Exploring the Influence of Gender and Post-Event Information on Memory for Suspects

by

Erin Rennie

B.A., The University of British Columbia, 1999

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Psychology, Cognitive Systems Area)

We accept this thesis as conforming to the required standard.

THE UNIVERSITY OF BRITISH COLUMBIA

July 2002

© Erin Rennie, 2002
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Psychology
The University of British Columbia
Vancouver, Canada

Date July 24, 2002
Abstract

The goal of this thesis was to investigate factors that influence eyewitnesses' memories for suspects. To investigate memory for suspects, I conducted two experiments. In Chapter 2, I report a study that focused on the influence of subject and suspect gender on witnesses' perceptions of and memories for suspects. In Chapter 3, I report a study that focused on the influence of post-event descriptions of suspects on witnesses' memories for suspects. Together, these studies revealed that women are better than men at remembering male and female suspects. In addition, women are especially good at remembering suspect's clothing. These findings are a first step towards understanding eyewitness memory for suspects.
Table of Contents

Abstract ............................................................................................................. ii
Table of Contents .......................................................................................... iii-iv
List of Tables ................................................................................................... v
List of Figures ................................................................................................... vi

CHAPTER 1: GENERAL INTRODUCTION .......................................................... 1-2

CHAPTER 2: GENDER INFLUENCES EYEWITNESS PERCEPTIONS OF AND MEMORY FOR SUSPECTS ........................................... 3-23
   Introduction .................................................................................................. 3-10
   Method ........................................................................................................ 10-15
   Scoring ....................................................................................................... 15-16
   Data Analyses ............................................................................................ 16-18
   Results ....................................................................................................... 18-20
   Discussion ................................................................................................. 21-23

CHAPTER 3: EXPLORING THE INFLUENCE OF POST-EVENT INFORMATION ON MEMORY FOR SUSPECTS ........................................... 24-50
   Introduction .................................................................................................. 24-30
   Phase 1: Likelihood Ratings ....................................................................... 30-34
      Method ..................................................................................................... 30-33
      Results & Discussion ............................................................................. 33-34
   Phase 2: The Influence of Post-event Information on Memory for Suspects
      Introduction .................................................................................................. 35
      Method ..................................................................................................... 35-41
      Results ....................................................................................................... 41-42
      Discussion ................................................................................................. 43-44

CHAPTER 4: CONCLUSION ............................................................................. 45

References ........................................................................................................ 46-49
Appendixes ....................................................................................................... 50-89
   Appendix A: Chapter 2 Suspect Photographs ............................................ 50-51
   Appendix B: Scoring Key ........................................................................... 52-65
   Appendix C: Chapter 3 Suspect Photographs ............................................ 66
   Appendix D: Likelihood Questionnaires ..................................................... 67-71
   Appendix E: Likelihood Questionnaire Data .............................................. 72-81
   Appendix F: Neutral Information for the Written Descriptions ................. 82-84
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix G</td>
<td>Written Post-event Descriptions</td>
<td>85-87</td>
</tr>
<tr>
<td>Appendix H</td>
<td>Memory Tests for Suspects A and B</td>
<td>88-97</td>
</tr>
<tr>
<td>Appendix I</td>
<td>Number of Correct Responses for Each Memory Test Item Across Suspect Photographs</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Tables</td>
<td>99-104</td>
</tr>
<tr>
<td></td>
<td>Figure Captions</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Figures</td>
<td>106-114</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table #</th>
<th>Title</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experiment Design for Chapter 2</td>
<td>99</td>
</tr>
<tr>
<td>2</td>
<td>Assignment of Subjects and Counterbalancing of materials</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>Order of tasks</td>
<td>101</td>
</tr>
<tr>
<td>4</td>
<td>Summary of Floor and Ceiling Effects for Study Phase and Test Phase Descriptions</td>
<td>102</td>
</tr>
<tr>
<td>5</td>
<td>Counterbalancing of Information Types Across Groups of Items</td>
<td>103</td>
</tr>
<tr>
<td>6</td>
<td>Subject Assignment and Presentation Sequence of Photographs and Written Descriptions</td>
<td>104</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure #</th>
<th>Title</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Subjects’ Study Phase Descriptions of Men and Women</td>
<td>106</td>
</tr>
<tr>
<td>2</td>
<td>Subjects’ Test Phase Descriptions of Men and Women</td>
<td>107</td>
</tr>
<tr>
<td>3</td>
<td>Subjects’ Study Phase Descriptions of Female Suspects</td>
<td>108</td>
</tr>
<tr>
<td>4</td>
<td>Subjects’ Study Phase Descriptions of Male Suspects</td>
<td>109</td>
</tr>
<tr>
<td>5</td>
<td>Subjects’ Test Phase Descriptions of Female Suspects</td>
<td>110</td>
</tr>
<tr>
<td>6</td>
<td>Subjects’ Test Phase Descriptions of Male Suspects</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>Mean Losses Between Study &amp; Test</td>
<td>112</td>
</tr>
<tr>
<td>8</td>
<td>Influence of Describing Out Loud and Silent Study on</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Recollections of Suspects</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Likelihood Rating Scale</td>
<td>114</td>
</tr>
</tbody>
</table>
CHAPTER 1
GENERAL INTRODUCTION

