteering your time to help!

And last but not least, thank you to my family, just because you are the absolute best.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	vii
List of Appendices.....	viii
Introduction.....	1
Overview of Eyewitness Identification Research.....	4
Recall.....	4
Recognition.....	5
Target-present and target-absent lineups.....	7
Lineup procedures.....	8
Familiarity.....	11
Recall of familiar-strangers.....	11
Event repetition.....	12
Exposure duration.....	14
Theories of recall memory.....	14
Recognition of familiar-strangers.....	17
Facial recognition studies.....	18
Eyewitness studies.....	21
Theories and regularities in recognition memory.....	24
Stereotypes.....	26
Gender stereotypes.....	27

Social role theory and occupational stereotypes.....	28
Stereotype consistent vs. stereotype inconsistent.....	32
Cognitive load.....	32
Inconsistent vs. consistent gender stereotypes.....	36
Stereotypes and eyewitness research.....	37
Stereotypes and familiar-strangers.....	39
Measuring cognitive load.....	40
Overview of the Current Study.....	43
Hypotheses.....	44
Method.....	45
Participants.....	45
Design.....	45
Materials.....	46
Procedure.....	52
Results.....	54
Stereotypes Manipulation Check.....	54
Cognitive Load Measures.....	54
Recall Data.....	55
Coding data.....	55
Recall accuracy.....	57
Stereotype ratings questionnaire.....	59
Cognitive load and recall memory.....	60
Identification Data.....	61

Identification accuracy.....	61
Stereotype ratings questionnaire.....	63
Cognitive load and identification accuracy.....	63
Discussion.....	64
Familiarity.....	66
Recall of familiar-strangers.....	66
Identifying familiar-strangers.....	70
Stereotypes.....	74
Stereotypes and recall accuracy.....	75
Cognitive load and recall accuracy.....	76
Stereotypes and identification accuracy.....	78
Cognitive load and identification accuracy.....	79
Limitations.....	81
Implications and Future Research.....	82
References.....	84

List of Tables

Table 1. Proportion (Frequency) of Lineup Members Chosen from the Lineup Fairness Test for Each Targets Lineup	50
Table 2. Intraclass Correlation Coefficient (ICC) Representing the Interrater Reliability of The Targets' Recall Data	57
Table 3. Means and Standard Deviation for Total Descriptors and Proportion of Correct Descriptors (%) by Familiarity Consistency.....	59
Table 4. Percentage of Correct Identification Decisions based on Familiarity Consistency and Target Presence.....	62
Table 5. Mean Difference between Stereotype Ratings for Various Occupations	129

List of Appendices

Appendix A: SONA Recruitment Notice	108
Appendix B: Informed Consent Form	109
Appendix C: Stereotype Profile Descriptions.....	111
Appendix D1: Stereotype Questionnaire	113
Appendix D2: Stereotype Questionnaire	113
Appendix E: Open Recall	114
Appendix F: Cognitive Load Measure (RSME).....	115
Appendix G: Cognitive Load Measure (NASA-RTLX).....	116
Appendix H: Secondary Informed Consent Form	118
Appendix I: Demographics Form	120
Appendix J: Description Form.....	121
Appendix K: Simultaneous Lineup Response Form.....	122
Appendix L: Study Checks	123
Appendix M: Debriefing Form	124
Appendix N: Material Testing Study Information.....	126
Subsection i Method	126
Subsection ii Results.....	128
Subsection iii Sona Recruitment Form	130
Subsection iv Informed Consent Form	131
Subsection v Stereotype Ratings Questionnaire	133
Subsection vi Debriefing Form.....	135

Stereotypes and the Familiar-Stranger: The Role of Previous Contact and Gender Stereotypes on Eyewitness Recall and Recognition Accuracy

Eyewitness identification evidence is frequent and persuasive in the investigation and prosecution of defendants in many criminal cases (e.g., Cutler & Penrod, 1995; Flowe, Mehta, & Ebbesen, 2011). However, an analysis of actual wrongful conviction cases in the U.S. reveals that misidentification on the part of eyewitnesses and victims is the most common mistake leading to the arrest and conviction of innocent people (Connors, Lundregan, Miller, & McEwen, 1996) with several Canadian cases suggesting similar issues (e.g., R.v. Hanemaayer, 2008; R. v. Sophonow, 1985). With the high prevalence of eyewitness misidentification sending the wrong people to jail, numerous studies have been conducted in order to try and inform best practices for gathering and preserving eyewitness evidence (Technical Working Group for Eyewitness Evidence, 1999, 2003; Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000).

To date, the majority of eyewitness research has focused on eyewitness recall and identification of strangers (i.e., never before seen targets). However, many crimes in which eyewitnesses are of use are not committed by strangers but rather are committed by people who are *known* or familiar to the witness (Flowe et al., 2011; Memon, Havard, Clifford, Gabbert, & Watt, 2011; Valentine, Pickering, & Darling, 2003). Moreover, not only are many crimes committed by non-strangers, but this factor also influences how cases are perceived by the courts. Research conducted by Flowe et al. (2011) randomly sampled 725 violent felony cases from a District Attorney's closed case archives (1991-2000) in California. Of these cases, 237 were classified as stranger cases while 488 were classified as acquaintance cases. Within the stranger cases, positive identification evidence from an eyewitness did not significantly increase the odds that a case was prosecuted. However, for the acquaintance cases, positive identification

evidence from an eyewitness did significantly increase the odds a case was prosecuted.

Therefore, it appears that not only are many crimes committed by non-strangers, but that this factor influences how cases are eventually processed through the criminal justice system.

Despite the statistics, eyewitness research rarely distinguishes between stranger and non-stranger identification. Specifically, eyewitness recall and recognition of *familiar-strangers* (i.e., individuals who are not well known to the witness but with whom the witness has had brief contact with prior to the crime) has been neglected in the research. Whether the eyewitness has previously come into contact with the perpetrator would likely affect their memory of the perpetrator, therefore it seems there is a gap in the literature on examining the accuracy of these types of witnesses. The current study will examine the influence of familiarity on recall and identification accuracy.

Closely linked to the concept of a familiar-stranger (in the sense of brief contact) is the idea of schema-activation and stereotyping. Person perception models indicate that the first step in perceiving an individual includes categorization (Brewer, 1988; Fiske & Neuberg, 1990). Categories and category membership allow for prejudgement as placing something in a category associates it with all the ideations of that category (Allport, 1954). Therefore, categorization is both a necessary and sufficient condition of stereotyping. Stereotypes can be defined as a widely held belief that persons of a specific group have certain characteristics as a result of their membership within that group (Hilton & von Hippel, 1996). Stereotypes are learned and over learned to the point of being automatically activated (i.e., without conscious awareness).

Research has demonstrated that categorizing as male or female is one of the first ways in which people sort others in social scenarios (e.g., Blair & Banaji, 1996; Brewer & Lui, 1989; Stangor & McMillan, 1992). Once the perceiver has activated their schema of male or female, it

is then that other incoming information can be regarded as consistent or inconsistent with the person's gender. It has been suggested that humans hold deeply ingrained ideas about gender roles and what are suitable behaviours for men and women, that then influence people's perception of gender-specific occupations (Shinar, 1975). Therefore, occupations tend to be highly gender role stereotyped. Stereotypes are known to affect memory through all processes (i.e., encoding, maintenance, retrieval), and whether an individual is perceived to be stereotype consistent or inconsistent effects how memories of them are stored and retrieved (Cano, Hopkins, & Islam, 1991; Koomen & Dijker, 1997). Thus, gender role stereotypes would likely influence eyewitness memory of a familiar-stranger. The currently study also will consider the consistency of stereotypes and their influence on recall and identification accuracy.

Therefore, it is the purpose of this program of research to examine the difference in eyewitness recall and identification of familiar-strangers compared to strangers. Furthermore, the current research will examine how occupational gender stereotypes may affect eyewitness recall and recognition of familiar-strangers. This is one of the first studies that has looked at eyewitness recall and recognition of familiar-strangers, and is the only known study that has looked at how occupational gender stereotypes influences eyewitness memory of familiar-strangers.

The following sections of this thesis will begin with an overview of eyewitness identification research focusing on best practices and procedures. A discussion of familiarity and the available research on recall and recognition memory for familiar-strangers will follow. Following the discussion of familiarity will be a discussion of stereotypes. Specifically, theories and research on gender stereotypes and occupational gender stereotypes will be reviewed. Hypotheses are provided following a brief summary of the literature. Next, the method will be presented, followed by the results and discussion of findings.

Overview of Eyewitness Identification Research

In addition to this brief overview of best practices regarding eyewitness recall and recognition memory, the discussion of research and theory on familiarity will be organized into recall memory and recognition memory.

Recall. After witnessing a crime, a witness will be asked to recall the details of the event. Eyewitness memories are based on what they perceived at the time of the crime but also partially on their expectations, beliefs, and prior knowledge (Haber & Haber, 2001). For this reason, eyewitness accounts need to be retrieved as soon as possible in order to get the most accurate report of what took place (Mackay & Paterson, 2015; Wang, Paterson, & Kemp, 2014). Further, this information should be gathered before any interaction with other witnesses as research has demonstrated that discussion between witnesses can distort the personal accounts of each witness (e.g., Gabbert, Memon, & Allan, 2003; Gabbert, Memon, Allan, & Wright, 2004; Paterson, Kemp, & Ng, 2011). Additionally, these accounts should be taken without any interruptions and in open-ended format (i.e., with little use of direct or suggestive questioning; e.g., Lipton, 1977). Research has shown that free recall renders the most accurate statements, while more direct questions can lead to more inaccuracies (e.g., Fisher, Geiselman, & Raymond, 1987; Lipton, 1977).

Unfortunately, research on eyewitnesses' ability to recall and describe a perpetrator suggests that witnesses provide few and vague details that could apply to several people, therefore contributing little in the way of narrowing down the suspect pool (Lindsay, Martin, & Webber, 1994; Meissner, Sporer, & Schooler, 2007; Odnot & Wolters, 2006). For example, Kuehn (1974) examined the amount and type of perpetrator descriptors reported by victims to the Seattle Police Department in 1967. A random sample of 100 violent crimes were examined for

various physical trait descriptors of the perpetrator: age, sex, height, weight, build, skin colour, hair, and eye colour. They found that the number of descriptors ranged from 0 to 9 and that the mean number of descriptors provided was 7.2. The most commonly provided descriptors were sex, age, height, and build, with the least provided descriptor being eye colour.

Lindsay and colleagues (1994) also examined perpetrator descriptors provided by eyewitnesses. In this study they looked at descriptions of 105 criminals published in a local newspaper in Kingston, Ontario. Similar to Kuehn (1974) the most commonly provided descriptors were sex, clothing, age, and height. Lindsay et al. were additionally interested in looking at how eyewitness descriptors in laboratory studies would compare to the eyewitness descriptors reported in actual crimes. In comparison to the newspaper report, the laboratory participants provided significantly more details ($M = 7.35$) than the real crime witnesses ($M = 3.94$). However, the descriptors most commonly reported by the laboratory participants were similar in nature to the real crime witnesses with clothing, hair colour, and height being most reported. Therefore, it seems that although real crime witnesses may provide fewer details than laboratory participants, the types of descriptors they are providing are similar.

Laboratory studies of eyewitness descriptors are beneficial since recall accuracy can be measured. Given that it is unknown in real crime cases who the perpetrator is, it is important to use laboratory experiments to assess recall accuracy as these descriptors can be verified. In the current study, eyewitness recall will be measured and descriptors will be assessed for accuracy.

Recognition. After providing a description of the event and the perpetrator, the investigating officers will try to find one or several suspects that fit the witness' description and details of the crime. Once one or more suspects have been identified, the police will construct lineups around these suspects. Lineups can either be in a mode of live, video, or photographic. In

Canada, photographic lineups are most prevalent and are seen to have a number of advantages. First, it is easier to construct “fair” (i.e., suspect does not stand out) photographic arrays due to the large photograph pools (mugshots) available and second, live and video lineups can be obstructed by suspects acting out in ways that draw attention to themselves (Turtle, Lindsay, & Wells, 2003). Turtle et al. (2003) used the extensive research on perpetrator lineups and various reports in the U.S. and Canada to formulate “best practice recommendations” for collecting and preserving eyewitness memory in lineups. When composing the lineup, Turtle et al. suggests that only one suspect is included in each identification procedure; that there be enough foils (i.e., individuals known to be innocent to the crime in question; usually between 5 to 11); that foils selected generally fit the description of the perpetrator; that investigators avoid using foils that too closely resemble the suspect (i.e., even a person very acquainted with the suspect would find it difficult to differentiate the suspect from the foils); and that the position of the suspect’s photo be different in each lineup (i.e., if multiple witnesses in the same case and across cases). Turtle et al. also provide guidelines for what should be said to a witness prior to viewing the lineup. This includes but is not limited to, informing the witness they will be asked to view a set of photographs, instructing the witness that the individuals depicted might not appear exactly as they did on the date of the incident (i.e., features such as head and facial hair could have changed), and informing the witness that the picture of the perpetrator may or may not be in the set of photographs provided. Additionally, it is also suggested that the investigator ask the witness to state how certain he or she is of the identification. This suggestion though, is the most controversial. The main concern with this suggestion is that the inclusion of a certainty measure reflects an attempt, on the part of the investigators, to make witnesses feel a heightened sense of confidence (Turtle et al., 2003). Further, there is concern that asking for a numerical value could

result in an arbitrarily lower value based on individual's not feeling comfortable using the scales' upper limits. Although previous research has not found a significant confidence-accuracy relationship (e.g., Deffenbacher, 1980), more recent research has found that the relationship between initial confidence and accuracy is strong (Sauerland, Sagana, & Sporer, 2012; Wixted, Mickes, Clark, Gronlund, & Roediger, 2015; Wixted & Wells, 2017). In addition, Turtle et al. argue that in real world criminal investigations, the reason for including a confidence estimate is based on common sense and common practice. As an investigation proceeds witnesses will receive many cues by the investigating officers as to whether they made the "correct" choice, by having a measure of confidence taken right after the choice is made, this reduces the risk of confidence inflation. In order to try and mimic what would occur in the real world, a measure of confidence was included in the present study.

Target-present and target-absent lineups. There are two types of lineups that can be constructed for research purposes in which the presence of the target is varied. The first is a lineup including the guilty suspect and is referred to as a target-present (TP) lineup. An alternative lineup is one that is known not to contain the guilty suspect and instead contains a designated innocent suspect (Wells, 1993); this is referred to as a target-absent (TA) lineup. In TP lineups, there are three possible eyewitness decisions that can be made. The first is that the eyewitness correctly chooses the guilty target, known as a correct identification. In research and in the real-world, this is the only correct decision that can be made when presented with a TP lineup. The second is that the eyewitness incorrectly chooses a foil, known as a foil identification. In the real world, a foil identification is the only decision that police are aware is incorrect (Lindsay & Wells, 1985). This decision may indicate to police that the eyewitness is not credible or has a poor memory of the perpetrator (Pozzulo & Lindsay, 1999). Choosing a foil

may also indicate that the suspect is innocent, as studies have demonstrated that choosing a foil is more likely to occur in TA lineups than in TP lineups (e.g., Clark, Howell, & Davey, 2008; Wells, Yang, & Smalarz, 2015). And last, the witness can incorrectly dismiss the lineup and not select anyone; this is known as a false rejection. In the real world, police are not aware of this incorrect decision and thus a false rejection may result in the guilty perpetrator remaining in the community.

In TA lineups, there are as well three possible eyewitness decisions that can be made. First, the eyewitness may correctly dismiss the lineup, known as a correct rejection. In research and in the real world, this is the only correct decision that can be made with a TA lineup. Second, the witness may again incorrectly choose a foil (foil identification). As with TP lineups, this would indicate to police that the eyewitness may not be credible and the identification will not be pursued. And last, the witness may incorrectly choose the innocent suspect, known as a false identification. In the real world, a false identification is perhaps the most severe incorrect decision an eyewitness can make. This type of identification may cause an innocent person to be sent to jail (Beal, Schmitt, & Dekle, 1995) and further, the guilty perpetrator may remain in the community (Malpass & Devine, 1981).

The lineups used in eyewitness identification research mimic real-life eyewitness lineups with the added benefit of knowing who is the guilty suspect. In a real-life investigation, the police are unaware of the guilt or innocence of a suspect and therefore do not know if they are presenting a TP or TA lineup. As a result, it is important to research how eyewitnesses perform in both TP and TA conditions. Both TP and TA lineups will be used in the current study.

Lineup procedures. In addition to presenting either TP or TA lineups, there is also a variety of ways in which lineups can be presented. The most common lineup procedures are the

simultaneous lineup and the sequential lineup. The simultaneous lineup is when a witness views all lineup members at the same time. The pictures of the lineup members are presented in one single line or in several lines. The simultaneous lineup is the most commonly used lineup in the United States (Police Executive Research Forum, 2013). This type of lineup may encourage the witness to make a relative judgement; that is, comparing and contrasting the various lineup members and determining who looks *most* like the perpetrator (Wells, 1993). The reason a relative judgement may be made with a simultaneous lineup procedure is a result of the eyewitness being unable to fully accept the possibility that the perpetrator might not be in the set of photographs (Wells, 1993). Wells (1993) argues that a relative judgement is influenced by *ecphoric similarity*. Ecphoric similarity is a similarity judgement made between a stimulus and a memory trace (Tulving, 1981). If one lineup member displays greater ecphoric similarity than the others, there is a tendency to identify that lineup member even if the amount of ecphoric similarity is only slight. Wells demonstrates the use of relative judgements in simultaneous lineups by observing rates of identification in situations where a target is removed without replacement from a set of foils. In this study 200 participants were shown a video depicting a staged crime, half the participants were then shown a lineup containing the perpetrator (TP lineup), and the other half were shown the same lineup but with the perpetrator removed. All participants were told that the perpetrator may or may not be in the lineup. The results demonstrated that when the target was present in the lineup, 54% of participants correctly identified this lineup member as the perpetrator. When the target was removed without replacement, the lineup member who was the second-best choice in the TP lineup became the best choice in the target-removed lineup with 38% of participants identifying this target. Therefore, the results of this research are consistent with a relative judgment process. When the

target is in the lineup, they are most likely to be identified because they most resemble themselves, but when the target is removed, another person in the lineup becomes the one who most resembles the target (Wells, 1993). When the target is in the lineup, as with a TP lineup, a relative judgment is not problematic as the target is *most* likely to look like him- or her- self. The issue is in a situation where the suspect is not the guilty perpetrator; a relative judgement can lead to a false identification as the person who best resembles the perpetrator is likely to be chosen (Pozzulo, Reed, Pettalia, & Dempsey, 2015).

The sequential lineup is where a witness views one photograph at a time (i.e., serially) and must make a decision as to whether the person being shown is the perpetrator before they can move on to viewing the next photo (Lindsay & Wells, 1985). This lineup is the most commonly used procedure in Canada (Beaudry & Lindsay, 2006). This lineup procedure was designed by Lindsay and Wells (1985) to reduce the likelihood of making a relative judgment and encourage the use of making an absolute judgment. The use of a sequential lineup encourages absolute judgments because the witness must compare each picture to their memory of the perpetrator. Once the witness decides the photograph is not of the perpetrator, they are not able to go back and change their decision. Additionally, they are not aware of how many photos they will be shown. Therefore, although the eyewitness may view a lineup member and reason that this target resembles the perpetrator more than a lineup member previously presented, they cannot be sure that any subsequent lineup member (i.e., yet to be viewed) will not be a better match to the perpetrator than the target currently being viewed (Wells, 1993).

The debate as to which type of lineup procedure provides the most accurate identifications is ongoing. After its conception, the sequential procedure was popularized due to its ability to reduce false positive identifications compared to simultaneous lineups (e.g., Cutler

& Penrod, 1988; Lindsay, Lea, & Fulford, 1991; Lindsay & Wells, 1985; Melara, DeWitt-Rickards, & O'Brien, 1989; Parker & Ryan, 1993; Sporer, 1993). Additionally, meta-analysis conducted by Steblay, Dysart, Fulero, and Lindsay (2001) suggested that the sequential lineup, in comparison to the simultaneous lineup, was superior in almost all scenarios. Although, recent research contests the superiority of the sequential lineup, asserting that the simultaneous lineup is at least as accurate, if not more accurate, than the sequential (e.g., Dobolyi & Dodson, 2013; Mickes, Flowe, & Wixted, 2012). Extrapolating to the real world, the simultaneous lineup may result in more innocent suspects being convicted of crimes they did not commit, whereas the sequential lineup may result in more guilty suspects going free (Wells, 2014).

Although there have been arguments for both simultaneous and sequential, (see Amendola & Wixted, 2015a; Amendola & Wixted, 2015b; Gronlund, Carlson, Dailey, & Goodsell, 2009; Steblay, Dysart, & Wells, 2015; Wells, Dysart, & Steblay, 2015) it appears that the sequential lineup is an issue for some populations (e.g., children) and under some situations (e.g., when there is multiple perpetrators; Pozzulo & Lindsay, 1998). For example, the sequential lineup, in comparison to the simultaneous lineup, leads to larger discrepancies between children's and adults' correct rejections, with children making fewer rejections and adults making more (Pozzulo & Lindsay, 1998). Since it would be of interest in the future to extend the current program of research to include children, the simultaneous procedure was used in the present study.

Familiarity

Recall of familiar-strangers. Studies examining eyewitness recall have focused on the recall of strangers; no known studies to date have examined recall accuracy of familiar-strangers. In order to provide some insight into the recall accuracy of familiar-strangers, it may be

beneficial to look to research examining event repetition and exposure duration. An eyewitness may be exposed to a target very briefly (i.e., a few seconds/on one occasion) or for a longer period of time (i.e., several hours/several different occasions). Since prior exposure is an important aspect of familiarity (Mandler, 2008) and repeated and longer exposure would likely affect recall memory, a review of both the event repetition and exposure duration research is necessary.

Event repetition. With repeated exposure to a target, eyewitnesses will have the ability to see the target at various angles, with differing facial expressions, and in a variety of contexts. It has been demonstrated that this kind of face repetition tends to increase recognition accuracy (e.g., Bruce, Burton, & Hancock, 2007; Memon, Hope, & Bull, 2003). In contrast, little research has looked at face repetition and its influence on recall accuracy; the research available tends to look at recall accuracy after event repetition (i.e., being repeatedly exposed to similar events).

In a study conducted by McNichol, Shute, and Tucker (1999) recall accuracy was examined for children ($n = 58$) who experienced an event once versus children who experienced an event several times. In this study, a researcher spent 10-15 minutes alone with a child playing games and developing rapport either on one occasion (single event group) or on three occasions (multiple event group). Each event in the multiple event group was a week apart and had various items that changed from one event to the next, and various items that remained constant. Three weeks after the initial event, an interview was conducted with the child that mimicked an investigative interview that would occur after alleged child abuse. Initially the children were asked to recall everything about the “special events” taking place. Results indicated that the children in the repeated events group recalled significantly more items than the single event group. For the items that stayed the same, the children in the repeated-events group had

significantly less errors (commission and omission) than the children in the single event group. However, for items that changed, the repeated events group had significantly more errors than the single event group. The authors note that the errors made by the children in the repeated events conditions were due to interference errors as the children seemed to confuse the timing of details that changed across events. Regardless, these results suggest that repeated exposure to the same details of an event increases memory for those details. Although this study provides valuable information about event repetition, it was conducted with a child sample; it would be of interest to look to event repetition in adults since the current program of research utilizes an adult sample.

Theunissen, Meyer, Memon, and Weinsheimer (2017) conducted a recent study looking at event repetition with adults (ages 18-35). In this study they were looking at whether memory was better for participants who witnessed multiple emotional traffic accidents or participants who witnessed only one traffic accident. The results indicated that the participants in the single-event condition had significantly higher accuracy (i.e., reported more correct details and fewer commissions) than the participants in the repeated-event condition. An important difference to note between this study and the McNichol et al. (1999) study was that in this study there was no consistency between the events besides them being “emotional traffic accidents”. This would likely influence recall accuracy as no important details stayed the same.

Taken together the results of these studies seem to demonstrate that repeated exposure to an event enhances recall accuracy of these events when details of the events remain relatively stable. When details change, or there is no consistency between the events, repeated exposure to events may lead to source confusion resulting in errors in recall. Possible explanations for these

findings will be discussed when reviewing the relevant theory on eyewitness recall (see below: Theories of recall memory).

Exposure duration. As previously mentioned, an eyewitness can be exposed to a target for varying lengths of time and since longer exposure would likely affect recall memory, a review of exposure duration research is required. Although the majority of exposure duration literature involves recognition memory, there is limited research in which recall memory is assessed. Yarmey, Jacob, and Porter (2002) conducted a study in which 320 men and women were stopped on the street by a stranger for either 5 seconds or 30 seconds. Two minutes following this encounter the participants were approached and asked to answer 14 cued recall questions (seven questions about physical characteristics and seven questions about clothing characteristics). Results demonstrated that the participants who had a longer exposure to the target were more accurate in their descriptions of all seven clothing characteristics and were more accurate in their descriptions of eye colour. There were no significant differences between the two exposure groups for the other six physical characteristic descriptors. Based on this, it seems that longer exposure to a target, as with repeated exposure, may enhance eyewitness recall to *some extent*. Unfortunately, the research available is currently too limited to draw firm conclusions, and therefore looking to theory will provide further insight.

Theories of recall memory. An important theory that may supplement the limited research on eyewitness recall of familiar-strangers is the fuzzy trace theory (FTT; Brainerd & Reyna, 2002). FTT posits that there are two types of memory traces that form about a past event; they are referred to as gist and verbatim traces. These two categories of memory trace can be seen as the extremities of a memory continuum. At one extreme are gist traces, which are fuzzy representations of a past event (hence the name fuzzy-trace). Gist traces are vague, simple

representations that represent only the pattern of recently encoded information (Brainerd & Reyna, 1990). At the other extreme are verbatim traces, which are more detailed accounts of a past event. Verbatim traces are comprehensive, elaborate representations that preserve the meaning of recently encoded information with exactness (Brainerd & Reyna, 1990). FTT asserts that verbatim and gist information is processed, stored, and retrieved in dissociated parallel pathways. Put simply, both verbatim and gist traces are formed from the same stimuli but the way in which they are encoded and retrieved differs.

It is through the storage and retrieval of these two memory traces that help to explain eyewitness recall such that accessing verbatim traces over gist traces can increase or decrease recall accuracy and performance. Specifically, it is hypothesized that accessing verbatim traces would result in more accurate recall because specific detailed accounts are being activated as opposed to general information activated through gist traces (Reyna & Kiernan, 1994). Concerning relative accessibility, verbatim retrieval is favoured when the same stimulus is repeatedly presented, and gist retrieval is favoured when different variations of a stimuli are presented but no single stimuli is repeated (Brainerd & Reyna, 2002). Research has supported this claim demonstrating that verbatim memories are retrieved when the substance of an event matches the verbatim details of previous experiences (e.g., Reyna & Lloyd, 1997). Additionally, gist memories are retrieved when the semantic content of an event (e.g., the underlying meaning) matches with other previous experiences (e.g., Wolfe, Reyna, & Brainerd, 2005). This could explain why the children in the repeated events group in McNichol et al. (1999) had fewer errors recalling the items that stayed the same than did the children in the single event group. For the items that stayed the same, the children in the repeated events group were able to experience the same stimulus repeatedly and thus were able to retrieve the verbatim memories of those stimuli.

Similarly, this explains why the participants who were repeatedly exposed to similar events (but which contained differing verbatim details) in Theunissen et al. (2017) showed decreased recall accuracy compared to the group who only witnessed one event. In this study the meaning of the repeated events was the same, but the verbatim details differed, and thus gist memories were retrieved.

Using FTT and its relevant research to extrapolate to memory of people, eyewitness recall of familiar-strangers may be driven by verbatim traces. Seeing someone again would match the verbatim information of an earlier experience of meeting this person and verbatim memories would be retrieved. Furthermore, seeing them in a new setting (i.e., committing a crime) would not match other past experiences with them and therefore, gist memories of them are less likely to be accessed. Since the perpetrator has been presented (i.e., seen/encountered) more than once the verbatim information would likely be accessed and a more detailed description of the event and perpetrator would follow. On the other hand, it can be assumed that eyewitness recall of strangers would be guided by gist traces. Because of the amount of new incoming information (i.e., would not match verbatim details of previous experiences), an eyewitness would likely not be able to access strong verbatim traces and would rely on a general idea of what occurred during the criminal event.

Although FTT provides a theoretical framework to explain variations in recall accuracy, it is not without its critics. FTT cannot account for the social and emotional factors that influence the processing of stimuli and events (e.g., Davies, 1995). Recall memory is influenced not only by the cognitive processing that occurs at the time of the event but by a number of physical (e.g., lighting), social (e.g., pressure from others), and emotional (e.g., stress) factors that occur

surrounding the event. However, given the lack of research and alternative theories on recall accuracy, FTT was used cautiously when making hypotheses in the current program of research.

Recognition of familiar-strangers. In typical eyewitness research, an eyewitness paradigm is utilized. This paradigm involves participants viewing a never before seen target engaging in criminal activity such as a theft. After viewing the target, participants are then informed they have witnessed a mock crime and asked to recall details about the target and the crime and to identify the target via a lineup procedure. This paradigm is meant to mimic the process that occurs in actual criminal investigations so that the results can be used to inform legal procedures. Although most eyewitness studies focus on eyewitness performance when identifying a stranger (Pezdek & Stolzenberg, 2014), the majority of real life witnesses to a crime involve a perpetrator they have previously been exposed to and therefore are familiar with (Flowe et al., 2011).

Research has identified that the cognitive and neural systems (Burton, Bruce, & Hancock, 1999; Gobbini & Haxby, 2007; Klatzky & Forrest, 1984) involved in recognizing and examining familiar versus unfamiliar faces are intrinsically different (Althoff & Cohen, 1999; Pezdek & Stolzenberg, 2014). In reviewing the literature on memory for familiar compared to unfamiliar faces, Hancock, Bruce, and Burton (2000) concluded that familiar faces are recognized with little effort, even despite possibly large discrepancies in lighting, angles, expressions, and accessories (or “disguises”) such as beards, glasses or hats. Therefore, it is important to differentially study how an eyewitness would recognize a stranger as opposed to someone familiar to them.

Although there is extensive research examining stranger identifications, there are very few studies that have looked at how familiarity affects recognition. The studies that have looked at familiarity tend to focus on facial recognition without using an eyewitness paradigm (as

described above). In contrast to eyewitness research, facial recognition studies usually consist of a study phase where participants are presented with several faces (e.g., 20 or more) over a short period of time (i.e., seconds). Following the study phase, there is a testing phase where participants are presented with many (sometimes over 500) images of faces and asked to judge whether they believe that face appeared in the study phase or not (e.g., Burton, Wilson, Cowan, & Bruce, 1999; Clutterbuck & Johnston, 2005; Curran & Hancock, 2007). Considering there is little research on memory for familiar-strangers in eyewitness scenarios, research looking at facial recognition will be reviewed.

Facial recognition studies. Burton et al. (1999) examined whether personal familiarity with a target influenced the ability of a participant to recognize images of this target from a low (image) quality security system. In this study, participants were shown a set of 20 videos containing 20 different psychology professors from the University they attend; 10 male professors and 10 female professors. These videos were from a security system camera that is triggered when a person approaches the main entrance of the building. When a person enters or leaves this building, a security light is automatically turned on and the camera records approximately 4 seconds of video. The familiar group consisted of 20 students within the department of psychology who had been taught by all 20 targets. There were two unfamiliar groups; the first was an additional 20 students not in the department of psychology who had not taken any psychology courses, and the second unfamiliar group was 20 experienced police officers. Participants viewed 10 of the 20 video clips and were told they would have to identify the target in the video clips later. There was a short 2-3 second delay between each clip, and a minute rest period after all the videos had been viewed. In the test phase, participants were shown 20 high quality photos and told that half of these people had been present in the previous

video clips. Participants were asked to rate each photo from 1 to 7; 1 being the target was definitely not in the video clips and 7 indicating they were sure the target was in the video clips. Results demonstrated that the familiar group performed significantly better than both the unfamiliar groups. Therefore, the participants who had known the targets beforehand were better able to identify their faces as previously seen in the video clips. Unfortunately, in this study familiarity could not be experimentally controlled. Bruce, Henderson, Newman, and Burton (2001) address this methodological issue by experimentally manipulating familiarity in their study of facial recognition.

Bruce et al. (2001) looked at quality of video and familiarity of a target person in three different experiments. In Experiment 1 they looked at how well participants could recognize targets who were considered to be highly familiar to them (i.e., had been exposed to the targets for extensive periods of time on repeated occasions). They found that participants with high levels of familiarity with the targets performed significantly better on the recognition task than participants who were unfamiliar with the targets. In Experiment 2, they looked at whether low levels of familiarity (i.e., relatively brief prior exposure) would demonstrate the same effect on memory that the high levels of familiarity had on recognition memory. In the first phase of Experiment 2, participants were shown a 30-second video clip of a target moving their face from side to side and smiling. There were 24 videos in total and each participant was shown half (12) of the videos in this phase. The familiar group was shown the 12 videos twice, and the unfamiliar group was shown the videos once. In the test phase, low-quality images of both the familiarized and unseen targets were paired with a photograph of either the same target or a distractor. Therefore, when being tested, they saw a still from the video they had previously watched as well as one high quality picture; either the target or a similar looking distractor. They were then

asked whether they thought these two images were the same people, or two different people. The results indicated that there were no significant differences between the familiar and unfamiliar groups on the memory test. Therefore, combining the results from Experiment 1 and Experiment 2, it would seem that only relatively high levels of familiarization with a target increases recognition performance, and that brief (1 minute) exposure to a target is not long enough to create this enhanced recognition effect. With the results of this experiment, the researchers then wanted to examine the possibility that participants in Experiment 2 were not encoding the targets in a significant and deep way. The authors note that when people meet in everyday life, they usually interact with each other in meaningful ways. Therefore, in their final experiment they wanted to look at performance on a facial recognition task comparing participants who have had no exposure (unfamiliar group), participants who have had brief exposure (brief exposure group; analogous to the brief exposure group from Experiment 2), and participants who are viewed as socially exposed (social exposure group; they viewed the same videos but watched in pairs and were encouraged to talk about the people who appeared in the videos). The results indicated that the social exposure group performed significantly better than both the other two groups, and consistent with their results from Experiment 2, they found that the brief exposure group did not perform significantly better than the unfamiliar group. Therefore, Experiment 3 suggests that it is likely the type of exposure rather than the length of exposure that will lead to the memory advantage associated with familiarity. Research has demonstrated that focusing on personality features as opposed to physical attributes encourages better memory for faces (cf. Bower & Karlin, 1974; Patterson & Baddeley, 1977; Winograd, 1981), and it may be that this deep processing was only available in the socially exposed condition of the experiment. In line with this, research has also demonstrated that providing even very limited social information about a

target can lead to the improved recognition associated with familiarity (e.g., Curran & Hancock, 2007; Yovel & Paller, 2004).

Research on facial recognition seems to show that familiar faces are better recognized than never before seen faces. Bruce et al. (2001) added to this by demonstrating that being socially exposed might be a necessary condition in order to show the enhanced recognition for familiar faces; although this need not be a large interaction (e.g., Curran & Hancock, 2007; Yovel & Paller, 2004). Therefore, in the current program of research participants were not only shown faces of the targets during the familiarization phase, but the pictures were also paired with a 1 minute audio overlay in which the target describes aspects of their occupation in order to try and mimic a real-world social exposure.

Eyewitness studies. One study that did make use of an eyewitness paradigm to examine familiarity was conducted by Steblay, Dietrich, Ryan, Raczynski, and James (2011). In this study participants were shown a video of an actor stealing a purse. The actor in the video was a recent graduate of the school in which the study was conducted and therefore they used this to assess participant's familiarity with the target. In this study the participants were simply asked whether they had seen the target before today and dichotomized these groups into *yes*- they were familiar, and *no*- they had never seen them before. After viewing the video, participants were presented with a sequential photo lineup and had the option of reviewing the photo lineup once or twice. Therefore, the analyses addressed the witness' decisions at two points, their first decision and a final decision if the participant decided to view the lineup more than once. The results indicated that participants who were familiar with the target were more likely to correctly identify him from the lineup than were the participants for whom the target was a stranger, at first decision and at final decision. In addition, when the target was absent from the lineup, the participants

who were familiar were significantly less likely to pick a photo from the TA lineup (i.e., make a false positive identification) than were the unfamiliar participants. Therefore, this study demonstrates that familiarity with a target significantly improves performance in TA and TP lineups. This study, although helpful in providing a basis of evidence for eyewitness memory of familiar targets, has its limitations for its use in forensic settings. The main goal of this study was not to look at familiarity, thus the researchers did not assess the level of familiarity each participant had with the target. Familiarity with the target included people who knew him personally, interacted with him, or just saw him on campus. It is likely that differing amounts of familiarity would greatly impact performance in this type of identification, and therefore it is important to examine different levels of familiarity separately. Additionally, the researchers assume that the participants who are reporting being familiar with the target are correct in recognizing this person, and that the participants who are reporting being unfamiliar with the target, actually have not seen this person before. This could be an issue if participants are incorrect in their assumption of familiarity; a recent study by Pezdek and Stolzenberg (2014) questions whether subjective ratings of familiarity are actually diagnostic of prior contact.

Pezdek and Stolzenberg (2014) examined student's ability to recognize other students that had recently graduated from their high school. The participants consisted of students from two small private schools within the Los Angeles Metropolitan area. Since both schools were on small campuses (approximate enrolment of 750) the likelihood of being familiar with other students at their respective schools was quite high. Participants were presented with 40 yearbook pictures consisting of 20 students who graduated from their school one year previously (familiar targets) and 20 students from the other school (unfamiliar targets). The results indicated that 23% of the unfamiliar faces were misidentified as being familiar and only 42% of the familiar faces

were correctly identified as being familiar. Further analysis indicated that familiar faces were significantly more likely to be identified as familiar than were the unfamiliar faces. However, the authors note that the high diagnostic error rate (i.e., the amount of faces classified as familiar when they had not been seen before) suggests that familiarity judgments should be used cautiously in forensic settings. Similar to the Steblay et al. (2011) study, this study provides valuable information regarding adults' ability to diagnose familiarity, however because familiarity could not be experimentally controlled these studies cannot contribute fully to our understanding of the effects of familiarity on identification accuracy.

Exposure duration. Alternatively, there have been a number of studies that have used an eyewitness paradigm to examine the role of exposure duration on recognition accuracy. Given that prior exposure is an important aspect of familiarity (Mandler, 2008) and, as with recall memory, longer exposure would likely affect recognition memory, a review of the exposure duration research is necessary. Although most of the exposure duration literature has focused on child samples, there are a few studies that utilize adult samples. For example, Memon et al. (2003) examined a sample of 164 younger (ages 17-25) and older (ages 59-81) adults to see if longer exposure to a target in a simulated eyewitness setting would increase recognition accuracy in TA and TP lineups. Half the participants were exposed to the target for 12 seconds (short duration) and half were exposed for 45 seconds (long duration). Results demonstrated that longer exposure to the target increased recognition accuracy compared to the shorter exposure duration. Specifically, participants in the longer exposure duration conditions produced more correct identifications while the participants in the short exposure conditions made more false identifications.

In another study looking at exposure duration on adult's recognition memory, Horry, Halford, Brewer, Milne, and Bull (2014) examined 295 case files from an England police force. Within these case files, 833 identification lineups were administered to eyewitnesses. The rate of suspect identification was significantly higher for witnesses who were exposed to the perpetrator for a longer period of time (i.e., more than 60 seconds) than for witnesses who were exposed to the perpetrator for a shorter period of time (i.e., less than 60 seconds).

Consistent with the archival results from Horry et al. (2014); Bornstein, Deffenbacher, Penrod and McGorty (2012) conducted a meta-analysis looking at the influence of exposure time on identification accuracy. Extending the work of a previous meta-analysis conducted by Shapiro and Penrod (1986), the authors examined 25 facial recognition studies finding that longer exposure to a perpetrator resulted in significantly greater identification accuracy than did shorter exposure durations. However, Bornstein and colleagues (2012) note that the effect of exposure time was non-linear and that the effect of increased exposure may be most influential when exposure time is relatively short. Specifically, the largest effects appear when the exposure time is less than 30 seconds; once the exposure time exceeds 30 seconds, any further increase needs to be quite substantial to produce equivalent effects.

Theories and regularities in recognition memory. When discussing recognition memory it is important to discuss an empirical regularity known as the *mirror effect*. A mirror effect is said to occur when one experimental condition produces more hits (i.e., true positives) and fewer false alarms (i.e., false positives) than another condition (Cary & Reder, 2003). Dual-process theories of recognition memory are said to account for these mirror effects (Glanzer & Adams, 1985). Dual-process theories generally state that there are two processes involved in recognition: familiarity and recollection. Familiarity can be described as a perceptual process involving

pattern matching and is the mechanism responsible for the experience of prior occurrence.

Recollection is a general recall process that aids in the retrieval of the stored memory representation and facilitates the ability to identify the target item or event (Mandler, Pearlstone, & Koopmans, 1969).

According to dual-process theories, the hit component of mirror effects are a result of the ease with which one condition produces recollection-based judgments, while the false alarm component is a result of the conditions' differential dependence on familiarity-based judgements (Cary & Reder, 2003). Recollection-based recognition is the retrieval of the stored memory trace of a previously encoded event. When recollection is unsuccessful, the recognition judgment now relies on familiarity. With a familiarity judgement, a positive recognition may be made if the test item seems highly familiar (Cary & Reder, 2003). One mirror effect of particular importance for discussing familiar-stranger identifications is the Strength Based Mirror Effect (SBME) which is produced by manipulating encoding strength (Stretch & Wixted, 1998). Encoding strength can be manipulated by adjusting the duration of time an item is presented or adjusting the number of presentations of an item (i.e., the number or repetitions). When examining the SBME, researchers generally test a strong encoding condition against a weak encoding condition. In the strong condition the test items are presented for longer periods of time (or alternatively presented several times) where in the weak condition the test items are presented for shorter periods of time (or fewer times; e.g., Stretch & Wixton, 1998). According to dual-process theories, the more times an item is presented, the stronger the memory trace becomes. Therefore, it should be easier for participants to recollect an item in the strong condition than an item in the weak condition (Cary & Reder, 2003). As a result, there will be more recollection based hits in the strong condition than in the weak condition at testing, thus reflecting the hit component of the SBME

(Stretch & Wixted, 1998). On the other hand, people will rely more on familiarity in the weak condition adopting a more liberal threshold at testing, resulting in the false alarm component of the SBME. Overall, research has supported this claim demonstrating that items from a strong encoding condition are recognized at a higher rate than items from a weak encoding condition (e.g., Ratcliff, Clark, & Shiffrin, 1990). Extending this to the current study, it is likely that seeing the target more than once would create a stronger representation of this person and the participants will rely more on recollection than on familiarity. In this way, these participants will have more hits (i.e., correct identifications) and fewer false alarms (i.e., false positive identifications). On the other hand, participants who are unfamiliar with the target and only get to see them during the commission of the crime will likely have to rely on a familiarity judgement. Therefore, they will likely have fewer hits and more false alarms than those participants in the familiar condition.

Overall, research and the relevant theory on recognition memory and eyewitness identification suggest that longer exposure duration and increased frequency of exposure may result in improved recognition accuracy (Gross & Hayne, 1996; Leippe, Romanczyk, & Manion, 1991); specifically, when a social exchange takes place (Bruce et al., 2001). Based on theory and previous research, it may be predicted that familiar-stranger eyewitnesses would demonstrate higher correct identification rates (in TP lineups) and higher correct rejections rates (in TA lineups) than stranger eyewitnesses.

Stereotypes

There are a number of social factors that can affect memory and retrieval of a familiar-stranger. One of which is peoples' pre-conceived notions of what is appropriate for given social groups, i.e., stereotypes.

Stereotypes have been defined as beliefs or general knowledge structures regarding the characteristics, traits, and behaviours of people within various out-groups (Hilton & von Hippel, 1996). Stereotype thinking serves many purposes that reflect various cognitive and motivational processes (Hilton & von Hippel, 1996). Stereotyping can emerge as a way of simplifying the demands of the perceiver (Bodenhausen, Kramer, & Süsler, 1994; Bodenhausen, Sheppard, & Kramer, 1994; Macrae, Milne, & Bodenhausen, 1994). Additionally, stereotypes make information processing easier as they allow the perceiver to rely on previously acquired knowledge rather than incoming information. Stereotypes also may arise in response to situational factors such as differing social and gender roles (Eagly, 1995), group conflicts (Robinson, Keltner, Ward, & Ross, 1995), and shifts in power (Fiske, 1993). In other instances, stereotypes arise as a way of legitimizing the existing state of affairs (Jost & Banaji, 1994), or as a way to satisfy the need for social identity (Hogg & Abrams, 1988).

When examining stereotypes, it is important to discuss them according to how they emerge (i.e., stereotype formation), are maintained (i.e., stereotype maintenance), and applied (i.e., stereotype application; Hilton & von Hippel, 1996). Stereotype formation occurs when perceivers use generalizations and associations about the world around them in order to create mental shortcuts. Once formed, stereotypes are then maintained through various reinforcements in everyday life. Stereotype application involves applying these formed stereotypes to out-group members. This process from formation to application becomes automatic and occurs in various social scenarios and settings (Blair & Banaji, 1996).

Gender stereotypes. Research has shown that categorizing as male or female is most often the first grouping that people sort others into in social contexts (Blair & Banaji, 1996; Brewer & Lui, 1989; Stangor & McMillan, 1992). Subsequent categorizations within other social

dimensions are embedded within the perceiver's prior understanding of themselves and others as male or female (Brewer & Lui, 1989; Stangor & McMillan, 1992). This principle distinction facilitates the formation and maintenance of gender stereotypes (Ridgeway & Correll, 2004). In order to get a more comprehensive understanding of gender stereotypes, and more specifically, occupational gender stereotypes, a brief review of social role theory will be provided.

Social role theory and occupational stereotypes. Social role theory states that sex differences that occur in social interactions follow from the typical characteristics of roles generally held by women versus men (Eagly, Wood, & Diekmann, 2000). One principle guiding these differences is a result of the differing activities associated with the common roles of each sex. Women and men adjust to their roles by acquiring the skills and social tendencies necessary to succeed in their role. For example, with the homemaker-provider division of labour, females learn domestic skills such as cooking and cleaning, while males acquire skills that are useful in the paid economy. Particularly important to the role of females is the assignment of most childrearing responsibilities; this task requires women to be nurturing and caring to children and other dependents (Eagly et al., 2000). This requirement promotes the development of superior social skills and the ability to associate nonverbally. In contrast, men's adaption to the employment role encourages more assertive and independent tendencies (Eagly & Steffen, 1984).

Therefore, gender roles are the result of men and women accommodating their behaviour to succeed in their sex-typical occupations and family roles (Eagly et al., 2000). Thus, as women have been commonly known to perform roles that require communal, domestic, and subordinate behaviours, these characteristics become stereotypic of females and their gender role. In contrast, because men have been associated with roles that require assertive and dominant behaviours,

these tendencies become stereotypic for males and their gender role (Eagly et al., 2000). As a result of gender role stereotypes and the traits associated with each gender, occupations have also become gender stereotyped. Gender role stereotypes influence what people believe to be appropriate occupational roles for men and women and occupational gender stereotypes are produced. Occupational gender stereotypes are generalizations about the gender-relevant features associated with various occupations (Glick, Wilk, & Perreault, 1995; Shinar, 1975). In line with gender roles, male stereotyped occupations are those that require decisiveness, coldness, and toughness; while female stereotyped occupations are those that require patience, nurturing, and caring (Heilman, 2001).