My goal in this thesis was to investigate factors that influence eyewitnesses’ memories for people such as suspects. Many factors are known to influence eyewitness memory: emotions, drugs, the presence of a weapon, misleading information, other witnesses, gender, etc. For example, a witness who is emotionally traumatized may have difficulty remembering the suspect or surroundings of the crime. When any of these factors influence a witness, there is a risk that the witness may remember the suspect incorrectly or incompletely. When a witness provides inconsistent or incomplete testimony they may mislead the police into apprehending the wrong suspect or they may mislead a jury into reaching the wrong verdict. Convicting an innocent person or acquitting someone who is guilty has serious consequences for the suspect, the victims, the justice system and society.

In this thesis, I examine only some of these factors. Throughout the following chapters, I will use the word “suspects” to refer to the people that subjects view and remember. Chapter 2 focuses on the influence of witness and suspect gender on eyewitness perceptions of and memory for suspects. This chapter begins with a brief review of the relevant literature on this topic. This review is followed by a description of an experiment that investigates witnesses’ perceptions of and memory for suspect attributes.

Chapter 3 focuses on how post-event information about a suspect’s physical appearance can influence a witness’s memory for the suspect. After witnessing a crime or accident, a witness will be faced with information from many different sources regarding the suspect they have seen. This information may aid or distort their memory for the suspects and thus influence their testimony in court. In this chapter, I begin by
discussing literature that is relevant to this issue. This discussion is followed by a
description of an experiment and a discussion of the results of this experiment.
CHAPTER 2
THE INFLUENCE OF GENDER ON EYEWITNESSES' PERCEPTIONS OF AND MEMORIES FOR SUSPECTS

Introduction

The goal of this study was to investigate the influence of gender on witnesses' perceptions of and memories for suspects. The main question of interest was whether men and women perceive and remember the same physical attributes of men and women equally often (e.g., hair, clothes, height). If men and women notice and remember different details about members of their own gender versus the opposite gender, perhaps they differ in their abilities to supply testimony about male and female suspects. Police may be able to use information about witnesses' strengths and weaknesses to select those witnesses who are most likely to provide consistent, complete and accurate testimony about a particular suspect. For example, if women remember more details about other women, then perhaps police could rely more on female witnesses when a description of a female suspect is needed.

To understand whether women or men observe or remember the physical attributes of men and women differently, it is necessary to focus on perception and retrieval. For my purposes in this thesis, perception refers to those suspect attributes that witnesses observe (i.e., notice) while they are actually viewing the suspect. Retrieval refers to suspect attributes that witnesses remember (e.g., recall or recognize) when the suspect is no longer in view. By focusing on perception and recall, I will be able to determine exactly which specific physical attributes (e.g., hair color, weight, face) men and women notice and which of these attributes are remembered later on. Examining both perception and recall may also shed light on whether differences in memory can be explained by differences that occur during encoding or retrieval. In the
following sections, I will review the literature on gender differences in perception of suspects and then I will review the literature on gender differences in memory for suspects.