Occupational stereotypes may be explicit or implicit in nature (Greenwald & Banaji, 1995). Earlier research examining occupational gender stereotypes has focused on explicit processes where stereotypes are treated as a collection of characteristics or attitudes that the participant consciously associates with members of certain groups (Beggs & Doolittle, 1993; Shinar, 1975). More recent research has recognized that implicit processes are also important in stereotyping (e.g., White & White, 2006). Greenwald and Banaji (1995) have defined implicit stereotypes “as the introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category” (p. 15). These implicit cognitive thoughts are the result of remaining explicit beliefs that, although consciously disregarded, continue to influence thoughts and perceptions (White & White, 2006). This is often beyond conscious control (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) as research has demonstrated that appropriate priming can trigger implicit stereotyping even among those who explicitly disagree with group categorization (Banaji, Hardin, & Rothman, 1993). Stereotypes

may thus continue to influence perceptions implicitly, even if they are not present explicitly (Kunda & Spencer, 2003).

White and White (2006) looked at occupational stereotypes by using the Implicit Association Test (IAT). The IAT assesses implicit stereotypes by measuring unconscious associations between various concepts (White & White, 2006). This procedure involved having participants discriminate between two targets. In one phase, the targets were occupation-gender pairings reflecting consistent stereotypes. For example, “Engineer or Male” would appear on the left side and “Elementary School Teacher or Female” would appear on the right side. In another phase, the occupation-gender pairings were reversed such that the stereotypical occupation was paired with the inconsistent gender (e.g., “Engineer or Female” and “Elementary School Teacher or Male”). In both phases, a male or female name would appear below the target pairings and the task for the participant was to decide if the name was associated with the target on the right (by pressing the right key) or was associated with the target on the left (by pressing the left key). After completion of the IAT, participants were asked to rate on a 7-point Likert-scale how masculine and feminine they thought each occupation was (i.e., explicit stereotype ratings). The implicit stereotypes are inferred from the response speeds in each of the IAT’s tasks. The response speeds to the stereotype consistent pairs (i.e., Engineer or Male and Elementary School Teacher or Female) should be faster than the response speeds to the inconsistent pairs (i.e., Engineer or Female and Elementary School Teacher or Male). If this were the case, this would demonstrate that the stereotype consistent pairs are more easily associated than the stereotype inconsistent pairs. Looking at the IAT results, they found that participants responded significantly faster to the consistent pairing than they did the inconsistent pairings. Additionally, looking at the explicit ratings, engineer was rated the most masculine ($M = 2.3$), and elementary

school teacher being the most feminine ($M = 5.6$). Therefore, it appears that the occupations utilized in this study were gender stereotyped. Engineer was implicitly as well as explicitly seen as masculine and elementary school teacher was implicitly as well as explicitly seen as feminine.

Although implicit measures may provide a more accurate representation of an individual's occupational stereotypes, the bulk of research on gender stereotypes has been conducted through explicit measures. While explicit measures may underestimate occupational stereotypes, these measures are not likely to overestimate their effect, as explicit measures are more prone to social desirability and self-presentational effects (White & White, 2006).

An early study of explicit occupational stereotypes was conducted by Shinar (1975) where he asked 120 college students to rate the extent to which they perceived various occupations to be masculine, feminine, or gender neutral on a 7-point Likert-scale. Beggs and Doolittle (1993) replicated the study finding similar results. Although in the replication the mean ratings for some occupations were considered more gender neutral than previously rated in Shinar's study, in both studies many occupations were highly and significantly gender typed. For example, in both studies, auto mechanic, carpenter, and construction worker were rated as significantly masculine, while secretary, receptionist, and registered nurse were rated as significantly feminine. Although these two studies are dated, in a recent study by DiDonato and Strough (2013) they found that college students continued to rate stereotypically masculine jobs as appropriate for men, and stereotypically feminine jobs as appropriate for women.

In the current program of research, material testing of several occupations was used to determine that the occupations chosen for the targets were considered stereotypical of a certain gender (see Appendix N). This was done using an explicit measure, and as previously alluded to,

occupations rated as highly gender stereotyped explicitly would likely be rated just as high, if not higher implicitly (Kunda & Spencer, 2003).

Stereotype consistent vs. stereotype inconsistent. Information acquired in social settings can either be consistent with our pre-existing beliefs and stereotypes, or inconsistent with our beliefs and stereotypes (Crocker, Hanna, & Weber, 1983). Research on whether stereotype consistent or stereotype inconsistent information has an advantage in memory has been mixed. Much research has found there to be a memory advantage for stereotype consistent information (Cano et al., 1991; Fyock & Stangor, 1994; Koomen & Dijker, 1997; Stangor & Ruble, 1989) while several other studies have found there to be a memory advantage for stereotype inconsistent information (Bargh & Thein, 1985; Driscoll, Hamilton, & Sorrentino, 1991; Hastie & Kumar, 1979; Sherman, Stroessner, Conrey, & Azam, 2005; Srull, 1981; Srull, Lichtenstein, & Rothbart, 1985). More recently, however, it has been discovered that these inconsistencies were due to methodological differences between studies.

Cognitive load. The most notable reason for the inconsistencies in results was the discovery that participants under a higher cognitive load performed differently than participants under a lower cognitive load. Cognitive load refers to the amount of effort being used in working memory (Sweller, van Merriënboer, & Paas, 1998). Cognitive Load Theory (CLT) assumes that working memory is limited. The theory highlights that working memory has capacity and duration limitations and these limitations only apply to novel information obtained through sensory memory (van Merriënboer & Sweller, 2005). How easily information may be processed through working memory is at the core of CLT. Cognitive load can either be affected by the inherent qualities of a task (i.e., intrinsic cognitive load) or by the way in which a task is presented (i.e., extraneous cognitive load; van Merriënboer & Sweller, 2005). Intrinsic cognitive

load is dependent on the number of details that need to be processed concurrently in working memory. Extraneous cognitive load is a result of the way in which information is presented; the presentation can either reduce or increase the amount of resources required to process the information (van Merriënboer & Sweller, 2005). Therefore, the amount of load (intrinsic or extraneous) experienced during a task will affect memory and performance on said task. If there are many novel elements being processed during a task, memory and performance on that task will be lower than if the task was a familiar one or one with less elements. Therefore, it is from examining CLT that researchers on memory for stereotyped information suggest that the inconsistent findings may be attributed to differences in the amount of cognitive load.

From reviewing the literature on CLT, it was suggested that memory for stereotype consistent and inconsistent information is differentially impacted by the amount of cognitive load experienced. This suggestion was based on the discovery that participants who were under high cognitive load during the study/learning phase of the experiment (e.g., had to remember an 8-digit number) were better able to recall stereotype consistent details than stereotype inconsistent details (Dijksterhuis & van Knippenberg, 1995; Macrae, Hewstone, & Griffiths, 1993; Sherman & Hamilton, 1994). Sherman and Frost (2000) tested this theory by manipulating participant's cognitive load during a recall task. Participants were told they would be learning about a man named Bob. Half the participants were told that Bob was a priest, and half were told that Bob was a skinhead. Each participant was presented with 30 behaviours that Bob had reportedly engaged in: 10 were kind, 10 were neutral, and 10 were unkind behaviours. The kind behaviours were considered stereotype consistent for the priest (and inconsistent for the skinhead), while the unkind behaviours were considered stereotype consistent for the skinhead (and inconsistent for the priest). Before the presentation of the behaviours, half the participants were asked to

remember an 8-digit number (i.e., an intrinsic cognitive load) during the presentation of the behaviours while the other half of the participants were not given any additional tasks. Following the presentation of the behaviours and a five-minute filler task, participants were asked to recall as many behaviours as possible. The free recall task demonstrated that participants under cognitive load (i.e., the participants who had to remember the 8-digit number) remembered more stereotype consistent words than stereotype inconsistent words. For the participants not under a cognitive load, more stereotype inconsistent words were remembered than stereotype consistent words (Sherman & Frost, 2000). Since this discovery it seems the literature has resolved its debate and most researchers in the field agree that under conditions in which cognitive load is low, stereotype inconsistent information is better remembered than stereotype consistent information. And conversely, when cognitive load is high, stereotype consistent information is better remembered than stereotype inconsistent information.

The reason for this memory advantage for stereotype inconsistent information (when cognitive load is low) seems somewhat paradoxical. Human memory has limited resources to process and store information and because stereotypes are mental shortcuts in person formation it should follow that individuals would remember traits that are stereotypical and follow their own mental schemas (Araya, Akrami, & Ekehammar, 2003). Since this does not appear to be the case, the question follows of why memory is not always biased for stereotypically consistent information?

One reason for stereotype inconsistent memory advantage under lower cognitive load could be due to the von Restorff effect or the distinctiveness effect. The von Restorff effect states that people attend to and remember more about stimuli that are unique or distinct in a sequence of more related items (von Restorff, 1933). This could explain why memory is enhanced for

stereotype inconsistent information under full attention (i.e., when cognitive load is low). When presented with items that do not match expectations and when afforded the ability to attend differentially, perceivers focus more on the inconsistent information because it is unique and unexpected. Research looking at the difference between people under a higher cognitive load compared to people under a lower cognitive load supports this notion. Multiple researchers have demonstrated that the memory advantage for stereotype inconsistent information reverses to a memory advantage for consistent information when under increased cognitive load (Bargh & Thein, 1985; Dijksterhuis & van Knippenberg, 1995; Macrae et al., 1993; Sherman et al., 2005; Srull, 1981). Research indicates that people typically spend more time processing stereotype inconsistent (versus consistent) information (e.g., Stangor & McMillan, 1992; Stern, Marrs, Millar, & Cole, 1984) and that stereotype inconsistent information is processed more elaborately. Specifically, people are likely to encode the general theme of consistent or familiar events and disregard its specific details (e.g., Johnston & Hawley, 1994). Rather, the specific details can be recreated and filled in later by extracting from existing knowledge structures (e.g., Bruner, 1957; Koriat, Goldsmith, & Pansky, 2000; Sherman, Lee, Bessenoff, & Frost, 1998).

Another possible reason for the memory advantage for stereotype inconsistent information was proposed by Srull et al. (1985) who theorized that inconsistent stereotype information has more associative paths in memory. Srull et al. suggested that stereotype consistent information is linked only with other stereotype consistent information; whereas stereotype inconsistent information is linked with both consistent and inconsistent stereotype information. This theory was supported by research conducted by Srull (1981) and Srull et al. (1985) where the amount of stereotype inconsistent information was experimentally manipulated during the study phase of the experiments. Srull and colleagues (1981; 1985) found that as the

amount of stereotype inconsistent information increased, the more stereotype consistent information was recalled. Thus, stereotype inconsistent information may receive considerable processing during encoding as the perceiver tries to understand and make sense of the contradictory information (e.g., Srull & Wyer, 1989; Stangor & Duan, 1991). Accordingly, enhanced recall for inconsistent details would depend on the amount of attentional resources available that could allow such extensive encoding (Araya et al., 2003).

Inconsistent vs. consistent gender stereotypes. The only known studies to date that have looked at how consistent versus inconsistent gender role stereotypes influence memory have all been conducted using child samples. Research on whether consistent or inconsistent gender role stereotypes have an advantage in children's memory has been mixed. While some earlier research has found a memory advantage for stereotype inconsistent gender roles (e.g., Jennings, 1975) others have found a memory advantage for consistent gender roles (e.g., Koblinsky, Cruse, & Sugawara, 1978; Liben & Signorella, 1980). Additionally, more recent research has found no difference for memory of inconsistent versus consistent gender roles (Wilbourn & Kee, 2010). These inconsistencies could be due to the fact that stereotype formation in development is not a straight forward process. Specifically, stereotype beliefs are not linear in development but instead show phase life shifts (Ullian, 1976); children tend to go from being quite rigid in their application of gender role stereotypes, to being more flexible as they get older, and again shift back to more rigid and strict beliefs as they approach adolescence (Martin, Ruble, & Szkrybalo, 2002; Ruble & Martin, 1998; Welch-Ross & Schmidt, 1996). Although the literature generally agrees that various phases exist, there is disagreement about exactly when the shifts occur in development. Therefore, it seems stereotype formation and application is quite complicated in

children, thus making it difficult to identify how children react to and remember consistent and inconsistent gender role information.

Although there is no previous research examining adult's memory for gender role consistent and inconsistent information, this effect may be inferred from the literature looking at other stereotype information. Thus, based on previous research it may be predicted that under conditions in which attention is not restricted, inconsistent gender role stereotypes are likely to have an advantage in adult's memory over consistent gender role stereotypes (Sherman & Frost, 2000).

Stereotypes and eyewitness research. The research available that has looked at stereotypes in eyewitness paradigms focuses on stereotypes activated on the part of the witness during the commission of the crime and later when recalling the crime (e.g., Osborne & Davies, 2013). In these situations, the participant is exposed to a never before seen target, and thus the stereotype activation occurs while watching the crime or thinking back to the crime. This is very different from the current program of research where the target is a previously encountered person that the eyewitness has possibly already formed memories of.

In the previous studies examining never before seen targets, the relevant stereotype information usually is regarding people's cognitive schemas about what typically occurs in crimes and about criminal's physical characteristics (e.g., Bargh, 1999; Blair, Judd, & Fallman, 2004; Gilbert & Hixon, 1991; Osborne & Davies, 2013). There is evidence that people have firm stereotypes about criminals including what is typical of their physical appearance and personality traits (Bowers & MacLin, 2004; MacLin & Herrera, 2006; Madriz, 1997; Reed & Reed, 1973). For example, stereotypes of an adult criminal by college students consists of them being male, Caucasian, ranging in age from 16-35 years, having dark hair, typically masculine physical

characteristics (i.e., short hair, average build, tall in height) and typically masculine personality traits (e.g., aggressive; Bowers & MacLin, 2004; Greenberg, Westcott, & Bailey, 1998; MacLin & Herrera, 2006). Eyewitness situations would likely be one of high cognitive demand (i.e., high cognitive load), there are many novel items being processed through working memory as this is likely an unusual, high intensity situation. Therefore, it would follow that eyewitness recall is more accurate for information that is stereotype consistent than stereotype inconsistent. This assumption does appear to be supported by research. Research that has manipulated the *type of crime* to be stereotype consistent versus inconsistent for the target (e.g., Osborne & Davies, 2013) as well as research that has manipulated the *target* to be stereotype consistent versus inconsistent (Shapiro, 2009) has shown that more stereotype consistent information is recalled/recognized than stereotype inconsistent information. It has been suggested that in eyewitness situations, participants fail to process stereotype inconsistent details, instead focusing on stereotype consistent details. Eyewitnesses will fill in the gaps in their memory with stereotypically consistent but inaccurate details about the perpetrator and the crime (Shapiro, 2009). Accordingly, eyewitnesses may provide incomplete reports by omitting stereotype inconsistent details or produce false reports by changing inconsistent details to reflect their stereotypes (Harris, Lee, Hensely, & Schoen, 1988; Holst & Pezdek, 1992; Smith & Studebaker, 1996).

Although this assumption follows for crimes involving a stranger eyewitness, there may be a different outcome when the witness has previously encountered the target. Given that the target has already been introduced, there might not be any activation of stereotypes at the time of the crime, nor may the witness rely on their schemas of what they believe the perpetrator should look like. What effect would a previously formed stereotype of this individual have on a person's

memory for them later? Because the idea of a pre-formed stereotype necessarily requires familiarity with the target, a discussion of stereotypes and eyewitness memory of familiar targets will follow. Specifically, someone met briefly (e.g., a familiar-stranger) as this type of person formation is quick and increases the likelihood of stereotype formation (Kaplan, Wanshula, & Zanna, 1993).

Stereotypes and familiar-strangers. Based on the research reviewed, familiarity with a perpetrator is likely to affect how an eyewitness will recall and recognize the perpetrator (e.g., Bruce et al., 2001; Gross & Hayne, 1996; Leippe et al., 1991; Yarmey et al., 2002).

Familiarization with an individual fosters person formation which requires the use of mental shortcuts as a way to deal with the limited capacity of human memory. When individuals are pressed for time or have a quick meeting with someone—as with the concept of a familiar-stranger—they are likely to rely on stereotypes for efficiency (Kaplan et al., 1993). Thus, in a sense, familiar-strangers and stereotyping go hand in hand. Overall, there has been no research to date that has looked at factors that may be affecting eyewitness memory of familiar-strangers. In order to better understand recall and recognition of familiar-strangers, it is necessary to look at factors that may be affecting this relationship. The current program of research will be the first to address one key factor in person formation, stereotypes (specifically, occupational gender stereotypes), and their impact on the relationship between familiarity and eyewitness memory.

In the current study, the participants in the familiar condition were primed with stereotype consistent or stereotype inconsistent information about a target before witnessing them engage in a crime. This was done via an audio-recording where a male or female confederate discusses their occupation. For the stereotype-consistent priming there were two videos, one where a male discussed his job as a construction worker and one where a female

discussed her job as a nurse. For the stereotype-inconsistent priming, there were also two videos, one where a male discussed his job as a nurse, and one where a female discussed her job as a construction worker. In all videos the target described some of the duties and responsibilities associated with their respective jobs. As the participants learned about the targets under optimal conditions (i.e., alone, small room with limited distractions, etc.), cognitive load should be quite low, and therefore it is predicted that participants will have enhanced recall and recognition memory when the target is stereotype inconsistent (i.e., male nurse, female construction worker) compared to when the target is stereotype consistent (i.e., male construction worker, female nurse). Although previous research has found that stranger eyewitnesses typically recall more stereotype consistent information, this is likely because they are under high cognitive load during the first encounter with the target (i.e., during the commission of the crime). Being an eyewitness to a crime is likely to be one that is very stimulating to working memory, there are likely many novel elements that need to be processed resulting in an overload in working memory. The participants in the current study will likely not be under a high cognitive load when first introduced to the target (i.e., while watching the stereotype priming videos), therefore they should be able to retrieve details of the target that were encoded during their initial encounter and use those details to make an identification. In the current study cognitive load was not manipulated directly; however, participant's mental demand was measured after watching the stereotype video and it was predicted that this rating of cognitive load would be low compared to their rating after watching the crime video. I will now briefly discuss various ways to measure cognitive load and why it was necessary to include measures of mental workload.

Measuring cognitive load. The amount of cognitive load a person experiences from a task will depend on their personal characteristics, expertise, and experience with the task (e.g.,

Plass, Kalyuga, & Leutner, 2010). Therefore, individuals will differ on the amount of cognitive load experienced during an event or task. In this sense, it is important to measure an individual's cognitive load as this can greatly differ from one person to another. There are various ways to measure cognitive load and these usually take the form of mental workload assessments. Mental workload assessments evaluate the amount of effort (e.g., physical, mental, temporal) used when completing a task; the more effort that is used, the higher the cognitive load (Paas, Tuovinen, Tabbers, & van Gerven, 2003). These assessments include self-report measures, performance measures, and physiological measures. Self-report measures are participants' subjective ratings of their own mental workload during a task. Performance measures use the participants' performance on a task (e.g., number of errors made, speed of performance, reaction time) to determine how much workload was required for the task at hand. On the other hand, physiological methods (e.g., heart rate, EEG responses, eye fixations) measure participants' physiological responses to the task at hand in order to determine the mental workload. For the purposes of the current study, self-report measures seemed the most appropriate and feasible to include. There are many self-report measures of cognitive load that have been used in the past. One frequently used rating scale is the NASA Task Load Index (TLX; Hart & Staveland, 1988). This scale is a multidimensional scale in which the participants must rate their experience on several subscales (e.g., physical load, frustration). There is a two-step process required to obtain an overall workload rating for the TLX consisting of both weights and ratings. First, the participant must evaluate the contribution of each subscale to the workload of the task. Each subscale is compared to all other subscales (i.e., 15 comparisons in total) and the number of times each scale is selected is tallied. The second step is the participant selecting the numerical ratings for each scale that reflects the magnitude of workload experienced during the task. Each

of the six subscales is represented by a line with 20 equal intervals with the leftmost interval labeled “Very Low” and the rightmost interval labelled “Very High”¹. The overall workload score is then computed by multiplying the ratings for each subscale by the weight assigned for that subscale. Because of the lengthy process involved in acquiring the TLX measure, Byers, Bittner and Hill (1989) proposed a variation to the TLX called the Raw Task Load Index (RTLX) which does not require the first step of paired comparisons. The RTLX simply averages the six subscales to get a measure of workload. Research from Byers and colleagues has demonstrated that the TLX and RTLX have comparable means and standard deviations with correlations above $r = .95$. As the RTLX is easier to administer, the current study chose to use this simplified version of the scale to measure participants’ cognitive load. Another useful measure of cognitive load is the Rating Scale Mental Effort (RSME; Zijlstra & van Doorn, 1985; Zijlstra, 1993). This scale is a unidimensional scale consisting of a continuous line from 0 to 120. Raters mark an ‘X’ on the line at the point which corresponds to the amount of mental effort used during the task. Along the line at several anchor points are statements relating to the amount of invested effort (e.g., almost no effort, extreme effort). This scale is a more direct measure of mental workload and does not measure the more abstract aspects of mental workload as some others do (e.g., NASA TLX). Whether unidimensional scales or multidimensional scales are more useful for assessing workload depends on the information needed for a given task. As this is one of the first eyewitness studies to implement cognitive workload measures it is unclear as to what type of measure may be most appropriate. Therefore, in addition to using the NASA RTLX, the unidimensional RSME was also used to measure participants’ workload. The ratings from these scales were used to assess the amount of workload used when watching the stereotype

¹ With the exception of the fourth subscale “performance” which is anchored by *Good* on the left and *Poor* on the right.

video and the crime video. Although there are no cut-off scores for the RTLX or RSME indicating low, moderate, or high workload, it is predicted that participants' rating of mental effort will be on the lower ends of these scales for the stereotype video. It is also predicted that the crime video will have higher mental workload ratings than the stereotype video.

Overview of the Current Study

In summation, there seems to be a gap in the literature concerning eyewitness identification of familiar-strangers. There is little research to date that has used an eyewitness paradigm to examine whether eyewitnesses' familiarity with the target influences their recall or recognition accuracy. Considering the amount of crime that is committed by people who are known to an eyewitness (see Flowe et al., 2011), research on this topic is critical in order to inform policy and best practice in the criminal justice system.

The lack of empirical research on familiar eyewitnesses may be because the notion of person familiarity is not a well-defined concept (Mandler, 2008). In eyewitness practice, familiarity can be described broadly as *any* prior exposure to a target before the commission of a crime (Pozzulo, 2016). This definition allows for many operationalizations. In the current study, a familiar-stranger eyewitness was operationalized as someone having had one brief exposure to the perpetrator before the commission of the crime. Research has suggested that interactions which include personal information or social interaction may foster a stronger feeling of familiarity (e.g., Bower & Karlin, 1974; Bruce et al., 2001; Patterson & Baddeley, 1977; Winograd, 1981). Therefore, in order to try and develop a familiar-stranger relationship, participants were required to watch a short (1 minute) video of the target where the target provided information about their occupation.

The purpose of the current program was therefore to provide insight into the recall and identification accuracy of familiar-stranger eyewitnesses. Moreover, the current study also was interested in looking at factors that may affect the relationship between eyewitness memory and familiarity; in particular how occupational gender stereotypes may affect recall and recognition of familiar-strangers. Therefore, the present study also manipulated the targets' occupation to either be consistent with gender role stereotypes, or inconsistent with gender role stereotypes.

Hypotheses

1. Participants will provide a greater number of descriptors and more proportionately correct descriptors when they are familiar with the target compared to when they are unfamiliar with the target.
2. When familiar with the target, participants will provide a greater number of descriptors and more proportionately correct descriptors when the target is stereotype inconsistent compared to when the target is stereotype consistent.
3. Participants will be more likely to provide a correct decision in target-present lineups compared to target-absent lineups.
4. a) Participants will be more likely to provide a correct decision (i.e., correct identification in TP lineups or correct rejection in TA lineups) when they are familiar with the target compared to when they are unfamiliar with the target.
4. b) Conditions involving familiar targets' will produce better applied outcomes (i.e., increase perpetrator identifications in TP lineups and decrease innocent suspect identifications in TA lineups) than conditions involving unfamiliar targets.
5. a) When familiar with the target, participants will be more likely to provide a correct decision (i.e., correct identification in TP lineups or correct rejection in TA lineups)

when the target is stereotype inconsistent compared to when the target is stereotype consistent.