*Perception of Suspects*

It is sometimes assumed that women are better than men at noticing things about other people, however, no studies have investigated whether women or men observe different details about the same sex versus the opposite sex. Only one study has shed light on the accuracy of subjects' perceptions of others (Ebbeson & Rienick, 1998).

Ebbeson and Rienick (1988) conducted the only study that has investigated the accuracy of witnesses’ perceptions of suspects. This extensive study explored perception and memory. I will discuss only the part that dealt with perception of suspect attributes. Subjects were instructed to provide a written description of a suspect's gender, age, height, ethnicity, hair color, hairstyle, clothing top color and style while the person was standing and facing them. These self generated descriptions offered a way to measure which suspect attributes were perceived by subjects.

The results indicated that subjects described several attributes incorrectly in their written descriptions (Ebbeson & Rienick, 1998). For example, they may have described a suspect as having wavy hair, when in fact, the suspect had straight hair. Twenty three percent of the thirty subjects made description errors for facial features. Seventeen percent made errors for clothing style and seven percent were mistaken about ethnicity.

This study provides insight into the accuracy of subjects’ perceptions when they were asked to describe specific suspect attributes (e.g., hair color, clothing top color), but it tells us nothing about gender differences in perception of these attributes. To
examine whether men and women differ in their perceptions, both men and women subjects and suspects must be included in a study.

Because subjects in this study relied on cues from the experimenter, I do not know whether these cues biased subjects or whether subjects would be equally accurate if they were required to provide uncued (i.e., free, spontaneous) descriptions of suspects. Uncued descriptions could reveal the details that subjects perceived without any biasing influence from cues supplied by the experimenter and they would also be more realistic. In a crime situation, a witness would not be told which aspects of a suspect to look at as the crime was in progress and there is a possibility that witnesses will show more variation in what details they observe when they are not given any cues.

To truly understand whether men and women differ in the details that they notice about others, a study would require subjects and suspects of both genders. In addition, such a study should elicit uncued descriptions from subjects while the suspects are in view. The study that I report in this chapter meets both of these requirements.

**Remembering Suspects**

A search of the literature on memory for suspects reveals that many researchers have asked whether men and women are equally good at remembering members of the same sex versus the opposite sex. Most of this literature comes from face recognition studies. However, few studies have asked whether men and women differ in their recollections of physical attributes of men and women (e.g., hair color, clothing type, height).

The ability to recognize a suspect's face is important to a police investigation, however, the ability to provide a detailed description is even more critical. Police may need a very detailed description of a suspect to help them apprehend the individual or to create a sketch of them for public circulation. In these cases, a witness must be able
to provide a description of the suspect long before they are asked to recognize the individual. Despite the fact that police and the courts require witnesses who can provide descriptions of suspects, the literature reveals little about the influence of witness gender and suspect gender on witnesses' recollections of suspects' physical attributes (e.g., hair, eyes, clothes).

Because only one study has addressed memory for suspect attributes, I will review this study together with face recognition studies (Powers, Loftus & Andriks, 1979). The purpose of the following review is to highlight the major results of face recognition studies that have included subjects and suspects of both genders. I will also discuss explanations for these results. Finally, I will discuss whether or not these results and explanations can provide any clues to suggest whether men and women may differ in their memories for attributes of male and female suspects.

In face recognition studies subjects are typically asked to study photos of faces for a few seconds each. At some time after this study phase, subjects are given a recognition test. Face recognition tests usually require subjects to indicate orally, or by pressing a key, whether a face is old (i.e., seen at study) or new.