5. b) When familiar with the target, conditions involving stereotype inconsistent targets will produce better applied outcomes (i.e., increase perpetrator identifications in TP lineups and decrease innocent suspect identifications in TA lineups) than conditions involving stereotype consistent targets.

Method

Participants

The total number of participants consisted of 287 undergraduate students. However, 25 (8.7%) participants were discarded for their answers to various study check questions (see below: Study check form). In addition, 4 (1.4%) participants were discarded for not being able to adequately communicate with the researcher (i.e., English was poor), and 1 (0.3%) was discarded for being shown an incorrect lineup. The final sample consisted of 257 participants ranging in age from 17 to 49 years ($M = 20.07$, $SD = 3.61$). There were 176 (68.5%) females, 80 (31.1%) males, and 1 (0.4%) participant who identified as transgendered. Although the majority of the sample identified themselves as White/Caucasian (52.5%), participants' ethnicity was quite diverse, including people with Asian (21.9%), Black (14.0%), Latin American (2.3%), Aboriginal (1.6%), and Mixed/Other (7.7%) ethnic backgrounds.

Design

A 3 (familiarity consistency: consistent familiar vs. inconsistent familiar vs. unfamiliar) x 2 (target lineup presence: target present vs. target absent) between-subjects factorial design was used. There were six conditions with 40 to 46 participants per condition. There were three targets in total, two female and one male. Gender of the target was counterbalanced across conditions,

and within the female conditions the two targets were also counterbalanced. The primary dependent variables were recall accuracy (i.e., total number and proportion of correct descriptors), and recognition accuracy (i.e., proportion of correct identification decisions and proportion of suspect identifications).

Materials

Informed consent form. Participants were given a mildly deceptive consent form explaining that they will be asked to watch two videos and answer questions about the videos (Appendix B).

Occupation stereotype profile videos. All participants were shown one of eight occupation stereotype profile videos, all approximately 60 seconds in length. The videos described gender consistent or gender inconsistent occupational stereotypes. The occupations chosen were based on the results from a material testing study asking participants to rate how feminine or masculine they viewed various occupations (please see Appendix N for a description of the material testing study and the corresponding materials and results). The videos consisted of a picture of a male or female confederate with an audio overlay describing the confederate's occupation (i.e., nurse or construction worker; Appendix C).

Stereotype rating questionnaire. Once participants viewed the stereotype profile videos, participants were asked to fill out one of two questionnaires (both containing two questions) indicating the extent to which the person (question 1) and profile (question 2) were consistent with their perception of a nurse (Appendix D1) or construction worker (Appendix D2). Both questions were rated on a 7-point scale (1 = *consistent*, 7 = *inconsistent*). The first question was used to assess the stereotype manipulation and as well measure the personal stereotypes held by the participants.

Open recall task. Participants were asked to recall any details they could remember about the confederate in the stereotype profile video and their story. This form was used as a filler task (Appendix E).

Cognitive load measures. Participants were asked to rate their mental workload using the RSME (Appendix F) and the RTLX (Appendix G). These forms were used to assess the cognitive load experienced during the occupation stereotype video and the crime video.

Computer game. As an unrelated filler task, participants were asked to play “Bejeweled,” a Tetris-like computer game, where same coloured jewels are matched in sets of three or more in order to move from one level to the next. In the familiar conditions, this was used to create a delay between meeting the confederate (familiarization) and watching them commit a crime; in the unfamiliar conditions, this was simply used as a delay in order to keep all procedures as consistent as possible.

Crime video. All participants were shown a brief video (approximately 1 minute in length) depicting a staged, non-violent theft. The video portrayed a female victim and confederate (male or female) sitting in a room setup to look like a breakroom. The video began with the confederate sitting on a couch reading a book while the victim sat on a chair using a Mac laptop (victim faced away from the camera). After a few seconds, the victim put the laptop down on a side table to answer a phone call and left the room. The lens then focused on the confederate for approximately 20 seconds while he/she looked suspiciously at the laptop left unattended on the side table. The camera then panned in for a closer look at the confederate’s face for 6 seconds. The lens panned back out while the confederate continued to look suspiciously at the laptop. The confederate then got up from the couch, picked up the laptop, and left the room.

Secondary consent form. Participants were given a secondary consent form outlining the true purpose of the study and why mild deception was necessary for its completion (Appendix H).

Demographics form. After participants viewed the video they were asked to complete a demographics form requesting their age, sex, and ethnicity (Appendix I). Participants were able to refuse to answer any of the demographic questions without penalty.

Description form (free recall task 2). After participants viewed the crime video they were asked to complete a description form, asking them to describe everything they remembered about the video (event description) and everything they remembered about the criminal in the video (person description; event and person descriptions were counterbalanced across participants). This form was used to assess the participants' recall memory of the target (Appendix J).

Puzzle task. As a second unrelated filler task, participants were asked to complete a word search puzzle. This was used to simulate a delay that occurs in real life scenarios between witnessing a crime and participating in the identification task.

Photographic lineup. For each target lineup, six 4- x 6-inch photographs were taken of volunteers resembling the three targets. Five of these photographs were used as the foil photographs in all lineups, while the sixth photograph was used as the target replacement in the target-absent lineups. For each target, a 4 x 6 inch photograph was taken for the target-present lineups. All photos were in colour, head and upper body photographs taken against a white background. Persons photographed adopted a neutral facial expression in order to maintain consistency across photographs.

Creation of the photographic lineup. For each target, six photographs resembling the target were taken of volunteers or chosen from the Laboratory for Child Forensic Psychology's online database of mugshots. This database consists of volunteers from Carleton University who have given consent to have their photos taken for the purpose of eyewitness research studies. These pictures were chosen by the primary researcher, a second graduate student, and a professor. Next, three student volunteers independently watched the three crime videos and provided a written description of each of the perpetrators' (i.e., targets') physical appearance. For each target, these descriptions were combined and a composite description was created. These descriptions were created in order to test lineup bias (i.e., whether the suspect gets identified at a proportion higher than chance) and lineup effective size (i.e., the degree to which the lineup members are realistic choices to be identified) as per the suggestion by Malpass (1981). The targets' descriptions were as follows; Male target: Caucasian, early to mid-20s, average build, average height, brown eyes, short brown hair (longer on top); Female 1 target: Caucasian, long straight brown hair, slim build, brown eyes, taller than average, early to mid-20s; Female 2 target: Caucasian, long straight brown hair, slim build, average height, dark eyes, early to mid-20s. Once each composite description was created, 26 mock-witnesses were provided with the composite description and shown the seven (six foils and target) photos simultaneously and asked to choose the perpetrator based on that description. This was done by the same mock-witnesses three times for each of the three target lineups. Results indicated fair lineups with each lineup member being chosen a proportionate amount (see Table 1).

Table 1

Proportion (Frequency) of Lineup Members Chosen from the Lineup Fairness Test for Each Targets Lineup

	Lineup Member						Target
	1	2	3	4	5	6	
Target Lineup							
Male	0.08 (2)	0.19 (5)	0.15 (4)	0.12 (3)	0.12 (3)	0.19 (5)	0.15 (4)
Female 1	0.19 (5)	0.08 (2)	0.12 (3)	0.12 (3)	0.08 (2)	0.19 (5)	0.23 (6)
Female 2	0.15 (4)	0.23 (6)	0.08 (2)	0.12 (3)	0.12 (3)	0.15 (4)	0.15 (4)

A test for lineup bias was conducted for each lineup using the guidelines provided by Malpass (2015). This purpose of this test is to determine whether the suspect stands out in the lineup and whether the foils are adequate alternatives to the suspect (Malpass, 2015). Using the spreadsheet created by Malpass, the total number of mock-witnesses was entered, as well as the number of mock-witnesses who chose the target, and the total number of persons in the lineup. The spreadsheet then tells you the proportion expected to identify the suspect by chance alone and the proportion of participants who actually chose the suspect. If the proportion of participants identifying the suspect reveals bias, the critical ratio will exceed 1.96 at the .05 alpha level and will exceed 2.58 at the .01 alpha level. The critical ratios in the current study were 0.08 (male target), 0.99 (female 1 target), and 0.08 (female 2 target) thus suggesting non-biased lineups.

Tredoux's *E* was also calculated for each lineup. Tredoux's *E* tells us the number of lineup members who are realistic choices based on the description given of the perpetrator (Tredoux, 1998). Similar to the spreadsheet used to test for lineup fairness, a spreadsheet created

by Malpass (2015) was used to calculate Tredoux's E . In order to calculate Tredoux's E , the total number of mock-witnesses was entered, as well as the total number of persons in the lineup, and the number of times each lineup member was identified. This was done for all three target lineups. For the male lineup, $E' = 6.50$, 95% CI [5.49, 7.97]; for female 1 lineup, $E' = 6.04$, 95% CI [4.81, 8.11]; and for female 2 lineup, $E' = 6.38$, 95% CI [5.20, 8.26]. This indicates that for each lineup it is estimated that at least 6 of the lineup members are realistic choices based on the description given, thus suggesting fair lineups.

Simultaneous lineup procedure. The lineup photographs were presented simultaneously to the participant in two rows. Each row contained three photographs. A "Not Here" option appeared at the bottom of the two rows, centered. The location of the target or the replacement photograph was randomly determined, with the foils displayed in the same order relative to each other. Participants were provided with the following instructions prior to viewing the lineup: "*Think back to the video. Think back to what the criminal looks like. I am going to show you some pictures. Please look at the pictures. The criminal's picture may or may not be here. If you see the criminal's picture, please place a check mark in the box that corresponds with the criminal's lineup number. If you do not see the criminal, please place a check mark in the box labeled not here.*" Participants were provided with a form to make their selection. The form included the same instructions that were given orally to the participant. The form included seven boxes, numbered 1 to 6, and one labelled "Not Here". Following their identification decision, participants were asked to rate their confidence in the accuracy of their decision on a scale of 0 (*not at all confident*) – 100 (*very confident*; Appendix K).

Study check form. Upon completion of the identification task, participants were asked a number of questions to determine whether they knew the true nature of the study before

participating and whether they knew the target or any lineup members before participating (Appendix L). Overall, 21 participants (7.3%) said they were aware of the true purpose of the study and another 4 participants (1.4%) said they knew someone from the lineup. These participants were not used for analyses. Five participants (1.7%) said they had taken a class on eyewitness memory and another 29 participants (10.1%) stated that they had previously completed an identification lineup, however; these participants did not differ significantly from the rest of the participants in terms of their identification accuracy² and as such they were included in the subsequent analyses.

Procedure

After ethical approval was received from Carleton University, a recruitment notice was posted via SONA (Appendix A), an online recruitment tool. Once registered with SONA, participants scheduled an appointment to attend the Laboratory for Child Forensic Psychology to complete the study. Participants were randomly assigned to experimental conditions prior to arrival. Upon arrival, participants were asked to complete a deceptive informed consent form (Appendix B). They were then shown a stereotype profile video of a male or female confederate's picture and an audio story representing a gender consistent stereotype (male construction worker, female nurse) or a gender inconsistent stereotype (male nurse, female construction worker; Appendix C). Following the stereotype profile video, each participant was asked to complete a stereotype ratings questionnaire (Appendix D1 or D2) and to do an open recall task (Appendix E) where they wrote down everything they remembered about the confederate in the stereotype profile video. The stereotype ratings questionnaire task and the

² Having taken a class on eyewitness memory did not significantly influence identification accuracy, ($p = .67$, Fisher exact test). Also, previously completing a lineup identification did not significantly influence identification accuracy, $\chi^2(1, N = 257) = 0.05, p = .82, \phi = 0.01$.

open recall task were counterbalanced across conditions. Participants then filled out the RSME (Appendix F) and the RTLX (Appendix G) regarding their experience of watching the stereotype video. Participants were then asked to play the computer game “Bejewelled” and were made aware that their score in the game did not matter. The time from the end of the stereotype profile video and the crime video was 15 minutes, regardless of the time it took to complete the questionnaire, open recall, and cognitive load measures.

Participants were then asked to watch a second video, a crime taking place in which the participant was either familiar or unfamiliar with the perpetrator. In the familiar condition, the perpetrator in the crime video was the same confederate as shown in the stereotype profile video. In the unfamiliar condition, the perpetrator in the crime video was different than the confederate shown in the stereotype profile video. Participants were not made aware that their memory for the video would be tested prior to watching the video. Following the crime video, the participants were given a secondary consent form (Appendix H), explaining the true purpose of the study and why mild deception was necessary for its completion. Participants then filled out a demographics form (Appendix I) and the description form (Appendix J) followed by the RSME (Appendix F) and the RTLX (Appendix G) relating to their experience of watching the crime video. Following the completion of these measures, a second delay task (word search puzzle) was given. The delay lasted 10 minutes. After the delay, participants were presented with a simultaneous lineup in which the perpetrator from the video was either present or absent. They were asked to use the simultaneous lineup response form (Appendix K) to indicate which photograph from the lineup matches the perpetrator from the video. Participants then filled out the RSME (Appendix F) and the RTLX (Appendix G) for a third time, this time relating to their experience completing the lineup task. Participants then filled out the study check form

(Appendix L). Finally, participants were asked to read over a debriefing form (Appendix M) and thanked for their time.

Results

Stereotypes Manipulation Check

In addition to placing the participants in either the consistent or inconsistent stereotype conditions, the participants also were asked to subjectively rate how stereotype consistent they viewed the target. Participants were asked to rate how well the target fit their perception of what a nurse or construction worker looks like on a scale from 1 to 7 (1 = *consistent*; 7 = *inconsistent*). However, before any analyses were run the scores from this scale were reverse scored as it was more logical to discuss this measure in this way. Thus, lower scores on the scale now represent inconsistent ratings and higher scores represent consistent ratings (i.e., 1 = *inconsistent*; 7 = *consistent*). Using the scores from this scale, I conducted an independent samples *t*-test to determine whether participants who viewed a consistent stereotype video had higher stereotype consistency ratings than participants who viewed an inconsistent video. The results demonstrated that the participants who viewed the consistent videos had significantly higher stereotype consistency ratings ($M = 5.65$, $SD = 1.33$) compared to the participants who viewed the inconsistent videos ($M = 4.98$, $SD = 1.79$), $t(167) = -2.80$, $p = .006$. Thus, participants assigned to the stereotype consistent conditions rated the target as more consistent whereas the participants assigned to the stereotype inconsistent conditions rated the target as more inconsistent.

Cognitive Load Measures

Both the RSME and RTLX were used to assess cognitive load for the participants after the occupation video and the crime video. Initial analyses indicated that the RSME and the

RTLX scores following the occupation video were highly correlated, $r = .58, p < .001$. In addition, the RSME and RTLX scores following the crime video were highly correlated, $r = .65, p < .001$. Therefore, in an effort to report on the necessary analyses only the results from RSME scales are reported in the main document.³

The RSME scores were used to determine whether the participants were under cognitive load while watching the occupation video. It was expected that participants would rate their cognitive load low following the occupational video (OV). The mean RSME OV score was 52.40 ($SD = 24.35$) which may be considered low given the scale extends from 0 to 120. However, there are no cut-off scores for determining various “levels” (i.e., low, moderate, high) of cognitive load so it cannot be said with certainty that the participants were not feeling cognitively overwhelmed.

It was expected that participants would rate their cognitive load as higher for the crime video (CV) than for the OV. In order to determine whether participant’s cognitive load was higher for the CV than the OV, a paired t -test was conducted comparing the mean RSME OV score to the mean RSME CV score. The results indicated that the mean RSME OV score ($M = 52.56$) was not significantly different from the mean RSME CV score ($M = 53.57$), $t = -0.56, p = .58$. Thus indicating that participants did not feel their cognitive load was higher for the crime video than the occupation video.

Recall Data

Coding data. In order to create the three coding manuals (i.e., one for each of the three targets) two independent raters watched the crime videos and determined the descriptors that would be deemed accurate. Eyewitness recall accuracy was assessed based on their response to

³ All analyses involving mental workload measures were run using both the RSME and RTLX. It is noteworthy that similar results emerged for the RSME and RTLX.

an open-ended statement asking them to, "Please write down everything you can remember about the criminal in the video." The total number of descriptors as well as the total number of correct descriptors was recorded for each participant. Descriptors for height, weight, and age were marked correct if they fell within a specified range (i.e., height = +/- 2 inches, weight = +/- 10 pounds, age = +/- 3 years). All 257 responses were coded by the primary researcher. In order to assess inter-rater reliability, 20 randomly selected response sets for each target (i.e., 60 in total; 23.3%) were coded a second time by a trained undergraduate psychology student. Initial discrepancies between coders were discussed and a decision was reached on how to code those descriptors. Intraclass correlation coefficients (ICC) were calculated for both the total number of descriptors and the number of correct descriptors for each target. Any remaining discrepancies between coders were discussed until the resulting ICC values (Cronbach's Alpha) fell within the acceptable range of 0.8 to 1.00 (Pozzulo & Warren, 2003). Table 2 shows the final ICC values for each targets recall data.

Table 2

Intraclass Correlation Coefficient (ICC) Representing the Interrater Reliability of the Targets' Recall Data

	ICC value (Absolute Agreement)	95% CI
Male Target		
Total Number of Descriptors	1.00	(0.99 – 1.00)
Number of Correct Descriptors	0.99	(0.98 – 1.00)
Female Target 1		
Total Number of Descriptors	0.99	(0.97 – 1.00)
Number of Correct Descriptors	0.99	(0.97 – 1.00)
Female Target 2		
Total Number of Descriptors	0.99	(0.98 – 1.00)
Number of Correct Descriptors	0.99	(0.98 – 1.00)

Note. CI = Confidence Interval

Recall accuracy. Initially, the two primary dependent variables to be used when assessing recall accuracy were the total number of descriptors and the proportion of correct descriptors. After screening the variables for normality it was discovered that that the proportion of correct descriptors was severely negatively skewed. This was likely because the majority (52%) of participants had 100% accurate descriptors. In order to try and mediate this issue a new variable was created in an effort to spread out the distribution. For this new proportion variable it was not only the proportion that would influence the variable, but as well the number of correct descriptors. This variable was calculated by dividing correct descriptors by total descriptors + 1. That way the proportion could not reach 100% and the overall number of correct descriptors also would be influential, thus creating more variance in the distribution. This new variable, which

will be referred to as *proportion correct transformed*, was still skewed (skewness = - 1.92, $SD = .19$) and leptokurtic (kurtosis = 4.76, $SD = .37$). Z -scores were computed and four outliers were identified (i.e., z -scores = - 3.59, - 3.59, - 4.88, - 4.88) and removed. Normality was assessed for the new variable with the outliers removed; proportion correct transformed had skewness and kurtosis values within the normal range (skewness = - .967, $SE = .19$; kurtosis = .17, $SE = .38$). Therefore, it was decided that the two primary dependent variables to be used for the subsequent analyses were total descriptors and proportion correct transformed.

Two one-way ANOVAs were used to assess whether familiarity and stereotype consistency influenced the total number of recall descriptors reported and the proportion of correct descriptors reported (see Table 3 for means and standard deviations). It was hypothesized that participants in the inconsistent familiar condition would provide more descriptors and that their descriptors would be more accurate than those in the consistent familiar condition, and that participants in both the familiar conditions would provide a greater number of descriptors and be more accurate than participants in the unfamiliar condition. The first one-way ANOVA examined whether familiarity consistency predicted the total number of perpetrator descriptors recalled by participants. The omnibus F -test indicated no significant differences between groups, $F(2, 254) = 0.78, p = .46, \eta^2 = .006$. The second one-way ANOVA examined whether familiarity consistency predicted the proportion of correct descriptors recalled by participants using proportion transformed. Again, the omnibus F -test indicated no significant differences between groups, $F(2, 250) = 0.20, p = .82, \eta^2 = .002$.

Table 3

Means and Standard Deviation for Total Descriptors and Proportion of Correct Descriptors (%) by Familiarity Consistency

	Total Descriptors		Proportion Correct	
	<i>M</i>	SD	<i>M</i>	SD
Familiar Consistent	7.54	3.14	77.78	13.74
Familiar Inconsistent	6.97	3.09	76.54	13.34
Unfamiliar	7.14	2.91	76.85	12.55

Stereotype ratings questionnaire. In addition to examining whether the experimental stereotype groupings predicted recall accuracy, the subjective stereotype consistency ratings were examined to determine if they would predict recall accuracy. Two separate linear regression analyses were run to determine whether stereotype consistency ratings predicted the total descriptors as well as the proportion of correct descriptors. First a linear regression was run on the participants in the familiar conditions with complete data on the stereotypes ratings ($n = 169$) to determine whether stereotype consistency ratings predicted total descriptors. The stereotypes rating was a significant predictor of total number of descriptors, $R^2 = .04$, $F(1, 167) = 6.17$, $p = .01$. That is, the more consistent the participant rated the target, the more descriptors they provided. Next, a linear regression was run on the participants in the familiar conditions with complete data on the stereotypes ratings ($n = 165$) to determine whether stereotype consistency ratings predicted the proportion of correct descriptors using proportion transformed. Stereotypes ratings was marginally significant in predicting proportion transformed, $R^2 = .02$, $F(1, 163) = 3.59$, $p = .06$. Therefore, the more consistent the participant rated the target, there seems to be a trend towards providing more correct descriptors.

Cognitive load and recall memory. In addition to the previous analyses involving cognitive load, it also was of interest to determine whether the cognitive load scores could be used to predict the amount of information recalled by participants in the stereotype consistent and inconsistent groups. If cognitive load affects the number of descriptors remembered then participants' cognitive load should lead to more descriptors/correct descriptors in the inconsistent condition and fewer descriptors/correct descriptors in the consistent condition. In order to determine this, a hierarchical linear regression was run on the participants in the familiar conditions ($n = 169$) and examined whether RSME OV, stereotype groups, and the product term representing the two-way interaction predicted total number of descriptors. Block 1 contained the main effects while Block 2 contained the main effects and interaction term. The first model was not significant, $F(2, 166) = 2.07, p = .13$; additionally, the second model also was not significant, $F(3, 165) = 1.37, p = .25$.

Next, a hierarchical linear regression was run to determine if RSME OV, stereotype groups, and the product term representing the two-way interaction predicted proportion transformed. Again, the first model was not significant, $F(2, 162) = 0.50, p = .61$; and the second model also was not significant, $F(3, 161) = 0.46, p = .71$. Therefore, cognitive load did not influence the number of descriptors or proportion of correct descriptors reported in either the stereotype consistent condition or the stereotype inconsistent condition.

Before examining the identification data, a logistic regression was run to determine whether recall accuracy predicted recognition accuracy (correct identifications and correct rejections [Correct] were coded as 1; and false identifications, incorrect rejections, and foil identifications [Incorrect] were coded as 0). For recall accuracy both the total number of descriptors and the proportion of correct descriptors were used as the predictor variables. There

was a significant main effect of total descriptors; for each additional descriptor reported, the odds of making a correct identification decision increased 1.12 times, $b = 0.11$, $SE = 0.05$, Wald's $\chi^2(1) = 5.08$, $p = .02$, $e^b = 1.12$ (95% CI [1.02, 1.24]), however there was not a significant main effect of proportion of correct descriptors, $b = -1.40$, $SE = 1.14$, Wald's $\chi^2(1) = 1.50$, $p = .22$, $e^b = 0.25$ (95% CI [0.03, 2.31]).

Identification Data

Identification accuracy. It was hypothesized that eyewitnesses would be more likely to make a correct decision (i.e., correct identification or correct rejection) when they were familiar with the target compared to when they were unfamiliar with the target. In addition, it also was hypothesized that when familiar with the target, eyewitnesses would be more likely to provide a correct decision when the target was stereotype inconsistent compared to when the target was stereotype consistent. Thus, a hierarchical binary logistic regression was conducted to examine whether target presence, familiarity consistency (unfamiliar was used as the reference group), and the product term representing the two-way interaction predicted participants' identification decisions (Correct = 1, Incorrect = 0). Block 1 contained the main effects, and Block 2 contained the main effects and interaction term. There was no main effect of target presence, $b = -0.02$, $SE = 0.25$, Wald's $\chi^2(1) = 0.01$, $p = .94$, $e^b = 0.98$ (95% CI [0.60, 1.61]) and no main effect of familiarity consistency, Wald's $\chi^2(2) = 0.24$, $p = .89$. In addition, the interaction between target presence and familiarity consistency was not significant, Wald's $\chi^2(2) = 1.97$, $p = .37$. Therefore, these results demonstrate that familiarity and stereotype consistency did not influence eyewitnesses' correct identification decisions (see Table 4 for the percentage of correct decisions for each experimental group).

Table 4

Percentage of Correct Identification Decisions based on Familiarity Consistency and Target Presence

	Familiarity Consistency		
	Familiar Consistent	Familiar Inconsistent	Unfamiliar
Target Presence			
Target Absent	47.5	61.9	52.2
Target Present	56.8	50.0	53.7

In addition, it was hypothesized that familiarity between an eyewitness and perpetrator would produce better applied outcomes (i.e., increase perpetrator identifications in TP lineups and decrease innocent suspect identifications in TA lineups) than no prior familiarity. Further, it was hypothesized that when familiar with the target, a stereotype inconsistent target would produce better applied outcomes than a stereotype consistent target. Thus, a second hierarchical binary logistic regression was conducted to examine whether target presence, familiarity consistency (unfamiliar was used as the reference group), and the product term representing the two-way interaction predicted participants suspect identifications (Suspect ID = 1, Foil ID/Rejections = 0). Block 1 contained the main effects, and Block 2 contained the main effects and interaction term. There was a significant main effect of target presence, where participants in the target present condition were 10.19 times more likely to make a suspect identification than those in the target absent condition, $b = 2.32$, $SE = 0.34$, Wald's $\chi^2(1) = 46.04$, $p < .001$, $e^b = 10.19$ (95% CI [5.21, 19.93]). However, there was no main effect of familiarity consistency, Wald's $\chi^2(2) = 0.07$, $p = .97$. In addition, the interaction between target presence and familiarity consistency was not significant, Wald's $\chi^2(2) = 0.81$, $p = .67$. These results demonstrate that, as would be expected, the guilty suspect is more likely to be identified than the designated innocent

suspect. In addition, the lack of interactions demonstrates that neither familiarity nor stereotype consistency influenced applied outcomes in the manner hypothesized.

Stereotype ratings questionnaire. A regression was conducted to examine whether higher stereotype consistency ratings in the familiar conditions could predict accuracy in the target-present and target-absent lineups. Therefore, a hierarchical binary logistic regression was run on the participants in the familiar conditions ($n = 169$) to determine whether stereotype consistency ratings, target presence, and the product term representing the two-way interaction predicted participants correct identification decisions (Correct = 1; Incorrect = 0). Block 1 contained the main effects, and Block 2 contained the main effects and interaction term. There was a significant main effect of stereotype ratings, for each one unit decrease in stereotype consistency, the odds of making a correct identification decision increased 1.30 times, $b = 0.26$, $SE = 0.10$, Wald's $\chi^2(1) = 6.46$, $p = .01$, $e^b = 1.30$ (95% CI [1.06, 1.59]). However, there was no main effect of target presence, $b = 0.13$, $SE = 0.32$, Wald's $\chi^2(1) = 0.17$, $p = .68$, $e^b = 1.14$ (95% CI [0.61, 2.12]). In addition, the interaction between target presence and stereotypes rating was not significant, $b = -0.07$, $SE = 0.21$, Wald's $\chi^2(1) = 0.10$, $p = .76$, $e^b = 0.94$ (95% CI [0.62, 1.42]). This demonstrates that the participants' subjective stereotypes influenced their recognition accuracy. For both the target-present and target-absent conditions, the more inconsistent the participant viewed the target; the more likely they were to be correct in their identification decision. Even after controlling for which stereotypes' condition the participants were in, the subjective stereotypes ratings still had an effect, $b = 0.26$, $SE = 0.11$, Wald's $\chi^2(1) = 6.22$, $p = .01$, $e^b = 1.30$ (95% CI [1.06, 1.60]).

Cognitive load and identification accuracy. In addition to the previous analyses involving cognitive load, the cognitive load scores also could be used to determine whether

higher scores on the RSME OV leads to more correct identification decisions in the consistent condition and fewer correct identification decisions in the inconsistent condition. In order to examine this, a hierarchical logistic regression was run on the participants in the familiar conditions ($n = 169$) and examined whether RSME OV, stereotype groups, and the product term representing the two-way interaction predicted correct identification decisions (Correct = 1; Incorrect = 0). Block 1 contained the main effects while Block 2 contained the main effects and interaction term. There was no main effect of RSME OV, $b = -0.02$, $SE = 0.22$, Wald's $\chi^2(1) = 0.90$, $p = .34$, $e^b = 0.98$ (95% CI [0.94, 1.02]) and no main effect of stereotypes consistency, $b = -0.22$, $SE = 0.80$, Wald's $\chi^2(1) = 0.08$, $p = .78$, $e^b = 0.80$ (95% CI [0.17, 3.87]). In addition, the interaction between RSME OV and stereotypes consistency was not significant, $b = 0.01$, $SE = 0.01$, Wald's $\chi^2(1) = 0.22$, $p = .64$, $e^b = 1.01$ (95% CI [0.98, 1.03]).

Discussion

The purpose of the present study was to examine how brief familiarity in combination with target-presence and stereotypes would influence eyewitness recall and recognition accuracy. There is a wealth of research examining the recall and identification accuracy of eyewitnesses who have not previously seen the perpetrator before the crime (i.e., stranger eyewitnesses). This research has been integral for informing policy and best practice for dealing with eyewitnesses in the criminal justice system. However, research looking at the recall and identification abilities of an eyewitness who has previously had contact with the perpetrator has been neglected. It is important to begin examining the recall and recognition abilities of familiar eyewitnesses as research suggests that in many cases an eyewitness will state that they have previously encountered the suspect before the crime (e.g., Flowe et al., 2011; Memon et al., 2011; Valentine et al., 2003). For example, in a report examining the rate of criminal victimization, Statistics

Canada found that in the majority of cases (52%) the victim knew their attacker (Perrault, 2015); while Flowe and colleagues (2011) found that in 67% of the sampled cases at least one eyewitness was previously familiar with the suspect.

Familiarity has been broadly defined as *any* prior exposure to the perpetrator before the commission of a crime (Pozzulo, 2016). Familiarity can range from someone met once, to someone considered a close friend or relative. It is likely that varying levels of familiarity would differentially affect the recall and identification abilities of an eyewitness, therefore in order to provide research that would be forensically useful in the criminal justice system one type of relationship within the broad definition of familiarity was examined; familiar-strangers.

Not only is it important to examine familiar witnesses because of their prevalence, but it also is important as research has demonstrated that cases involving familiar witnesses are more likely to be prosecuted than stranger cases (Flowe et al., 2011). It seems that investigators and prosecutors may be more confident in the identification accuracy of someone who is familiar with the perpetrator. Given that there is limited research on the identification accuracy of familiar eyewitnesses, it is unclear whether this assumption is valid. In addition, the current policies for collecting and preserving eyewitness accounts are based on stranger eyewitness research; these policies may not be suitable for collecting the most accurate eyewitness information when the eyewitness is familiar with the perpetrator. Therefore, the results of the current study provide a necessary start to the understanding of familiar-stranger eyewitness' recall and identification abilities.

Further, in order to understand eyewitness recall and recognition of familiar-strangers it is important to examine factors that may influence this relationship. In the current study, a familiar-stranger eyewitness was defined as someone having had one brief exposure to the

perpetrator before the commission of the crime. With this level of familiarity examined, one factor that was considered to affect this relationship was stereotypes as stereotypes tend to be activated following a brief meeting with someone (Brewer, 1988; Fiske & Neuberg, 1990).

During a brief social exchange, information acquired can either be consistent with a pre-existing stereotype or inconsistent with a pre-existing stereotype. Therefore, the purpose of examining stereotypes in the current study was to determine whether a stereotype consistent familiar-stranger or a stereotype inconsistent familiar-stranger would have an advantage in eyewitness memory. If memory for a familiar perpetrator differs based on stereotypes, then future research examining familiar-strangers needs to account for the effects of this variable.

Familiarity

The main purpose of this study was to examine how familiarity between an eyewitness and a perpetrator influences eyewitness recall and identification accuracy. Because familiarity has been described quite broadly in the past (i.e., *any* prior exposure to a target before the commission of a crime), it is necessary to examine various levels of familiarity. In the current study, familiarity was manipulated to try and mimic a relationship somewhere between a stranger and acquaintance. In this sense, familiarity was at the less familiar end of the familiarity spectrum. Based on the results of the current study, it seems that low levels of prior familiarity between an eyewitness and perpetrator does not provide any enhancement in recall or identification abilities compared to a stranger eyewitness.

Recall of familiar-strangers. Contrary to what was expected, stranger eyewitnesses and familiar eyewitnesses provided a similar number of descriptors and a similar proportion of accurate descriptors at recall. In other words, participants who had a 1 minute “meeting” with the target 15 minutes before the crime took place, were no more likely to provide more descriptors

or more accurate descriptors about the target than were participants who were strangers with the target. The results of this study suggest that the memory of eyewitnesses who have had a brief exposure with the culprit may be no more informative than the memory of stranger eyewitnesses. The results of the current study are in contrast to what was predicted based on research examining event repetition and exposure duration (e.g., McNichol et al., 1999; Theunissen et al., 2017). Research on event repetition has demonstrated that recall of event details increases after being exposed to the same event several times in comparison to one exposure (McNichol et al., 1999). Based on this previous research, it was predicted that participants in the current study would be able to provide more descriptors and more correct descriptors about the perpetrator when they had a previous encounter with the target person. However, this was not found as there were no differences in recall accuracy between groups. It is possible that because details changed from the initial exposure (i.e., occupation video) to the second exposure (i.e., crime video; e.g., clothing, view of target, setting) that this impacted the participants' ability to retrieve the information they acquired from the initial meeting. Research on event repetition suggests that if important details of the event do not remain the same from one exposure to the next, recall accuracy may be lower than if the event was experienced only once (McNichol et al., 1999; Theunissen et al., 2017). It is difficult to determine whether this effect of inconsistencies in event repetition would occur with *person* repetition (i.e., repeated exposure to the same person). In Theunissen and colleagues' (2017) examination of event repetition, the information that was the same from one event to the next was the general theme of the event (i.e., all events were traffic accidents) and it was the other details (e.g., car colour, time of day) that were assessed during the memory test. With the person repetition in the current study, the information that stayed the same from one event to the next was also the main focus of the event and the main focus of the

memory test (i.e., the target). Therefore, it is possible that having the main focus of the event (i.e., the target) remain consistent from one event to the next would impact recall accuracy differently than when it is just the overall theme of the events that remains consistent. Currently there is too little research in this area to be able to determine whether inconsistencies between events would have the same effect on person recall as it does on event recall.

In terms of fuzzy trace theory (FTT), the comparable recall accuracy between witnesses in the familiar and unfamiliar conditions seems to suggest that both types of witnesses were relying on the same type of memory traces during recall; likely gist traces. Based on FTT it was predicted that familiar witnesses would rely on verbatim traces (detailed memories) while stranger witnesses would rely on gist traces (vague memories) of the criminal during recall; and thus familiar witnesses would have more accurate recall. This prediction was based on the impact of disintegration and event repetition on memory traces. FTT states that verbatim information is more susceptible to disintegration than gist information. Therefore, it would be expected that after a delay (as in the current study: 10 minutes) that the verbatim memories of the target would not be as accessible as the gist memories of the target for *both* the familiar and unfamiliar participants. However, FTT states that verbatim memories are strengthened when events are repeated and the verbatim information of the subsequent events matches with the prior events (Reyna & Lloyd, 1997). Therefore, because the target for the familiar eyewitnesses was the same in both videos it was expected that the witnesses' memory for the target would be strengthened and therefore these verbatim memories would be more accessible. Additionally, given there was no second event for the unfamiliar eyewitnesses, there was no opportunity for the verbatim memory traces to be strengthened for these witnesses, and therefore the verbatim memories would continue to disintegrate and they would rely on gist traces during recall. As

verbatim traces are more detailed memories of an event, it was predicted that familiar witnesses would have more accurate recall of the perpetrator. Given that the familiar witnesses were not any more accurate with their recall of the perpetrator than the stranger witnesses it seems a different explanation is necessary.

Examining FTT for possible explanations of these findings suggests it could be due to interference or contradiction within the familiar conditions. Interference can be anything presented that is not an exact representation of a previous memory or anything that can affect a previous memory (Brainerd & Reyna, 1993). It is possible that because the clothing of the target changed after the first encounter that this change caused interference in the familiar eyewitnesses' ability to retrieve the verbatim information about the target. Interference has been shown to be a factor that accounts for decreased accessibility of verbatim traces (e.g., Payne, Elie, Blackwell, & Neuschatz, 1996) but does not affect the accessibility of gist information (Reyna & Lloyd, 1997). In addition, contradiction is a source of interference that hinders the accessibility of verbatim memories whereas contradiction cues gist memories (Reyna & Lloyd, 1997). Therefore, seeing the familiar target committing a crime was likely a contradiction to the prior meeting with them as a nurse or construction worker, and thus gist memories were cued and retrieved during recall. However, the majority of research on interference and FTT in general tends to focus on memory of learned words; therefore until more research is conducted examining these ideas in conjunction with memory of learned people, this discussion is just conjecture.

Overall, the results of the current study seem to suggest that eyewitnesses who are briefly familiar with a perpetrator are no more accurate at providing criminal descriptors than someone who has never seen the perpetrator before the crime. Although these results provide a strong

basis for future researchers to build on, there is currently too little research on recall memory of familiar people to draw any firm conclusions at this time.

It is also important to consider a few caveats of the recall results. As previously mentioned, the clothing of the target did change from the occupation video to the crime video; therefore, we would not expect to see an increase in familiar eyewitnesses' ability to recall clothing descriptors. This is important to note as clothing has been reliably found to be one of the most commonly reported descriptors (Pozzulo & Warren, 2003; Sheahan, Pica, Pozzulo, & Nastasa, 2017). The lack of difference in number and correct descriptors in the current study could be due to the large amount of clothing descriptors reported. Future research should consider examining the accuracy of clothing descriptors separately to see if the exclusion of these descriptors would produce more significant results.

Identifying familiar-strangers. It was hypothesized that the eyewitnesses who were familiar with the perpetrator before the crime would have better identification accuracy than the eyewitnesses who had never seen the perpetrator before the crime. The results of the study did not support this hypothesis as the witnesses in the familiar condition and the stranger condition had similar correct identification and correct rejection rates. Based on the results of the study it appears that brief familiarity as manipulated in the current study does not enhance eyewitness identification. Participants who viewed a 1 minute video of the perpetrator 15 minutes before the crime were no more likely to make a correct identification decision (correct identification or correct rejection) than were participants who had never seen the perpetrator before.

This finding was surprising given research on facial recognition has demonstrated that familiar faces are better recognized than never before seen faces (e.g., Burton et al., 1999). More specifically, previous research from Bruce and colleagues (2001) has demonstrated that

familiarity alone (i.e., 30 second exposure to a target) may not increase facial recognition; however, when socially exposed to a target, recognition becomes significantly better compared to no prior exposure. Following the results of Bruce et al., the current study aimed not only to expose the participant to the target but also to create a social exposure through the target sharing personal information about themselves (i.e., talked for 1 minute about their occupation as either a nurse or construction worker). However, in contrast with previous research, the participants who had a brief social exposure to the target did not perform any better on the identification task than did the participants who had never seen the target before. This was surprising given the results of Bruce et al. and various other studies that have found that providing social information about a target leads to improved recognition of the target (e.g., Curran & Hancock, 2007; Yovel & Paller, 2004).

The difference in findings between the current study and past research could be a result of the difference in the type and amount of social exposure to the target. Specifically, the social exposure as manipulated in the current study may not have been strong enough to enhance memory. It is possible that having a target talk at you about themselves as opposed to having a conversation (i.e., back and forth with personal information) or having a conversation about the target with another person (as in Bruce et al., 2001) would not create the same feeling of familiarity. Similarly, in the current study the target was introduced in a still video (i.e., a still picture with an audio overlay) and thus the participants were not able to see the target's face in motion. Research has demonstrated that recognition for faces is enhanced when a face is shown in motion (e.g., Knight & Johnston, 1997; O'Toole, Roark, & Abdi, 2002). When presented with a static face only a limited amount of information can be extracted; facial motion adds to the structural information of a face which leads to optimal facial processing (Xiao, Perrotta, Quinn,

Wang, Sun, & Lee, 2014). Therefore, the participants in the current study were presented with limited facial information which likely hindered participants' ability to encode enough information to be able to recognize the target later. Future research should manipulate familiarity to include a live video of the target with a deeper social exchange component; this will not only better mimic a real-world meeting but should also enhance recognition in a realistic manner.

Another possible explanation for the discrepancy in findings could be as a result of the inherent differences in the nature of a facial recognition task compared to an eyewitness identification task. The facial recognition paradigm involves only exposure to the face (and possibly upper shoulders) whereas in an eyewitness paradigm, the witness is exposed to the entire person. In addition, in contrast to an eyewitness paradigm where a participant is exposed to one target person, facial recognition studies usually involve exposure to several target faces. Further, the consequences involved in an eyewitness task are more mentally taxing than the consequences involved in a facial recognition task. Participants would likely feel more pressure to make a correct decision in a situation where they are imparting the role of an eyewitness as the consequences associated with an incorrect decision are serious. They would likely not feel the same pressure in a task where they simply need to decide whether a face has been seen previously. Thus, it may not be surprising that the results regarding familiar facial recognition and the results regarding familiar eyewitness identification do not mimic one another.

A caveat to the familiarity findings in the current study is that the subjective familiarity felt by the participants was not assessed; therefore, it is not clear whether the familiarity manipulation was effective and participants in the familiar conditions actually felt familiar with the perpetrator. In other words, it cannot be determined whether the objective familiarity as manipulated by the researchers was effective in creating a feeling of subjective familiarity within

the participants. These results therefore should not be applied to subjective feelings of familiarity. Future research should measure the witnesses' perceived familiarity with the perpetrator as this kind of familiarity is likely to be how an eyewitness expresses their previous relationship with the perpetrator in the real world.

Examining the findings in terms of dual-process theories of recognition memory and the SBME, it seems that there was no SBME between conditions as the participants in both the familiar and unfamiliar conditions had similar correct identification and false identification rates. It was predicted that there would be a SBME where the participants in the familiar condition would have more hits (i.e., correct identifications) and fewer false alarms (i.e., false identifications) than the participants in the unfamiliar condition. This prediction was based on the idea that repetition of an item or event enhances the strength of the memory trace making it easier for a person to recollect the item or event (Cary & Reder, 2003). In the current study, because participants in the familiar condition had repeated exposure to the target, it was assumed that their memory trace of the target would be strong and therefore they would be able to use recollection-based recognition during the identification task. In this sense they would have more recollection based hits (i.e., correct identifications) than the participants in the unfamiliar condition who would have to rely on a familiarity judgment. However, it seems the participants in the familiar condition were not able to rely on recollection and instead, like the participants in the unfamiliar condition, had to rely on a familiarity judgment. Based on this conclusion, it is likely that the 1 minute video of the target prior to the crime video was not enough exposure to create a memory trace strong enough for the participants to rely on recollection during the identification task.

Overall, the results from the study suggest that brief prior familiarity between an eyewitness and a perpetrator does not impact identification accuracy. This is an important finding to consider as research has demonstrated that cases involving familiar eyewitnesses are more likely to be prosecuted than cases involving stranger eyewitnesses (Flowe et al., 2011). Thus, it appears as though investigators and prosecutors have more confidence in a positive identification from a familiar eyewitness than a stranger eyewitness. In addition, there has been some recent research on juror decision-making and perceptions of familiar eyewitnesses that suggest jurors also may feel more confident in the accuracy of a familiar eyewitness. For example, Sheahan, Pozzulo, Reed, and Pica (2017) found that mock jurors were more likely to render a guilty verdict and assign higher guilt ratings (on a scale from 0 = *Not Guilty* to 100 = *Guilty*) when the eyewitness and defendant were described as being familiar with one another compared to when they were described as strangers. Based on the null results of the current study, this confidence in familiar witnesses from prosecutors and jurors is concerning. As the results of the current study suggest that familiar eyewitnesses have a similar proportion of false identifications; this unfounded confidence could lead to serious problems with false identifications leading to the conviction of innocent persons. Although there are only a few studies that have looked at brief familiarity and its impact on identification accuracy, results from the current study suggests that members of the criminal justice system should use caution when attributing more weight to a familiar witness identification before more research has been conducted.

Stereotypes

In addition to looking at the memory abilities of familiar-stranger eyewitnesses, it was also of interest to begin to examine factors that may influence memory of familiar-strangers.

Therefore, a main purpose of the current study was to examine whether consistent versus inconsistent gender stereotypes would influence memory of familiar-strangers. Specifically, the current study aimed to examine whether the occupation of a target (gender consistent or gender inconsistent) would influence a familiar eyewitness' recall and recognition memory. Based on the results it seems that dichotomous stereotyping does not influence recall and recognition accuracy, however, subjective stereotyping may influence the amount of information recalled by an eyewitness and the likelihood of making a correct identification.

Prior to examining the participant's memory of the perpetrator, the subjective stereotype consistency ratings were examined to determine whether the stereotype manipulation was effective. Results from this analysis identified that participants did discriminate between consistent and inconsistent targets. The male construction worker and female nurse (i.e., consistent targets) were rated as more stereotype consistent than the male nurse and female construction worker (i.e., inconsistent targets).

Stereotypes and recall accuracy. It was predicted that participants in the inconsistent stereotype condition would be able to remember more details about the perpetrator than the participants in the consistent stereotype condition. The results demonstrated that there was no difference between the consistent and inconsistent groups in the number of descriptors reported or the proportion of correct descriptors reported. However, it was found that the more consistent the participants rated the target (irrelevant of the target they saw) the more descriptors they reported, although, these descriptors were not necessarily more accurate. In other words, the overall number of descriptors provided was larger whether those descriptors were correct or incorrect. It is likely that there was only a difference found for the subjective ratings and not the dichotomous grouping variable because the subjective ratings were more sensitive compared to

the discrete grouping variable. Stereotypes are more complicated than just consistent or inconsistent and there are individual differences that exist for how firm or fluid someone's stereotype beliefs are (e.g., Moskowitz, 1993). For example, someone with male relatives who are nurses would likely see a male nurse as fairly consistent with their schema of a nurse, whereas someone who has only ever seen female nurses would likely see a male nurse as contradictory and inconsistent to their schema of a nurse.

Based on past research (e.g., Shapiro, 2009; Sherman & Frost, 2000), it was predicted that eyewitnesses would recall more correct details the more stereotype inconsistent the target was. In the current study, this was not found as the proportion of correct descriptors was not affected by the participant's stereotype grouping and the subjective ratings actually demonstrated a trend in the opposite direction where the more proportionally correct descriptors were provided by participants who viewed the target as more stereotype consistent.