McKelvie's (1981) review of face recognition studies as well as more recent work in this area have revealed that interactions between sex of subject and sex of face are sometimes found. When these interactions are broken down, the most frequently reported result is a same sex effect for women. A same sex effect suggests that women are very good at recognizing women's faces (McKelvie, 1981; McKelvie, Standing, Lionel, St. Jean & Law, 1993). This effect has been found in numerous studies and seems fairly well established (Ellis, 1975; McKelvie, 1978, 1981). A same sex effect for men is less common (Caroo & Mozingo, 1989; McKelvie et al, 1993, Experiment 1; Powers. Loftus & Andriks, 1979). Same sex effects for women and for men are
sometimes found in the same study (McKelvie et al., 1993, Experiment 1; Powers, Loftus & Andriks, 1979). However, other studies have found that men are equally good at recognizing men’s and women’s faces while women show a same sex effect (Cross, Cross & Daly, 1971; Deffenbacher, 1978; Going & Read, 1974). Sometimes these interactions are accompanied by a main effect in which all subjects recognize significantly more female than male faces (McKelvie et al., 1993).

Other studies have not found a significant interaction, though they have found a main effect of subject gender. One of these studies found that men and women recognized more male than female faces, while another study found that men and women recognized more female than male faces (Borges & Vaughn, 1977; Brigham & Barkowitz, 1975).

Many explanations have been offered to account for these results, however, differences in interests and in experience with faces are the most common. Researchers have proposed that men and women are interested in different things and that these interests may explain why they differ in face recognition abilities (e.g., McKelvie, 1981). For example, women may be taught by society to be interested in and attend to stereotypically female things such as clothing, cosmetics or women’s faces which may partly explain the same sex effect for women (McKelvie, 1981; McKelvie et al., 1993; Cross, Cross & Daly, 1971). Along the same reasoning, men may become interested in stereotypically masculine objects such as cars, which may explain why the same sex effect for men is not as common (Davies & Robertson, 1993).

More often, researchers have proposed that different interests and different amounts of experience with faces both influence the results discussed previously (Davies & Robertson, 1993; McKelvie, 1981; McKelvie et al., 1993; Powers, Loftus & Andriks, 1979). According to the expertise account, humans are experts at recognizing
faces because they have had exposure to faces and they have experience at making fine discriminations between them (Diamond & Carey, 1986). The expertise account fits together neatly with the interest account described earlier to explain several results. For example, the same sex effect for women can be explained by women's increased interest in and experience with discriminating faces. To support this idea, Cross et al. (1971) suggested that women may be interested in female faces, especially attractive facial features, through exposure to cosmetic advertisements. Cross et al. (1971) also suggested that increased exposure to faces may lead to better discrimination between faces, resulting in improved recognition performance. Using the same reasoning, perhaps the same sex effect for men is less common than the same sex effect for women because men are not as interested in faces and because they have less experience at discriminating amongst faces.

The interest and expertise accounts can explain why individual results occur such as the same sex effect for women and the finding that men are equally good at recognizing faces of both genders, but they cannot explain why mixed results occur across studies. For example, some studies have shown a same sex effect for men, while others have found that men recognize male and female faces equally well (McKelvie et al, 1993; cf. Cross et al., 1971).

It has been suggested that these mixed findings could be connected to differences in face recognition difficulty (McKelvie, 1981). McKelvie (1981, p. 111) did not explain what he meant by "difficulty", however I will assume that he was referring to differences in the similarity between faces. Deffenbacher (1978) noted that very similar faces have been found to lead to poorer recognition performance. In support of Deffenbacher (1978), Brigham and Barkowitz (1975) found that all of their subjects recognized more male faces than female faces and they attributed this result to the fact
that the Caucasian women in their study all had very similar long hair making it difficult for subjects to distinguish between them (Brigham & Barkowitz, 1975).

Based on the face recognition literature, it is not possible to make exact predictions about how men and women may differ in their recollections of physical attributes of male and female suspects. It seems reasonable that both interest and experience play a role in face recognition and in memory for suspect attributes, however, this does not necessarily mean that the face recognition results will hold up when subjects are asked to describe a suspect in detail. Men and women might pay more attention to physical attributes of the opposite sex than to their own sex. This would suggest that same sex effects found in face recognition should not occur at all or that they should be weaker in memory for suspect attributes. Although the face recognition literature has provided clues as to what results could occur when investigating memory for suspect attributes, these clues are speculative.

My primary objective in this study was to investigate whether men and women perceive and freely recall suspect attributes (e.g., hair, eye color, clothes color) of male and female suspects equally often. To assess whether subjects perceived suspect attributes equally often, they were asked to describe photographs of men and women out loud (i.e., study phase descriptions) while viewing these photos. To assess whether subjects recalled attributes equally often, they were asked to describe the photographs out loud from memory (i.e., test phase descriptions). Investigating perceptions (i.e., study phase descriptions) and recollections also provided a way of exploring whether gender differences are due to what subjects perceive or to what they remember. Based on my own observations and the evidence in support of gender differences in face recognition, I expected that men and women would differ in the frequency with which
they described various suspect attributes (e.g., hair, clothing, facial features) in their perceptions and in their recollections.

My secondary objective in this study was to investigate whether different methods of studying photographs of suspects would influence later recall. Subjects were randomly assigned to describe photographs of men and women out loud during study or to silently study the photos. This manipulation was intended to determine whether anticipated differences at recall would be found only when subjects described photographs out loud or whether differences would hold up across a silent study condition.

Method

Subjects and Design

Subjects were 66 (18 men, 48 women) undergraduates recruited from first and second year psychology courses at the University of British Columbia. Subjects participated voluntarily in exchange for course credits.

This experiment was part of a large study and I will discuss only those factors that are relevant to my thesis. The two between-subjects factors were as follows: subject gender (men and women) and study task (describe out loud vs. silent study).

For the describe out loud (DL) condition, subjects described photographed suspects out loud while viewing their photos. For the silent study (SST) condition subjects silently studied photographed suspects. All subjects viewed photographs of men and photographs of women. The design is illustrated in Table 1. Photograph Gender and study task formed four main conditions. Each subject participated in two of these conditions.

Insert Table 1
Materials

Photographs. Ten 4” x 6” color photographs were used in this study. Two photographs were taken of each of five people: two Caucasian men and three Caucasian women (see Appendix A for samples). The photos of one of these women were used for a practice session. These individuals were recruited through a colleague. The mean age of the young men was 26 years, while the mean age of the women was 24.5. Each suspect was randomly assigned a gender neutral name (i.e., Casey, Kelly, Jamie, Ronnie, Pat) to help subjects during the study task and recall test. Gender neutral names were chosen to prevent gender bias in subjects’ recollections of the suspects.

The pair of photos for each suspect showed the individual (from head to toe) in different poses, specifically, in side profile and front view. The purpose of using two photos for each suspect was to provide subjects with more information about them and to prevent subjects’ perceptions and recollections from being biased by the specific pose and facial expressions of the suspects or by the perspective in which a particular photo was taken.

Picture Differences Task. This task consisted of a white sheet of 8.5” x 14” paper with a line drawn down the center. The form contained pairs of black and white line drawings of common objects (e.g., bathtub, candlestick). The pairs were split up so that half of the members were on one side of the black line, while the other members were on the other side of the line. Several pairs of objects differed with respect to one small detail. The task was used as a filler activity.

North American Adult Reading Test (NAART). This test measured verbal intelligence and consisted of 61 words (Spreen & Strauss, 1998). These words were presented on two white 8.5” x 11” sheets of paper. Each sheet contained two columns
of words. The NAART was intended to screen out subjects who were not proficient in English, however, the experimenter felt that all subjects were proficient enough to complete the tasks and deemed it unnecessary to use the NAART data.

A tape recorder was employed to record subjects' study phase and test phase descriptions as well as their NAART responses. A stopwatch was also used to time several tasks.