Cognitive load and recall accuracy. Previous research on memory for stereotype consistent and inconsistent information has not all agreed. Research has been mixed as whether memory is better for consistent or inconsistent information. Researchers trying to untangle these differing results have suggested the difference in findings may be due to cognitive load (e.g., Dijksterhuis & van Knippenberg, 1995; Macrae et al., 1993; Sherman & Hamilton, 1994). These researchers believe that memory will either be better for inconsistent or consistent information depending on the degree of cognitive load induced from a task. Research has supported this suggestion by demonstrating that when under a high cognitive load, participants remember more inconsistent material and when under a low cognitive load, participants remember more consistent material (Dijksterhuis & van Knippenberg, 1995; Macrae et al., 1993; Sherman & Hamilton, 1994; Sherman & Frost, 2000). In order to try and assess whether cognitive load

influenced memory for consistent and inconsistent information in the current study, participants' cognitive load was measured via mental workload assessments. It was predicted that participants would rate their cognitive load low for the occupation video and thus participants who viewed the inconsistent target would be able to recall more details about the target than those who viewed the consistent target. Although the participants' cognitive load ratings after the occupation video did seem to be on the lower ends of the scales, there are no cut-off scores for these mental workload assessments and therefore it cannot be said with certainty that participants were not feeling cognitively overwhelmed during the video. Therefore, in addition to simply looking at the mean score of the workload assessments, regression analyses were also run to see whether scores on the workload assessments predicted the number of descriptors reported and the proportion of correct descriptors reported. Results from these analyses suggested that cognitive load did not influence the number or proportion of correct descriptors. These results suggest that it may not be cognitive load that influenced eyewitness recall memory for consistent information in the current study. However, a caveat to note about the cognitive load assessments used in the current study is that they are designed to assess mental workload after performing in a task or solving a problem within a task. Because there was no obvious "task" involved in watching the occupation or crime video, it seemed the participants were not entirely sure how to fill out the mental workload assessments following these videos. Based on this, it cannot be concluded that cognitive load had no influence on the results in the current study. Instead, it could be that the measures used in the current study were unable to accurately assess the participants' mental workload following these visual and audio activities. Nevertheless, because it is unclear whether cognitive load played a role in the current study; it is also important to provide an explanation of the results in terms of research unrelated to cognitive load.

Specifically, the results may also be explained by prior research examining memory accuracy and stereotype encoding.

Research on encoding abilities and memory for stereotyped information has suggested that people encode specific details about inconsistent events and encode only the general theme of consistent events (e.g., Bruner, 1957; Koriat et al., 2000; Sherman et al., 1998). For consistent events the specific details are filled in later by extracting from existing knowledge structures. Therefore, it would follow that in the current study the participants who viewed the target as more stereotype consistent encoded only the general theme and thus used their schemas of a typical nurse or construction worker to fill in more details about the target during recall. On the other hand, those who viewed the target as more inconsistent likely encoded the specific details and could not draw details from their schema of nurse/construction worker because the target did not fit their schema. Thus, during recall, they provided only the specific details they previously encoded. As recognition memory may not be affected by participants trying to fill in details from their schemas, it is not surprising that the identification results do not reflect this same pattern as I will discuss below.

Stereotypes and identification accuracy. As the majority of research on stereotype consistent and inconsistent information is based on recall, the recognition hypotheses of the current study are extended from the recall research. It was predicted that participants who learned about a stereotype inconsistent target would be more accurate in their identification decisions than would the participants who learned about a stereotype consistent target. Based on the stereotype condition alone, this was not found. There was no difference in identification decisions between the stereotype consistent or inconsistent participants. However, when asked to rate the stereotype consistency of the target (i.e., using the continuous stereotype ratings) it was

found that the more consistent the target was rated, the less accurate participants were in their identification decisions. In other words, the more the participants felt the target did not fit their stereotype as either a nurse or construction worker, the better they were at picking the target out of a lineup or rejecting the lineup if they were not there.

These results are consistent with prior research that has demonstrated that there is an advantage in memory for stereotype inconsistent information (Bargh & Thein, 1985; Driscoll et al., 1991; Hastie & Kumar, 1979; Macrae et al., 1993; Sherman et al., 2005; Srull, 1981; Srull et al., 1985). Research indicates that people typically spend more time processing stereotype inconsistent information over consistent information (e.g., Stangor & McMillan, 1992; Stern et al., 1984) and that stereotype inconsistent information is processed more elaborately. Further, whereas people encode specific details about an inconsistent event, people are likely to encode the general theme of consistent events and disregard the specific details (e.g., Johnston & Hawley, 1994). Therefore, the participants who viewed the target as inconsistent likely spent more time processing the target and encoding the target's specific details. On the other hand, the participants who viewed the target as consistent likely just encoded general details about the target. As a result, during the identification task, the more extensive encoding of specific details likely led to the increase in identification accuracy.

Cognitive load and identification accuracy. The finding of the stereotype inconsistent target having an advantage on memory was predicted to be due to cognitive load. Based on prior research on cognitive load, it was predicted that participants in the inconsistent condition would have better memory for the target (and therefore better identification accuracy) than participants in the consistent condition. This was hypothesized to be due to the amount of cognitive load they were experiencing during the study/learning phase of the experiment. As mentioned previously,

although the cognitive load ratings did seem to be low after watching the stereotype video, there are no cut-off scores for these mental workload assessments. Thus, as with the recall data, a regression was run to determine whether scores on the workload assessments predicted identification accuracy. The results of this analysis suggested that cognitive demand did not influence identification accuracy. However, as previously mentioned, the mental workload assessments used in the current study may not have been able to accurately measure the workload associated with the tasks at hand.

Overall, it appears that subjective stereotyping influences the identification accuracy of a familiar-stranger eyewitness. However, whether this finding can be attributed to cognitive load is not clear at this time. It could be assumed that watching a short video of someone discussing their occupation would not be overwhelming to working memory, however, due to individual differences it could be more demanding for some individuals than others and thus generalizations should not be made.

To summarize the results on the influence of occupational gender role stereotypes on eyewitness memory, it seems that it is only subjective stereotype ratings that influence recall and recognition. Neither recall accuracy nor recognition accuracy differed based on the stereotype condition the participants were placed in. However, for recall memory, the more stereotype consistent the participants rated the target, the more overall descriptors they provided. For recognition memory, the more inconsistent the participants rated the target, the more likely they were to make a correct identification decision. In other words, participants were able to provide more descriptors when they saw the target as consistent, but were better able to identify the target (or reject the lineup) when they saw the target as inconsistent. Although this may not seem to make sense on the surface, based on the research looking at recall of inconsistent and

consistent information, it is not unforeseen. Research suggests that people encode the general theme of a consistent event while with a less familiar or inconsistent event, people tend to encode the specific details of the event (e.g., Bruner, 1957; Koriat et al., 2000; Sherman et al., 1998). Based on this, it could be assumed that participants who saw the target as consistent likely encoded only the general theme of the video and when later asked to recall the details they recalled what they could remember but also included details that may not have been present in the video but are present in their mental schema of a typical nurse/construction worker. This led to these participants providing more details, although not necessarily more proportionately correct details. On the other hand, those who viewed the target as inconsistent encoded the specific details and when asked to describe the perpetrator recalled those specific details but could not elaborate by pulling information from their schemas of a nurse/construction worker because the target did not fit that schema. Therefore, they may have provided fewer details. Further, during the recognition task, encoding specific details of the target over the general theme likely helped with making a correct identification decision.

Overall, the results examining how stereotyping may impact familiar-strangers recall and recognition accuracy suggests that these relationships are complex and need more research. Future research examining familiar-stranger eyewitnesses should continue to examine factors that may influence the accuracy of these types of eyewitness. It is important to explore possible influential factors in order to obtain a more thorough understanding of how these witnesses may perform in the criminal justice system.

Limitations

It is important to note some limitations of the current study. First, as previously mentioned, participants' subjective familiarity with the target was not measured. Therefore, it is

unclear whether the familiarity manipulation was effective and participants in the familiar conditions actually felt familiar with the perpetrator. It is important for future research on familiar-stranger eyewitnesses to try and assess participant's perceived level of familiarity. However, because familiarity is not a well-defined concept, there currently is no agreed upon way to measure familiarity. It would be beneficial for researchers to try and develop a universal way of assessing familiarity in order to create a standard when examining familiar-stranger eyewitnesses.

In addition, there was only a brief delay (i.e., 10 minutes) between the crime video and the lineup task. In the real world, it may be several weeks or several months before an eyewitness is asked to make an identification. Research from both laboratory and field studies has demonstrated that the length of delay between witnessing a crime and making an identification may influence identification accuracy (e.g., Deffenbacher, Bornstein, McGorty, & Penrod, 2008; Horry, Memon, Wright, & Milne, 2012). Therefore, future research should look at longer delay times in order to better mimic a real-world eyewitness investigation.

There are also other limitations that are common among most eyewitness studies, specifically issues concerning low external validity. In the current study the conditions in which the participants watched the crime video were without threat. The participants were told they would be watching a video and were in a private room where there was little to distract them. This is not likely to occur in the real world where the eyewitness would likely be surrounded by many distractions as well as feel some stress or anxiety from the undesirable situation. In addition, participants were aware their decision in the lineup task would have no real-world consequences, therefore they may have made a less cautious decision.

Implications and Future Research

There are important real-world implications from the results of this study that need to be discussed. Preliminary results looking at brief familiarity between an eyewitness and a perpetrator suggests that these types of eyewitnesses are no more accurate at making a correct identification or correct rejection than stranger eyewitnesses. Therefore, investigators and juries should not be more confident in the identification accuracy of an eyewitness who claims to have briefly met the perpetrator before. Instead, someone who is briefly familiar with the perpetrator should be assumed to have the same recall and identification accuracy as someone who has never seen the perpetrator prior to the crime. Given that familiarity exists on a continuum research needs to focus on various levels of familiarity to understand how different levels may impact eyewitness performance. As this study attempted to look at the less familiar end of the continuum, future research should try and replicate these findings and further the research by examining “stronger” manipulations of familiarity. In addition, because familiarity is highly subjective, future research should assess perceived familiarity following a crime and examine how perceived familiarity from an eyewitness may impact eyewitness performance as this may be more diagnostic in practice.

Further, this study also provides insight into how stereotyping may influence eyewitness’ memory of familiar-strangers. It seems with familiar-strangers that it is only subjective stereotyping that influences an eyewitness’ recall and identification. However, because it is not yet clear how brief familiarity may influence eyewitness memory, more research needs to be conducted examining familiarity before any conclusions can be drawn on the impact of various stereotypes. As familiarity between an eyewitness and perpetrator has been neglected in research, the current study and the future research examining this variable is of great importance in order to inform legal practice.

References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 997-1010. doi:10.1037/0278-7393.25.4.997
- Amendola, K. L., & Wixted, J. T. (2015a). Comparing the diagnostic accuracy of suspect identifications made by actual eyewitnesses from simultaneous and sequential lineups in a randomized field trial. *Journal of Experimental Criminology*, *11*(2), 263-284. doi:10.1007/s11292-014-9219-2
- Amendola, K. L., & Wixted, J. T. (2015b). No possibility of a selection bias, but direct evidence of a simultaneous superiority effect: A reply to wells et al. *Journal of Experimental Criminology*, *11*(2), 291-294. doi:10.1007/s11292-015-9227-x
- Araya, T., Akrami, N., & Ekehammar, B. (2003). Forgetting congruent and incongruent stereotypical information. *The Journal of Social Psychology*, *143*(4), 433-449. doi:10.1080/00224540309598455
- Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*(2), 272-281. doi:10.1037/0022-3514.65.2.272
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY: Guilford Press.

- Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology, 49*(5), 1129-1146. doi:10.1037/0022-3514.49.5.1129
- Beal, C. R., Schmitt, K. L., & Dekle, D. J. (1995). Eyewitness identification of children: Effects of absolute judgments, nonverbal response options, and event encoding. *Law and Human Behavior, 19*(2), 197-216. doi:10.1007/BF01499325
- Beaudry, J. L., & Lindsay, R. C. L. (2006). Current identification procedure practices: A survey of Ontario police officers. *The Canadian Journal of Police and Security Services, 4*(4), 178-183.
- Beggs, J. M., & Doolittle, D. C. (1993). Perceptions now and then of occupational sex typing: A replication of Shinar's 1975 study. *Journal of Applied Social Psychology, 23*(17), 1435-1453. doi:10.1111/j.1559-1816.1993.tb01042.x
- Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*(6), 1142-1163. doi:10.1037/0022-3514.70.6.1142
- Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87*(6), 763-778. doi:10.1037/0022-3514.87.6.763
- Bodenhausen, G. V., Kramer, G. P., & Süsler, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology, 66*(4), 621-632. doi:10.1037/0022-3514.66.4.621

- Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology, 24*(1), 45-62. doi:10.1002/ejsp.2420240104
- Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory strength. *Psychology, Crime & Law, 18*(5), 473-490. doi:10.1080/1068316X.2010.508458
- Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology, 103*(4), 751-757. doi:10.1037/h0037190
- Bowers, S., & MacLin, K. (2004, March). *Sources of criminal stereotypes and perceptions of the criminal*. Poster session presented at the American Psychology-Law Society conference, Scottsdale, Arizona.
- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3-47. doi:10.1016/0273-2297(90)90003-M
- Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100*(1), 42-67. doi:10.1037/0033-295X.100.1.42
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*(5), 164-169. doi:10.1111/1467-8721.00192
- Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull and R. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum Associates.

- Brewer, M. B., & Lui, L. N. (1989). The primacy of age and sex in the structure of person categories. *Social Cognition*, 7(3), 262-274. doi:10.1521/soco.1989.7.3.262
- Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, 12(1), 11-30. doi:10.1037/1076-898X.12.1.11
- Bruce, V., Burton, M., & Hancock, P. (2007). Remembering faces. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology: Vol. 2. Memory for people* (pp. 87-100). Mahwah, NJ: Erlbaum Associates.
- Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, 7(3), 207-218. doi:10.1037/1076-898X.7.3.207
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64(2), 123-152. doi:10.1037/h0043805
- Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science*, 23, 1-31. doi:10.1016/S0364-0213(99)80050-0
- Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, 10(3), 243-248. doi:10.1111/1467-9280.00144
- Byers, J. C., Bittner, A. C. Jr., & Hill, S. G. (1989). Traditional and raw Task Load Index (TLX) correlations: Are paired comparisons necessary? In A. Mital (Ed.), *Advances in industrial ergonomics and safety I* (pp. 481-485). London, England: Taylor & Francis.

- Cano, I., Hopkins, N., & Islam, M. R. (1991). Memory for stereotype-related material—a replication study with real-life social groups. *European Journal of Social Psychology, 21*(4), 349-357. doi:10.1002/ejsp.2420210407
- Cary, M., & Reder, L. M. (2003). A dual-process account of the list-length and strength-based mirror effects in recognition. *Journal of Memory and Language, 49*(2), 231-248. doi:10.1016/S0749-596X(03)00061-5
- Clark, S. E., Howell, R. T., & Davey, S. L. (2008). Regularities in eyewitness identification. *Law and Human Behavior, 32*(3), 187-218. doi:10.1007/s10979-006-9082-4
- Clutterbuck, R., & Johnston, R. A. (2005). Demonstrating how unfamiliar faces become familiar using a face matching task. *European Journal of Cognitive Psychology, 17*(1), 97-116. doi:10.1080/09541440340000439
- Connors, E., Lundregan, T., Miller, N., & McEwen, T. (1996). *Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial*. Washington, DC: US Department of Justice, National Institute of Justice.
- Crocker, J., Hannah, D. B., & Weber, R. (1983). Person memory and causal attributions. *Journal of Personality and Social Psychology, 44*(1), 55-66. doi:10.1037/0022-3514.44.1.55
- Curran, T., & Hancock, J. (2007). The FN400 indexes familiarity-based recognition of faces. *Neuroimage, 36*(2), 464-471. doi:10.1016/j.neuroimage.2006.12.016
- Cutler, B., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. England: Cambridge University Press.
- Cutler, B. L., & Penrod, S. D. (1988). Improving the reliability of eyewitness identification: Lineup construction and presentation. *Journal of Applied Psychology, 73*(2), 281-290. doi:10.1037/0021-9010.73.2.281

- Davies, G. M. (1995). Fuzzy-trace theory and eyewitness memory. *Learning and Individual Differences, 7*(2), 111-114. doi:10.1016/1041-6080(95)90022-5
- Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law and Human Behavior, 4*(4), 243-260. doi:10.1007/BF01040617
- Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14*, 139-150. doi:10.1037/1076-898X.14.2.139
- DiDonato, L., & Strough, J. (2013). Do college students' gender-typed attitudes about occupations predict their real-world decisions? *Sex Roles, 68*(9), 536-549. doi:10.1007/s11199-013-0275-2
- Dijksterhuis, A., & van Knippenberg, A. (1995). Memory for stereotype-consistent and stereotype-inconsistent information as a function of processing pace. *European Journal of Social Psychology, 25*(6), 689-693. doi:10.1002/ejsp.2420250607
- Dobolyi, D. G., & Dodson, C. S. (2013). Eyewitness confidence in simultaneous and sequential lineups: A criterion shift account for sequential mistaken identification overconfidence. *Journal of Experimental Psychology: Applied, 19*(4), 345-357. doi:10.1037/a0034596
- Driscoll, D. M., Hamilton, D. L., & Sorrentino, R. M. (1991). Uncertainty orientation and recall of person-descriptive information. *Personality and Social Psychology Bulletin, 17*(5), 494-500. doi:10.1177/0146167291175003
- Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50*(3), 145-158. doi:10.1037/0003-066X.50.3.145

- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735-754. doi:10.1037/0022-3514.46.4.735
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123-174). Mahwah, NJ: Erlbaum Associates.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229-238. doi: 10.1037//0022-3514.50.2.229
- Fisher, R. P., Geiselman, R. E., & Raymond, D. S. (1987). Critical analysis of police interview techniques. *Journal of Police Science and Administration*, *15*(3), 177-185.
- Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*(6), 621-628. doi:10.1037/0003-066X.48.6.621
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). New York, NY: Academic Press.
- Flowe, H. R., Mehta, A., & Ebbesen, E. B. (2011). The role of eyewitness identification evidence in felony case dispositions. *Psychology, Public Policy, and Law*, *17*, 140-159. doi: 10.1097/a00121311
- Fyock, J., & Stangor, C. (1994). The role of memory biases in stereotype maintenance. *British Journal of Social Psychology*, *33*(3), 331-343. doi: 10.1111/j.2044-8309.1994.tb01029.x

- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*(5), 533-543.
doi:10.1002/acp.885
- Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9*(2), 215-227. doi:10.1348/1355325041719428
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509-517.
doi:10.1037/0022-3514.60.4.509
- Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition, 13*(1), 8-20. doi: 10.3758/BF03198438
- Glick, P., Wilk, K., & Perreault, M. (1995). Images of occupations: Components of gender and status in occupational stereotypes. *Sex Roles, 32*(9-10), 565-582.
doi:10.1007/BF01544212
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*, 32-41. doi:10.1016/j.neuropsychologia.2006.04.015
- Greenberg, M. S., Westcott, D. R., & Bailey, S. E. (1998). When believing is seeing: The effect of scripts on eyewitness memory. *Law and Human Behavior, 22*(6), 685-694.
doi:10.1023/A:1025758807624
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27. doi:10.1037/0033-295X.102.1.4

- Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology: Applied*, *15*(2), 140-152. doi:10.1037/a0015082
- Gross, G., & Hayne, H. (1996). Eyewitness identification by 5- to 6-year-old children. *Law and Human Behavior*, *20*(3), 359-373. doi:10.1007/BF01499028
- Haber, R. N., & Haber, L. (2001, November). *A meta-analysis of research on eyewitness lineup identification accuracy*. Paper presented at the Annual Convention of the Psychonomics Society, Orlando, FL.
- Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*(9), 330-337. doi:10.1016/S1364-6613(00)01519-9
- Harris, R. J., Lee, D. J., Hensley, D. L., & Schoen, L. M. (1988). The effect of cultural script knowledge on memory for stories over time. *Discourse Processes*, *11*(4), 413-431. doi:10.1080/01638538809544711
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam, Netherlands: Elsevier.
- Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*(1), 25-38. doi:10.1037/0022-3514.37.1.25
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*(4), 657-674. doi:10.1111/0022-4537.00234

- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47(1), 237-271. doi:10.1146/annurev.psych.47.1.237
- Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes* (1st ed.). London, England: Routledge.
- Holst, V. F., & Pezdek, K. (1992). Scripts for typical crimes and their effects on memory for eyewitness testimony. *Applied Cognitive Psychology*, 6(7), 573-587.
doi:10.1002/acp.2350060702
- Horry, R., Halford, P., Brewer, N., Milne, R., & Bull, R. (2014). Archival analyses of eyewitness identification test outcomes: What can they tell us about eyewitness memory? *Law and Human Behavior*, 38(1), 94-108. doi:10.1037/lhb0000060
- Horry, R., Memon, A., Wright, D. B., & Milne, R. (2012). Predictors of eyewitness identification decisions from video lineups in England: A field study. *Law and Human Behavior*, 36, 257-265. doi: 10.1037/h0093959.
- Jennings, S. A. (1975). Effects of sex typing in children's stories on preference and recall. *Child Development*, 46(1), 220-223. doi:10.1111/j.1467-8624.1975.tb03294.x
- Johnston, W. A., & Hawley, K. J. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin and Review*, 1, 56-72. doi: 10.3758/BF03200761.
- Jost, T. J., & Banaji, R. M. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, 33(1), 1-27. doi: 10.1111/j.2044-8309.1994.tb01008.x
- Kaplan, M. F., Wanshula, L. T., & Zanna, M. P. (1993). Time pressure and information integration in social judgment: The effect of need for structure. In O. Svenson & A. J.

- Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 255-267). New York, NY: Plenum Press.
- Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition*, *12*, 60-70. doi:10.3758/BF03196998
- Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, *4*(3), 265-273. doi:10.1080/713756764
- Koblinsky, S. G., Cruse, D. F., & Sugawara, A. I. (1978). Sex role stereotypes and children's memory for story content. *Child Development*, *49*(2), 452-458. doi:10.1111/j.1467-8624.1978.tb02335.x
- Koomen, W., & Dijkster, A. J. (1997). Ingroup and outgroup stereotypes and selective processing. *European Journal of Social Psychology*, *27*(5), 589-601. doi:10.1002/(SICI)1099-0992(199709/10)27:5<589::AID-EJSP840>3.0.CO;2-Y
- Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*(1), 481-537. doi:10.1146/annurev.psych.51.1.481
- Kuehn, L. L. (1974). Looking down a gun barrel: Person perception and violent crime. *Perceptual and Motor Skills*, *39*(3) 1159-1164. doi:10.2466/pms.1974.39.3.1159
- Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*(4), 522-544. doi:10.1037/0033-2909.129.4.522
- Leippe, M. R., Romanczyk, A., & Manion, A. P. (1991). Eyewitness memory for a touching experience: Accuracy differences between child and adult witnesses. *Journal of Applied Psychology*, *76*(3), 367-379. doi:10.1037/0021-9010.76.3.367

- Liben, L. S., & Signorella, M. L. (1980). Gender-related schemata and constructive memory in children. *Child Development, 51*(1), 11-18. doi:10.1111/j.1467-8624.1980.tb02503.x
- Lindsay, R. C. L., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology, 76*(5), 741-745. doi:10.1037/0021-9010.76.5.741
- Lindsay, R. C. L., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions: A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior, 18*(5), 527-541. doi: 10.1007/BF01499172
- Lindsay, R. C. L., & Wells, G. L. (1985). Improving eyewitness identification from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology, 70*(3), 556-564. doi:10.1037/0021-9010.70.3.556
- Lipton, J. P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology, 62*(1), 90-95. doi:10.1037/0021-9010.62.1.90
- Luus, C. A. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior, 15*, 43-57. doi: 10.1007/BF01044829
- Mackay, T. L., & Paterson, H. M. (2015). How does timing of recall affect eyewitness memory and psychological distress? *Journal of Police and Criminal Psychology, 30*(4), 242-253. doi:10.1007/s11896-014-9156-z
- MacLin, M. K., & Herrera, V. (2006). The criminal stereotype. *North American Journal of Psychology, 8*(2), 197-208.
- Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology, 23*(1), 77-87. doi:10.1002/ejsp.2420230107

- Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66*(1), 37-47. doi:10.1037/0022-3514.66.1.37
- Madriz, E. I. (1997). Images of criminals and victims: A study on women's fear and social control. *Gender and Society, 11*(3), 342-356. doi:10.1177/089124397011003005
- Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior, 5*(4), 299-309. doi:10.1007/BF01044945
- Malpass, R. S. (2015). A lineup evaluation “Do it yourself kit” for attorneys and law enforcement. Retrieved on June 18, 2018 from: <http://eyewitness.utep.edu/diy.html>
- Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*(4), 482-489. doi:10.1037/0021-9010.66.4.482
- Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual-process theories of recognition. *Perspectives on Psychological Science, 3*, 390-399. doi: 10.1111/j.1745-6924.2008.00087.x
- Mandler, G., Pearlstone, Z., & Koopmans, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior, 8*(3), 410-423. doi:10.1016/S0022-5371(69)80134-9
- Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*(6), 903-933. doi:10.1037//0033-2909.128.6.903
- McNichol, S., Shute, R., & Tucker, A. (1999). Children's eyewitness memory for a repeated event. *Child Abuse & Neglect, 23*(11), 1127-1139. doi:10.1016/S0145-2134(99)00084-8

- Meissner, C. A., Sporer, S. L., & Schooler, J. W. (2007). Person descriptions as eyewitness evidence. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *The handbook of eyewitness psychology: Vol. 2. Memory for people* (pp. 3-34). Mahwah, NJ: Erlbaum Associates.
- Melara, R. D., DeWitt-Rickards, T. S., & O'Brien, T. P. (1989). Enhancing lineup identification accuracy: Two codes are better than one. *Journal of Applied Psychology, 74*(5), 706-713. doi:10.1037/0021-9010.74.5.706
- Memon, A., Havard, C., Clifford, B., Gabbert, F., & Watt, M. (2011). A field evaluation of the VIPER system: A new technique for eliciting eyewitness identification evidence. *Psychology, Crime, & Law, 17*(8), 711-729. doi: 10.1080/10683160903524333
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology, 94*(3), 339-354. doi: 10.1348/000712603767876262
- Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied, 18*(4), 361-376. doi:10.1037/a0030609
- Moskowitz, G. B. (1993). Individual differences in social categorization: The influence of personal need for structure on spontaneous trait inferences. *Journal of Personality and Social Psychology, 65*(1), 132-142. doi:10.1037/0022-3514.65.1.132
- O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences, 6*(6), 261-266. doi:10.1167/2.7.604

- Odinot, G., & Wolters, G. (2006). Repeated recall, retention interval and the accuracy-confidence relation in eyewitness memory. *Applied Cognitive Psychology, 20*(7), 973-985. doi: 10.1002/acp.1263
- Osborne, D., & Davies, P. G. (2013). Eyewitness identifications are affected by stereotypes about a suspect's level of perceived stereotypicality. *Group Processes & Intergroup Relations, 16*(4), 488-504. doi:10.1177/1368430212454927
- Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63-71. doi:10.1207/S15326985EP3801_8
- Parker, J. E., & Ryan, V. (1993). An attempt to reduce guessing behaviour in children's and adults' eyewitness identifications. *Law and Human Behavior, 17*(1), 11-26. doi:10.1007/BF01044534
- Paterson, H. M., Kemp, R. I., & Ng, J. R. (2011). Combating Co-witness contamination: Attempting to decrease the negative effects of discussion on eyewitness memory. *Applied Cognitive Psychology, 25*(1), 43-52. doi:10.1002/acp.1640
- Patterson, K. E., & Baddeley, A. D. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory, 3*(4), 406-417. doi:10.1037/0278-7393.3.4.406
- Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory and Language, 35*(2), 261-285. doi:10.1006/jmla.1996.0015

- Perrault, S. (2015). *Criminal Victimization in Canada, 2014*. (Catalogue number 85-002-X201500114241). Retrieved from Statistics Canada Website
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2015001/article/14241-eng.htm#a8>
- Pezdek, K., & Stolzenberg, S. (2014). Are individuals' familiarity judgments diagnostic of prior contact? *Psychology, Crime & Law*, 20(4), 302-314. doi:10.1080/1068316X.2013.772181
- Plass, J. L., Kalyuga, S., & Leutner, D. (2010). Individual differences and cognitive load theory. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 65 – 87). Cambridge, NY: Cambridge University Press.
- Police Executive Research Forum. (2013). *A national survey of eyewitness identification procedures in law enforcement agencies* (Document No. 242617). Retrieved from
<https://www.ncjrs.gov/pdffiles1/nij/grants/242617.pdf>
- Pozzulo, J. D. (2016). *The Young Eyewitness – How Well Do Children and Adolescents Describe and Identify Perpetrators*. Washington, DC: American Psychological Association.
- Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22(5), 549-570.
doi:10.1023/A:1025739514042
- Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child eyewitnesses. *Journal of Applied Psychology*, 84(2), 167-176.
doi:10.1037/0021-9010.84.2.167
- Pozzulo, J. D., Reed, J., Pettalia, J., & Dempsey, J. (2015). Simultaneous, sequential, elimination, and wildcard: A comparison of lineup procedures. *Journal of Police and Criminal Psychology*, 31, 71-80. doi:10.1007/s11896-015-9168-3

- Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology, 88*(2), 315-323.
doi:10.1037/0021-9010.88.2.315
- R. v. Hanemaayer, ONCA 580 (2008)
- R. v. Sophonow, 2 S.C.R. 524. (1985).
- Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 163-178.
doi:10.1037/0278-7393.16.2.163
- Reed, J. P., & Reed, R. S. (1973). Status, images, and consequences: Once a criminal always a criminal. *Sociology and Social Research, 57*(4), 460-472.
- Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*(2), 178-191. doi:10.1037/0012-1649.30.2.178
- Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences, 9*(2), 95-123. doi:10.1016/S1041-6080(97)90002-9
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender and Society, 18*(4), 510-531.
doi:10.1177/0891243204265269
- Robinson, R. J., Keltner, D., Ward, A., & Ross, L. (1995). Actual versus assumed differences in construal: "naive realism" in intergroup perception and conflict. *Journal of Personality and Social Psychology, 68*(3), 404-417. doi:10.1037/0022-3514.68.3.404

- Ruble, D. N., & Martin, C. L. (1998). Gender development. In W. Damon (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 933-1016). New York, NY: Wiley.
- Sauerland, M., Sagana, A., & Sporer, S. L. (2012). Assessing nonchoosers' eyewitness identification accuracy from photographic showups by using confidence and response times. *Law and Human Behavior, 36*(5), 394-403. doi:10.1037/h0093926
- Shapiro, L. R. (2009). Eyewitness testimony for a simulated juvenile crime by male and female criminals with consistent or inconsistent gender-role characteristics. *Journal of Applied Developmental Psychology, 30*(6), 649-666. doi:10.1016/j.appdev.2009.07.007
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139-156. doi:10.1037/0033-2909.100.2.139
- Sheahan, C. L., Pica, E., Pozzulo, J. D., & Nastasa, C. (2017). Eyewitness recall and identification abilities of adolescent and young-adults. *Journal of Applied Developmental Psychology, 53*, 86-95. doi:10.1016/j.appdev.2017.09.008
- Sheahan, C. L., Pozzulo, J. D., Reed, J. E., & Pica, E. (2018). The role of familiarity with the defendant, type of descriptor discrepancy, and eyewitness age on mock jurors' perceptions of eyewitness testimony. *Journal of Police and Criminal Psychology, 33*(1), 35-44. doi:10.1007/s11896-017-9232-2
- Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype-relevant information under cognitive load. *Personality and Social Psychology Bulletin, 26*(1), 26-34. doi:10.1177/0146167200261003

- Sherman, J. W., & Hamilton, D. L. (1994). On the formation of interitem associative links in person memory. *Journal of Experimental Social Psychology, 30*(3), 203-217. doi:10.1006/jesp.1994.1010
- Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology, 75*(3), 589-606. doi:10.1037/0022-3514.75.3.589
- Sherman, J. W., Stroessner, S. J., Conrey, F. R., & Azam, O. A. (2005). Prejudice and stereotype maintenance processes: Attention, attribution, and individuation. *Journal of Personality and Social Psychology, 89*(4), 607-622. doi:10.1037/0022-3514.89.4.607
- Shinar, E. H. (1975). Sexual stereotypes of occupations. *Journal of Vocational Behavior, 7*(1), 99-111. doi:10.1016/0001-8791(75)90037-8
- Smith, V. L., & Studebaker, C. A. (1996). What do you expect?: The influence of people's prior knowledge of crime categories on fact-finding. *Law and Human Behavior, 20*(5), 517-532. doi:10.1007/BF01499039
- Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78*(1), 22-33. doi:10.1037/0021-9010.78.1.22
- Srull, T. K. (1981). Person memory: Some tests of associate storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 440-463. doi:10.1037/0278-7393.7.6.440
- Srull, T. K., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(2), 316-345. doi:10.1037/0278-7393.11.2.316

- Strull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37*(10), 1660-1672. doi:10.1037/0022-3514.37.10.1660
- Strull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review, 96*(1), 58-83. doi:10.1037/0033-295X.96.1.58
- Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27*(4), 357-378. doi:10.1016/0022-1031(91)90031-Z
- Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*(1), 42-61. doi:10.1037/0033-2909.111.1.42
- Stangor, C., & Ruble, D. N. (1989). Stereotype development and memory: What we remember depends on how much we know. *Journal of Experimental Social Psychology, 25*, 18-35.
- Stebly, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior, 25*(5), 459-473. doi:10.1023/A:1012888715007
- Stebly, N. K., Dietrich, H. L., Ryan, S. L., Raczynski, J. L., & James, K. A. (2011). Sequential lineup laps and eyewitness accuracy. *Law and Human Behavior, 35*(4), 262-274. doi:10.1007/s10979-010-9236-2
- Stebly, N. K., Dysart, J. E., & Wells, G. L. (2015). An unrepresentative sample is unrepresentative regardless of the reason: A rejoinder to amendola and wixted. *Journal of Experimental Criminology, 11*(2), 295-298. doi:10.1007/s11292-015-9233-z

- Stern, L. D., Marrs, S., Millar, M. G., & Cole, E. (1984). Processing time and the recall of inconsistent and consistent behaviors of individuals and groups. *Journal of Personality and Social Psychology, 47*(2), 253-262. doi:10.1037/0022-3514.47.2.253
- Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1379-1396. doi:10.1037/0278-7393.24.6.1379
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296. doi:10.1023/A:1022193728205
- Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A guide for law enforcement*. National Institute of Justice, U.S. Department of Justice.
- Technical Working Group for Eyewitness Evidence. (2003). *Eyewitness evidence: A trainer's manual for law enforcement*. National Institute of Justice, U.S. Department of Justice.
- Theunissen, T. P. M., Meyer, T., Memon, A., & Weinsheimer, C. C. (2017). Adult eyewitness memory for single versus repeated traumatic events. *Applied Cognitive Psychology, 31*(2), 164-174. doi:10.1002/acp.3314
- Tredoux, C. G. (1998). Statistical inference on measures of lineup fairness. *Law and Human Behavior, 22*(2), 217-237. doi:10.1023/A:1025746220886
- Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior, 20*(5), 479-496. doi:10.1016/S0022-5371(81)90129-8
- Turtle, J., Lindsay, R. C. L., & Wells, G. L. (2003). Best practice recommendations for eyewitness evidence procedures: New ideas for the oldest way to solve a case. *Canadian Journal of Police and Security Services, 1*, 5-18.

- Ullian, D. Z. (1976). The development of conceptions of masculinity and femininity. In B. Lloyd, & J. Archer (Eds.), *Exploring sex differences* (pp. 25-47). London, England: Academic Press.
- Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology, 17*, 969-993. doi:10.1002/acp.939
- van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177. doi:10.1007/s10648-005-3951-0
- von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld [The effects of field formation in the trace field]. *Psychologische Forschung, 18*, 299-342.
- Wang, E., Paterson, H., & Kemp, R. (2014). The effects of immediate recall on eyewitness accuracy and susceptibility to misinformation. *Psychology, Crime, & Law, 20*(7), 619-634. doi: 10.1080/1068316x.2013.854788
- Welch-Ross, M. K., & Schmidt, C. R. (1996). Gender-schema development and children's constructive story memory: Evidence for a developmental model. *Child Development, 67*(3), 820-835. doi:10.1111/j.1467-8624.1996.tb01766.x
- Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48*(5), 553-571. doi:10.1037/0003-066X.48.5.553
- Wells, G. L. (2014). Eyewitness identification: Probative value, criterion shifts, and policy regarding the sequential lineup. *Current Directions in Psychological Science, 23*(1), 11-16. doi:10.1177/0963721413504781

- Wells, G. L., Dysart, J. E., & Steblay, N. K. (2015). The flaw in amendola and wixted's conclusion on simultaneous versus sequential lineups. *Journal of Experimental Criminology, 11*(2), 285. doi:10.1007/s11292-014-9225-4
- Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55*(6), 581-598. doi:10.1037/0003-066X.55.6.581
- Wells, G. L., Yang, Y., & Smalarz, L. (2015). Eyewitness identification: Bayesian information gain, base-rate effect equivalency curves, and reasonable suspicion. *Law and Human Behavior, 39*(2), 99-122. doi:10.1037/lhb0000125
- White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles, 55*(3), 259-266. doi:10.1007/s11199-006-9078-z
- Wilbourn, M. P., & Kee, D. W. (2010). Henry the nurse is a doctor too: Implicitly examining children's gender stereotypes for male and female occupational roles. *Sex Roles, 62*(10), 670-683. doi:10.1007/s11199-010-9773-7
- Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory, 7*(3), 181-190. doi:10.1037/0278-7393.7.3.181
- Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *The American Psychologist, 70*(6), 515-526. doi:10.1037/a0039510
- Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest, 18*(1), 10-65. doi:10.1177/1529100616686966

- Wolfe, C. R., Reyna, V. F., & Brainerd, C. J. (2005). Fuzzy-trace theory: Implications for transfer in teaching and learning. In J.M. Royer (Series Ed.), *Current Perspectives on Cognition, Learning, and Instruction: Transfer of learning from a modern multidisciplinary perspective* (pp. 53-87). Greenwich, CT: Information Age.
- Xiao, N. G., Perrotta, S., Quinn, P. C., Wang, Z., Sun, Y. P., & Lee, K. (2014). On the facilitative effects of face motion on face recognition and its development. *Frontiers in Psychology, 5*, 1-16. doi:10.3389/fpsyg.2014.00633
- Yarmey, A. D., Jacob, J., & Porter, A. (2002). Person recall in field settings. *Journal of Applied Social Psychology, 32*(11), 2354-2367. doi:10.1111/j.1559-1816.2002.tb01866.x
- Yovel, G., & Paller, K. A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *Neuroimage, 21*(2), 789-800. doi:10.1016/j.neuroimage.2003.09.034
- Zijlstra, F. R. H. (1993). *Efficiency in work behavior: A design approach for modern tools* (Doctoral thesis, Delft University of Technology, Delft, The Netherlands). Retrieved from <https://repository.tudelft.nl/islandora/object/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f?collection=research>
- Zijlstra, F. R. H., van Doorn, L. (1985). *The construction of a scale to measure perceived effort*. Technical Report. Delft University of Technology.

Appendix A**SONA Recruitment Notice**

Study Name: I Spy: Fill in the missing details

Description: Participation in this study will require that you watch short videos and answer questions about those videos.

Eligibility Requirements: Must be 18-years-old or older. This is a visual activity. If you have any vision problems that interfere with your ability to view the video, you are ineligible to participate in the current study.

Duration: 60 minutes

Location: Social Sciences Research Building, Room 111

Compensation: 1% in PSYC 1001, PSYC 1002, PSYC 2001, PSYC 2002, NEUR 2001, or NEUR2002

Researchers

Primary Investigator: Lauren Thompson, Psychology Department, MA Student

Phone: 613-520-2600, ext. 3695

Email: LaurenE.Thompson@carleton.ca

Faculty Advisor: Dr. Joanna Pozzulo

This study has received clearance by the Carleton University Research Ethics Board-B (105782)

Appendix B

Informed Consent Form

The purpose of informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. Informed consent must provide sufficient information such that you have the opportunity to determine whether or not you wish to participate in the study.

Present study: I Spy: Fill in the missing details

Research personnel: The following people will be involved in this research project and may be contacted at any time: Lauren Thompson (Principal Investigator, laurene.thompson@carleton.ca, 613-520-2600, ext. 3695) or Dr. Joanna Pozzulo (Faculty Advisor, joanna.pozzulo@carleton.ca, 613-520-2600, ext. 1412).

Concerns: If you should have any ethical concerns about this study please contact, Dr. Andy Adler (Chair, Carleton University Research Ethics Board-B, adler@sce.carleton.ca, 613-520-2600, ext. 8785). You can also contact the Carleton University Research Compliance Office at ethics@carleton.ca for any other concerns.

Purpose: The purpose of this study is to examine how you view and process details of a scene.

Task requirements: You will be asked to watch videos and then complete questionnaires. You must be 18-years-old or older. This is a visual activity. If you have any vision problems that prohibit you from viewing a video, you are ineligible to participate in the current study.

Duration and locale: Testing will take place in Room 111, Social Sciences Research Building, Carleton University. This study will be completed in one testing session, lasting approximately 60 minutes.

Token for participation: You will receive a 1% increase in your final grade of PSYC 1001, PSYC 1002, PSYC 2001, PSYC 2002, NEUR 2001, or NEUR2002 for participating in this study.

Potential risk/discomfort: There are no potential risks involved in this experiment. Should you experience any unease, you have the right to withdraw from the study and still receive course credit.

Anonymity/Confidentiality: All the information you provide will be strictly confidential. Data will only be used for research at Carleton University. Your answers will NOT be linked to your name or signature (i.e., consent form) and your responses will be coded in such a way that you cannot be identified.

Protection of Personal Information: This Informed Consent Form will be kept in the Lab for seven years. It will be placed in a room that has restricted access and is kept locked and closed at all times. At the seven-year mark, it will be shredded and disposed of.

Right to withdraw: Your participation is strictly voluntary. At any point during the study you have the right not to complete certain questions or to withdraw from the study without any penalty whatsoever by informing the researcher. However, once the study session has finished, participants cannot withdraw because the data is not linked to any names.

This study has received clearance by the Carleton University Research Ethics Board-B (105782)

Signatures: *I have read the above form and hereby consent to participate in this study. The data in this study will be used for research publications and/or teaching purposes. I am aware that the data collected in this study will be kept strictly confidential and anonymous. My signature indicates that I understand the above and wish to participate in this study.*

Participant's Name (print): _____

Participant's Signature: _____

Researcher's Name (print): _____

Researcher's Signature: _____

Date: _____

Appendix C

Stereotype Profile Descriptions

Female Nurse:

“Hi, my name is Jessie. I’ve been working as a nurse in the Intensive Care Unit at the Civic Hospital in Ottawa for the last two years. My time as a nurse requires the care for extremely ill patients. I provide patient care, which among many other things can include bathing and grooming, pain management, taking vital signs, and making my patients as comfortable as I possibly can. Most of the individuals I care for are suffering due to cardiac disease or brain injury. We see a lot of car accident victims in the critical care unit. I work closely with the doctors, hospital team members, and the patient’s loved ones. My job requires intensive mental and emotional stamina as it is quite a difficult unit to be a part of.”

Male Nurse:

“Hi, my name is Jessie. I’ve been working as a nurse in the Intensive Care Unit at the Civic Hospital in Ottawa for the last two years. My time as a nurse requires the care for extremely ill patients. I provide patient care, which among many other things can include bathing and grooming, pain management, taking vital signs, and making my patients as comfortable as I possibly can. Most of the individuals I care for are suffering due to cardiac disease or brain injury. We see a lot of car accident victims in the critical care unit. I work closely with the doctors, hospital team members, and the patient’s loved ones. My job requires intensive mental and emotional stamina as it is quite a difficult unit to be a part of.”

Female Construction Worker:

“Hi, my name is Jessie. I started working in construction as a full time job during the summers between school years. Now I’m a full time construction worker. I have been responsible for several construction jobs including, framing, drywalling, and roofing. My job title now is construction manager. Generally speaking I am in charge of a group of construction workers and projects. I have to schedule projects to meet deadlines, dispatch workers to construction sites, and oversee the projects in the field so that they stay on schedule and the project is completed according to specification. I am required to obtain necessary permits and licenses and to assure that my workers are following the proper building and safety codes. My number one responsibility is the safety of my crew members.”

Male Construction Worker:

“Hi, my name is Jessie. I started working in construction as a full time job during the summers between school years. Now I’m a full time construction worker. I have been responsible for several construction jobs including, framing, drywalling, and roofing. My job title now is construction manager. Generally speaking I am in charge of a group of construction workers and projects. I have to schedule projects to meet deadlines, dispatch workers to construction sites, and oversee the projects in the field so that they stay on schedule and the project is completed according to specification. I am required to obtain necessary permits and

licenses and to assure that my workers are following the proper building and safety codes. My number one responsibility is the safety of my crew members.”

Appendix D1 - Stereotype Questionnaire

Please read the questions provided below. For each question, circle the most appropriate answer that conforms to your opinion.

1. How well did Jessie's appearance fit your perception of what a nurse looks like?

1	2	3	4	5	6	7
Consistent						Inconsistent

2. How well did Jessie's job description fit your perception of what a nurse does?

1	2	3	4	5	6	7
Consistent						Inconsistent

Appendix D2 - Stereotype Questionnaire

Please read the questions provided below. For each question, circle the most appropriate answer that conforms to your opinion.

1. How well did Jessie's appearance fit your perception of what a construction worker looks like?

1	2	3	4	5	6	7
Consistent						Inconsistent

2. How well did Jessie's job description fit your perception of what a construction worker/manager does?

1	2	3	4	5	6	7
Consistent						Inconsistent

Appendix E

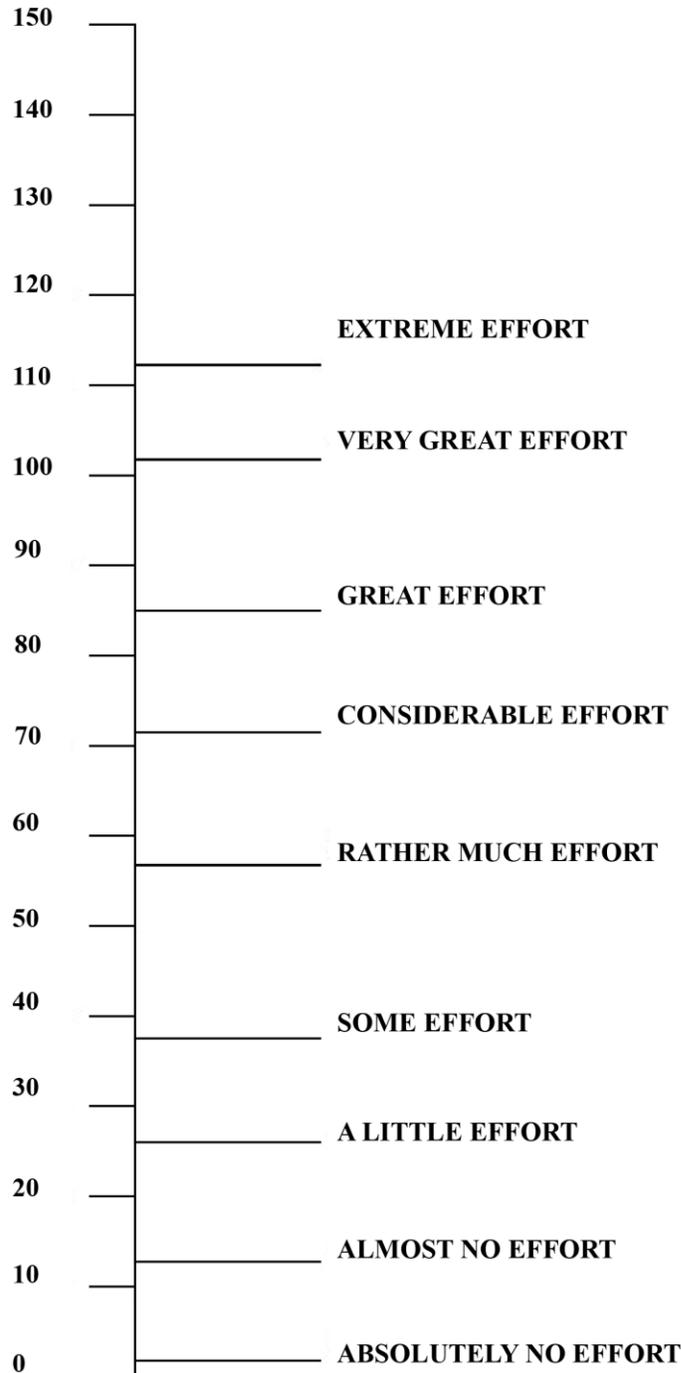
Open Recall

Write down everything you remember about the person in the video.