Counterbalancing of Materials

The four pairs of photographs were counterbalanced with a 4 x 4 Latin square to prevent order effects. Each row of the Latin square was crossed with the basic between-subject conditions (see Table 2). In Table 2, the photographs of the men are designated “A” and “B”, while the women are “C” and “D”. The numbers in the Latin squares indicate the order in which the sets of photos were presented to each subject (Table 2).

Procedure

Subjects were tested in one 1-hour session. The order in which tasks were administered to subjects is shown in Table 3.

Study Phase Instructions. All subjects were informed that they would be shown several sets of photographs, each showing a different suspect and that they would see only one suspect at a time.

Subjects in the DL condition were instructed to say the name of the photographed suspect out loud and then to describe the photographed suspect out loud to the best of their ability. Subjects were told to imagine that they were eyewitnesses to
a crime or accident and that the police were asking them to describe the suspect. The completeness and accuracy of each description was emphasized. Subjects were given three minutes to describe each suspect out loud. They were told that they would be given a memory test for the same suspects later in the experiment. They were told that they would engage in a practice session before the study phase began.

Subjects in the SST condition were given the same instructions as those in the DL condition, except they were asked to study each suspect silently for three minutes, rather than describe each one out loud.

**Study Phase Practice.** The study phase practice session immediately followed the instructions. Regardless of their assigned condition, all subjects described one Caucasian woman out loud for practice. Subjects were given three minutes to describe the suspect. During this practice session, the experimenter stressed the eyewitness scenario and the need for complete and accurate descriptions as needed to motivate subjects to provide descriptions which included different attributes (e.g., hair, face shape, height). The experimenter did not ask subjects to describe specific attributes of the suspects (e.g., age, height, body shape, etc.). If the subjects indicated that they were finished their description before the three minutes were up, the experimenter encouraged them to report everything possible and told them to keep trying to describe the person until the three minutes were up.

**Study Phase.** This task immediately followed the practice. Subjects were told that they would see four different suspects, presented one at a time. Subjects in the DL condition were reminded to say each suspect’s name out loud and then to describe each one out loud. Subjects in the SST condition were reminded that they must say each suspect’s name out loud and study them silently. All subjects were reminded of the three minute time limit to describe/study each suspect and that they would be given a
memory test later in the experiment. Following these brief instructions, each set of photographs was placed in front of the subject, one set at a time, while subjects described or silently studied the photos. As the photos were placed in front of the subject, the experimenter started the stopwatch. After three minutes had elapsed, the experimenter removed the photos from the table and reminded subjects to make their descriptions as complete and as accurate as possible before placing the next set of photos in front of the subject. The study phase lasted for about 13-15 minutes.

Filled Delay Tasks. The North American Adult Reading Test (NAART) was administered immediately following the Study Phase (see Table 3). The NAART was administered according to published guidelines (Spreen & Strauss, 1998). Subjects’ responses were tape recorded so that the responses could be scored at a later time. The NAART took less than five minutes to complete.

The Picture Differences Task was administered to all subjects immediately following the NAART. For this test, subjects were instructed that they would be given 1.5 minutes to find all of the pairs of line drawings that differed. They were asked to write the order in which they found each different pair beside the drawings.

Test Phase. Immediately following the Picture Differences Task, all subjects were given a free recall test for the photographed suspects that they had described out loud or silently studied earlier in the experiment. Subjects were instructed to describe the same suspects that they had described/studied earlier in whatever order they wished. The experimenter again instructed subjects to imagine that they were eyewitnesses and reminded them to be as complete and accurate as possible. After subjects described a suspect, the experimenter asked if they could remember anything else at all. No time limit was imposed on the length of each description. The experiment
ended when subjects indicated that they had described all of the suspects they could remember.