Appendix F

Cognitive Load Measure (RSME)

Directions: Based on the previous task of watching the video of the person **explaining their job/stealing the laptop**, consider how demanding it was to watch and pay attention to the various details of the video. Mark the line with an “X” at the point corresponding to the amount of mental effort used when watching the video and paying attention to the video.



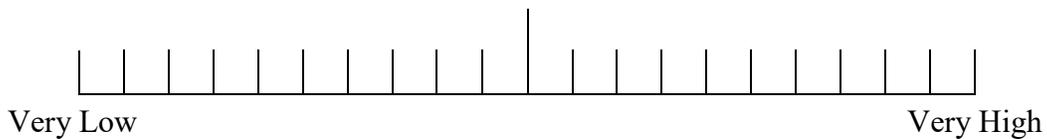
Appendix G

Cognitive Load Measure (NASA-RTLX)

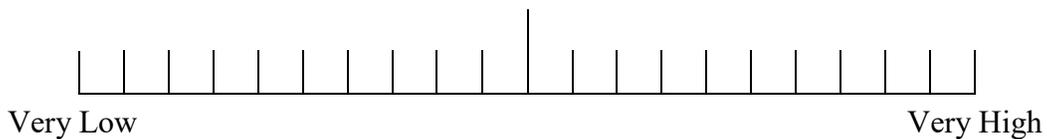
Directions: Based on the previous task of watching the video of the person **explaining their job/stealing the laptop**, consider how demanding it was to watch and pay attention to the various details of the video. Evaluate this task by putting an “X” on each of the six scales at the point which matches your experience.

Note: “PERFORMANCE” goes from “good” on the left to “poor” on the right. This order has been confusing for some people.

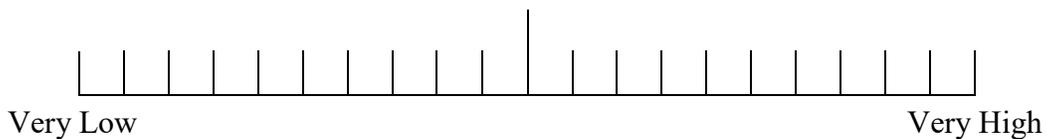
1. MENTAL DEMAND: How mentally demanding was the task?



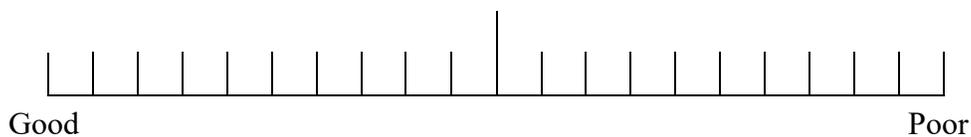
2. PHYSICAL DEMAND: How physically demanding was the task?



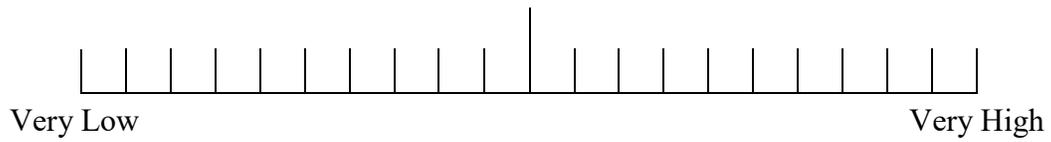
3. TEMPORAL DEMAND: How hurried or rushed was the pace of the task?



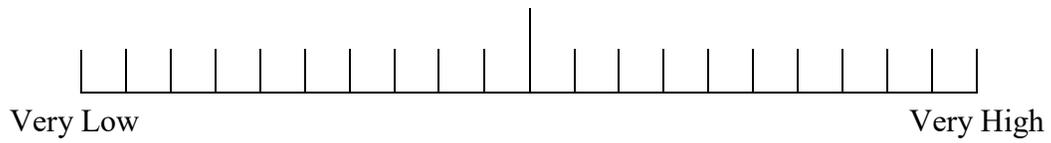
4. PERFORMANCE: How successful were you in accomplishing what you were asked to do?



5. EFFORT: How hard did you have to work to accomplish your level of performance?



6. FRUSTRATION: How insecure, discouraged, irritated, stressed, and annoyed were you?



Appendix H

Secondary Informed Consent Form

The purpose of informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. Informed consent must provide sufficient information such that you have the opportunity to determine whether or not you wish to participate in the study.

Present study: Stereotypes, Familiarity and the Effects on Eyewitness Identification

Research personnel: The following people will be involved in this research project and may be contacted at any time: Lauren Thompson (Principal Investigator, laurene.thompson@carleton.ca, 613-520-2600, ext. 3695) or Dr. Joanna Pozzulo (Faculty Advisor, joanna.pozzulo@carleton.ca, 613-520-2600, ext. 1412).

Concerns: If you should have any ethical concerns about this study please contact, Dr. Andy Adler (Chair, Carleton University Research Ethics Board-B, adler@sce.carleton.ca, 613-520-2600, ext. 8785). You can also contact the Carleton University Research Compliance Office at ethics@carleton.ca for any other concerns.

Purpose: The purpose of this study is to examine different factors potentially influencing accuracy with remembering what and who you saw. You will be asked to describe the video and you will then be presented with photographs. When you signed up to participate in this study you were told that the purpose was to examine your ability to view and process details of a scene; this was a deceptive tool used to prevent you from knowing the true purpose of the study, to measure eyewitness accuracy. This was done in order to mimic a real life eyewitness situation as best as possible. In real life, witnessing a crime is unforeseen and unpredictable. Had you known you were intended to watch a crime video your attention to specific details of the crime, such as the perpetrator's characteristics, may have been altered. As a result, your recall and identification may not have represented witnessing a real criminal act. At this time, you have the opportunity to withdraw your data without penalty. However, if you wish to sign this informed consent form, you consent to finishing the study having known its true nature.

Task requirements: You will be asked to watch videos and then complete questionnaires regarding the viewed videos. You must be 18-years-old or older. This is a visual activity. If you have any vision problems that prohibit you from viewing a video, you are ineligible to participate in the current study.

Duration and locale: Testing will take place in Room 111, Social Sciences Research Building, Carleton University. This study will be completed in one testing session, lasting approximately 45 minutes.

Token for participation: You will receive a 1% increase in your final grade of PSYC 1001, PSYC 1002, PSYC 2001, PSYC 2002, NEUR 2001, or NEUR2002 for participating in this study.

Potential risk/discomfort: There are no potential risks involved in this experiment. Should you experience any unease, you have the right to withdraw from the study and still receive course credit.

Anonymity/Confidentiality: All the information you provide will be strictly confidential. Data will only be used for research at Carleton University. Your answers will NOT be linked to your name or signature (i.e., consent form) and your responses will be coded in such a way that you cannot be identified.

Protection of Personal Information: This Informed Consent Form will be kept in the Lab for seven years. It will be placed in a room that has restricted access and is kept locked and closed at all times. At the seven year mark, it will be shredded and disposed of.

Right to withdraw: Your participation is strictly voluntary. At any point during the study you have the right not to complete certain questions or to withdraw from the study without any penalty whatsoever by informing the researcher. However, once the study session has finished, participants cannot withdraw because the data is not linked to any names.

This study has received clearance by the Carleton University Research Ethics Board-B (105782)

Signature: *I have read the above form, and now that I know the true purpose of this study:*

- I consent to the use of my data
- I do not consent to the use of my data

Participant's Name (print): _____

Participant's Signature: _____

Researcher's Name (print): _____

Researcher's Signature: _____

Date: _____

Appendix I
Demographics Form

Your age: _____

Your sex: _____

Ethnicity: Please indicate which ethnic group you would consider yourself to belong to by checking the appropriate box (optional):

- White (e.g., European)
- Black (e.g., African, African American, African Canadian, Caribbean)
- East Asian (e.g., Chinese, Japanese, Korean, Polynesian)
- South Asian (e.g., Indian, Pakistani, Sri Lankan, Bangladeshi)
- Southeast Asian (e.g., Burmese, Cambodian, Filipino, Laotian, Malaysian, Thai, Vietnamese)
- West Asian (e.g., Arabian, Armenian, Iranian, Israeli, Lebanese, Palestinian, Syrian, Turkish)
- Latin American (e.g., Mexican, Indigenous Central, South American)
- Aboriginal Canadian/Native Canadian/First Nations
- Mixed origin, please specify: _____
- Other: _____

Appendix J
Description Form

Event Description

Please write down everything that you can remember about the video.

Person Description

Please write down everything you can remember about the criminal in the video.

Appendix K**Simultaneous Lineup Response Form**

Think back to the video. Think back to what the criminal looks like. I am going to show you some pictures. Please look at the pictures. The criminal's picture may or may not be here. If you see the criminal's picture, please place a check mark in the box that corresponds with the criminal's lineup number. If you do not see the criminal, please place a check mark in the box labeled 'not here'.



Photo 1



Photo 2



Photo 3



Photo 4



Photo 5



Photo 6



Not Here

Please rate your confidence in the accuracy of your decision using a numerical value between 0 (*not at all confident*) and 100 (*very confident*): _____

Appendix L
Study Checks

1. Did you know that you would be acting as an eyewitness to a crime before coming to the study session today? Please circle the appropriate response.

YES NO

2. Did you know anyone from the criminal photo lineup (e.g., one of the boys/girls is your friend)?

YES NO

3. Before today, have you ever completed a photo lineup identification procedure? If yes, please explain.

YES NO

4. Have you ever taken a class on eyewitness memory?

YES NO

Appendix M

Debriefing Form

What are we trying to learn in this research?

The purpose of the present study is to examine whether gender stereotypes and familiarity with the perpetrator influence eyewitness recall and identification accuracy. Eyewitness research has focused on eyewitness' ability to recall and identify a stranger; little research has looked at how familiarity with a perpetrator affects eyewitness accuracy. Familiarity in this context can be understood as an individual who is not well known but with whom the witness has had brief previous contact with (i.e., familiar-stranger). Whether an eyewitness has previously come into contact with a perpetrator would likely affect their memory of the perpetrator. Additionally, there are factors that may effect how a witness encodes and remembers people met briefly. When meeting someone briefly it is likely that stereotypes will emerge. If this person occupies a gender inconsistent occupation versus a gender consistent occupation, this will likely affect your memory of them. Therefore, the purpose of this study is to look at how brief previous contact with a perpetrator (i.e. familiarity) will affect eyewitness recall and recognition memory. Additionally, we are interested in looking at factors that may affect the relationship between eyewitness accuracy and familiarity, specifically how inconsistent versus consistent occupational stereotypes influence accuracy.

Why is this important to scientists or to the general public?

This research is important because in many cases where an eyewitness is utilized the eyewitness has briefly come into contact with perpetrator previously. Understanding how familiarity and factors that may affect familiarity (e.g., stereotypes) influence eyewitness recall and recognition memory will help inform best practices within the criminal justice system.

What are our hypotheses and predictions?

We predict adults will provide more descriptors and more correct descriptors when they are familiar with the perpetrator compared to when the perpetrator is a stranger. Additionally, we predict that adults will be more accurate with identification when they are familiar with the perpetrator. We also predict that adults will show enhanced recall and recognition memory when they have had previous contact with the perpetrator (i.e., when familiar with the perpetrator) and the perpetrator has a stereotype inconsistent occupation compared to stereotype consistent occupation.

Where can I learn more?

Asai, C. (2001). The effect of familiarity on eyewitness identification testimony: The relationship between accuracy and confidence. *Japanese Journal of Psychology*, 72(4), 283-289.

Shapiro, L. R. (2009). Eyewitness testimony for a simulated juvenile crime by male and female criminals with consistent or inconsistent gender-role characteristics. *Journal of Applied Developmental Psychology*, 30(6), 649-666. doi:10.1016/j.appdev.2009.07.007

Why didn't you tell me I was going to watch a video of a crime?

We did not tell you that you would be watching a videotape of a theft because we wanted to mimic real life as much as possible and real life crime happens unexpectedly. Had you been forewarned of the crime aspect of the study, you may have paid attention to the video in a different way and, in turn, this would have influenced your responses to our questions. In this case, your responses would not have been useful to us because they would not be representative of how a real life witness would experience a similar situation.

Note. The video you watched was completely fictional, no crime was actually committed.

What if I feel distress or anxiety after participating in this study?

If you feel any distress or anxiety after participating in this study, please feel free to contact the Carleton University Health and Counseling Services at: 613-520-6674, or the Distress Centre of Ottawa and Region at 613-238-3311 (<http://www.dcottawa.on.ca>).

What if I have questions later?

If you wish to discuss this research any further feel free to contact any one of the following people: Lauren Thompson (Principal Investigator, laurene.thompson@carleton.ca, 613-520-2600, ext. 3695) or Dr. Joanna Pozzulo (Faculty Advisor, joanna.pozzulo@carleton.ca, 613-520-2600, ext. 1412).

What if I have concerns?

Concerns: If you should have any ethical concerns about this study please contact, Dr. Andy Adler (Chair, Carleton University Research Ethics Board-B, adler@sce.carleton.ca, 613-520-2600, ext. 8785). You can also contact the Carleton University Research Compliance Office at ethics@carleton.ca for any other concerns.

This research was cleared by Carleton University Research Ethics Board-B (105782).

At this time we would like to thank you for taking the time to take part in this study. Your participation has been greatly appreciated!

Appendix N, Subsection i

Method (Material Testing Study)

Participants

Undergraduate psychology students ($N = 36$) were recruited from Carleton University via an online participation pool (i.e., SONA; Appendix N, Subsection iii). Participants' age ranged from 18 to 31 ($M = 20.17$, $SD = 3.25$). There were 23 (63.9%) males and 13 (36.1%) females. The majority of the sample identified themselves as White/Caucasian (36.0%), with small numbers of Asian (27.8 %), Black (27.8%), Aboriginal (5.6%), and Mixed (2.8%) ethnic backgrounds.

Materials

Informed Consent form. Participants were given a consent form explaining that they will be asked to complete a series of questions on perceived stereotype consistencies (Appendix N, Subsection iv).

Stereotype Ratings Questionnaire. Participants were asked to fill out a questionnaire indicating the extent to which a profession is consistent with a male or female gender role. This was rated on a 7-point scale (1 = *male*, 7 = *female*; Appendix N, Subsection v).

Debriefing form. Participants were asked to read a debriefing form upon the completion of the study. The debriefing form outlined the purpose of the study, the research question and hypotheses, and the researcher's contact information (Appendix N, Subsection vi).

Procedure

Participants scheduled an appointment with the primary researcher to attend the Laboratory for Child Forensic Psychology to complete the study. Upon arrival, participants were asked to complete an informed consent form (Appendix N, Subsection iv). They were then given

the stereotype questionnaire (Appendix N, Subsection v). Finally, participants were asked to read over a debriefing form (Appendix N, Subsection vi). At this time, they were able to ask any questions they may have regarding the study and were thanked for their time.

Appendix N, Subsection ii**Results (Material Testing Study)**

The results from the stereotype questionnaire were used to determine what occupations would be chosen for the stereotypical male occupation and the stereotypical female occupation for the main study. A one-way repeated measures ANOVA was conducted to determine whether there were any differences in mean stereotype ratings across the various occupations. There was a significant main effect for the stereotype ratings, $F(9, 315) = 113.42, p < .001$. Pairwise comparisons were conducted using a Bonferroni correction for multiple comparisons. Paired comparisons revealed that several of the occupations were rated significantly different from one another (see Table 5 for mean differences). Mechanic and construction worker were rated the most masculine ($M = 2.44$; $M = 2.22$, respectively) while nurse and daycare worker were rated the most feminine ($M = 5.64$; $M = 6.22$, respectively). Based on these results, the construction worker was chosen to be the stereotypical male job and the nurse was chosen to be the stereotypical female job for the main study. Nurse was chosen over daycare worker as the job title “daycare worker” seemed too generic. Results indicated that the nurse was rated significantly more feminine than the construction worker (mean difference = $\pm 3.42, p < .001$). Thus, the main study moved forward using the nurse and construction worker profiles to represent the stereotype profiles.

Table 5

Mean Difference between Stereotype Ratings for Various Occupations

	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Mechanic	-2.94*	0.22	-2.58*	-0.72	-2.19*	-0.06	-0.58	-3.19*	-3.78*
2. Librarian	—	3.17*	0.36	2.22*	0.75	2.89*	2.36*	-0.25	-0.83
3. Construction Worker	—	—	-2.81*	-0.94	-2.42*	-0.28	-0.81	-3.42*	-4.00*
4. Hair Stylist	—	—	—	1.86*	0.39	2.53*	2.00*	-0.61	-1.19*
5. CEO	—	—	—	—	-1.47*	0.67	0.14	-2.47*	-3.06*
6. Teacher	—	—	—	—	—	2.14*	1.61*	-1.00*	-1.58*
7. Computer Technician	—	—	—	—	—	—	-0.53	-3.14*	-3.72*
8. Engineer	—	—	—	—	—	—	—	-2.61*	-3.19*
9. Nurse	—	—	—	—	—	—	—	—	-0.58
10. Daycare Worker	—	—	—	—	—	—	—	—	—

Note. * $p < .001$

Appendix N, Subsection iii**Sona Recruitment Form (Material Testing Study)**

Study Name: Do They Fit the Stereotype?

Description: Participation in this study will require that you watch short videos and answer questions about those videos.

Eligibility Requirements: Must be 18-years-old or older. This is a visual activity. If you have any vision problems that interfere with your ability to view the video, you are ineligible to participate in the current study.

Duration: 30 minutes

Location: Social Sciences Research Building, Room 111

Compensation: 0.5% in PSYC 1001, PSYC 1002, PSYC 2001, PSYC 2002, NEUR 2001, or NEUR2002

Researchers

Primary Investigator: Lauren Thompson, Psychology Department, Graduate Student

Phone: 613-520-2600, ext. 3695

Email: LaurenE.Thompson@carleton.ca

Faculty Advisor: Dr. Joanna Pozzulo

This study has received clearance by the Carleton University Research Ethics Board-B (105782)

Appendix N, Subsection iv**Informed Consent Form (Material Testing Study)**

The purpose of informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. Informed consent must provide sufficient information such that you have the opportunity to determine whether or not you wish to participate in the study.

Present study: Do They Fit the Stereotype?

Research personnel: The following people will be involved in this research project and may be contacted at any time: Lauren Thompson (Principal Investigator, laurene.thompson@carleton.ca, 613-520-2600, ext. 3695) or Dr. Joanna Pozzulo (Faculty Advisor, joanna.pozzulo@carleton.ca, 613-520-2600, ext. 1412).

Concerns: If you should have any ethical concerns about this study please contact, Dr. Andy Adler (Chair, Carleton University Research Ethics Board-B, adler@sce.carleton.ca, 613-520-2600, ext. 8785). You can also contact the Carleton University Research Compliance Office at ethics@carleton.ca for any other concerns.

Purpose: The purpose of this study is to examine the perceived consistency of several stereotypes.

Task requirements: You will be asked to watch videos and then complete questionnaires regarding the viewed videos. You must be 18-years-old or older. This is a visual activity. If you have any vision problems that prohibit you from viewing a video, you are ineligible to participate in the current study.

Duration and locale: Testing will take place in Room 111, Social Sciences Research Building, Carleton University. This study will be completed in one testing session, lasting approximately 30 minutes.

Token for participation: You will receive a 0.5% increase in your final grade of PSYC 1001, PSYC 1002, PSYC 2001, PSYC 2002, NEUR 2001, or NEUR2002 for participating in this study.

Potential risk/discomfort: There are no potential risks involved in this experiment. Should you experience any unease, you have the right to withdraw from the study and still receive course credit.

Anonymity/Confidentiality: All the information you provide will be strictly confidential. Data will only be used for research at Carleton University. Your answers will NOT be linked to your name or signature (i.e., consent form) and your responses will be coded in such a way that you cannot be identified.

Protection of Personal Information: This Informed Consent Form will be kept in the Lab for seven years. It will be placed in a room that has restricted access and is kept locked and closed at all times. At the seven-year mark, it will be shredded and disposed of.

Right to withdraw: Your participation is strictly voluntary. At any point during the study you have the right not to complete certain questions or to withdraw from the study without any penalty whatsoever by informing the researcher. However, once the study session has finished, participants cannot withdraw because the data is not linked to any names.

This study has received clearance by the Carleton University Research Ethics Board-B (105782)

Signatures: *I have read the above form and hereby consent to participate in this study. The data in this study will be used for research publications and/or teaching purposes. I am aware that the data collected in this study will be kept strictly confidential and anonymous. My signature indicates that I understand the above and wish to participate in this study.*

Participant's Name (print): _____

Participant's Signature: _____

Researcher's Name (print): _____

Researcher's Signature: _____

Date: _____

Appendix N, Subsection v

Stereotype Ratings Questionnaire (Material Testing Study)

Read each question below. Using the scale provided, circle the answer that appropriately represents your view of each job occupation.

How would you rate a:

Mechanic?

1	2	3	4	5	6	7
Male						Female

Librarian?

1	2	3	4	5	6	7
Male						Female

Construction Worker?

1	2	3	4	5	6	7
Male						Female

Hair Stylist?

1	2	3	4	5	6	7
Male						Female

CEO?

1	2	3	4	5	6	7
Male						Female

Teacher?

1	2	3	4	5	6	7
Male						Female

Computer Technician?

1	2	3	4	5	6	7
Male						Female

Engineer?

1	2	3	4	5	6	7
Male						Female

Nurse?

1	2	3	4	5	6	7
Male						Female

Daycare Worker?

1	2	3	4	5	6	7
Male						Female

Appendix N, Subsection vi**Debriefing Form (Material Testing Study)*****What are we trying to learn in this research?***

The purpose of the present study is to examine the extent to which a profession is consistent with a male or female gender role.

Why is this important to scientists or to the general public?

This research is important because it helps better understand how factors such as gender and occupation play a role in stereotype consistency.

What are our hypotheses and predictions?

We predict that individuals will be more likely to rate professions that are often dominated by males (e.g. construction workers) as more stereotypically male whereas professions dominated by females (e.g. nurse) will be more likely to be rated as stereotypically female.

What if I feel distress or anxiety after participating in this study?

If you feel any distress or anxiety after participating in this study, please feel free to contact the Carleton University Health and Counseling Services at: 613-520-6674, or the Distress Centre of Ottawa and Region at 613-238-3311 (<http://www.dcottawa.on.ca>).

What if I have questions later?

If you wish to discuss this research any further feel free to contact any one of the following people: Lauren Thompson (Psychology Department, Graduate Student, Principal Investigator, LaurenE.Thompson@carleton.ca, 613-520-2600, ext. 3695) or Dr. Joanna Pozzulo (Faculty Advisor, Joanna.Pozzulo@carleton.ca, 613-520-2600, ext. 1412).

What if I have concerns?

Concerns: If you should have any ethical concerns about this study please contact, Dr. Andy Adler (Chair, Carleton University Research Ethics Board-B, adler@sce.carleton.ca, 613-520-2600, ext. 8785). You can also contact the Carleton University Research Compliance Office at ethics@carleton.ca for any other concerns.

This research was cleared by Carleton University Research Ethics Board-B (105782).

At this time we would like to thank you for taking the time to take part in this study. Your participation has been greatly appreciated!