Scoring

Subjects' study phase and test phase descriptions of the photographed suspects were transcribed and then individually scored for 27 different suspect attributes (e.g., hair descriptors, clothing color, body-build) according to a scoring key (Appendix B). The scoring key contained detailed instructions on how to categorize words and phrases into each of the 27 suspect attribute categories. The purpose of the scoring key was to determine the number of times that each attribute occurred in each suspect description. For each attribute, the number of occurrences was counted and then recorded for each study phase description and each test phase description. For example, if a subject mentioned three different ideas about a person's hair (e.g., color, length), then the frequency of hair descriptors would be three (Appendix B). The sample below describes how to score the suspect attribute called hair descriptors.

Hair Descriptors. A count of the number of words and phrases used for describing hair. For this count, include reference to hair length, hair loss, hairstyle or hairdo (e.g., bangs, ponytail, buzz cut) and hair color. Also include references to where the hair is and/or how it hangs or is parted etc.

DO NOT include references to facial hair or lack of facial hair.

DO NOT count references to hair clips or pins. These will be counted in Jewelry/Accessories.

Examples

"He has blonde hair." "He is balding" 2 points

"He has black hair, balding at the front..." 3 points
"She has long, dark, straight hair, hanging about three inches below her shoulder." 4 points (Appendix B).

In addition, the number matches between each subject’s study and test descriptions were scored for each attribute. Matches were defined as the number of ideas that matched between the study phase descriptions and the test phase descriptions (see Appendix B).

Pearson's $r$ was used to calculate inter-rater reliability correlations between two raters for each of the 27 suspect attributes. Both raters scored 24 study phase descriptions collected from six different subjects for all 27 attributes. They also scored 44 test phase descriptions collected from 11 different subjects for all attributes. For the study phase descriptions and the test phase descriptions, inter-rater reliability was at least .70 for all attributes that did not have floor or ceiling effects. Most attributes were above .80.

The raters scored the matches between study and test descriptions for all 27 attributes based on 24 protocols from a total of six subjects. Only the matches for attributes that achieved an $r$ of at least .70 will be reported in this paper.

**Data Analyses**

**Study Phase & Test Phase Analyses.** There were a total of 136 spoken study phase descriptions and 264 test phase descriptions collected from 66 subjects. The study phase descriptions were analyzed to determine whether men or women described the same type of suspect attribute and whether they described each attribute equally often. The test phase descriptions were also analyzed to determine the type and quantity of each suspect attribute described by men and by women. These data analyses focused on the absolute mean frequency (number of times) that each subject mentioned a particular attribute in the study phase and in the test phase descriptions.
Data analyses focused on the frequency of suspect attributes rather than on the proportion of subjects who mentioned each attribute because this allowed me to determine whether male and female witnesses described the same attributes of males and females equally often at study and at test. Ninety-five percent confidence intervals were calculated for each attribute.

Subjects’ study phase and test phase descriptions were examined to determine the average amount of forgetting or losses between study and test for each suspect attribute. Losses were determined by using the scoring key to code the matches between study and test and then subtracting these matches from the study phase descriptions.

An error in subjects’ study phase and test phase descriptions was defined as a word or phrase that described an attribute incorrectly. For example, an error was scored and recorded when a subject described a shirt as red when it was actually black. Because errors were extremely rare, only accurate data are reported in this paper.

**Floor & Ceiling Effects.** Several suspect attributes had floor or ceiling effects. A floor effect was defined when 20% of the study or test descriptions contained a score of 0 for a particular attribute. Table 4 indicates that eight attributes showed floor effects for the study phase descriptions. Nine attributes showed floor effects for the test phase descriptions.

A ceiling effect was defined when at least 20% of the study phase descriptions or test phase descriptions contained the maximum attainable score for a particular attribute. Six suspect attributes (hair, age, eyes/eyebrows, gender, ethnicity, name) had a limit on the number of occurrences that could be counted as defined in the scoring key (see Appendix B). For example, “Hair” had a maximum attainable score of 1 for each individual study or test description. Thus, “Hair” reached ceiling because at least 20% of
the study phase descriptions and the test phase descriptions had a score of 1. Table 4 indicates that all six of these attributes showed ceiling effects in the study phase. Five of these attributes showed ceiling effects at test (Table 4). The rest of the 27 attributes had no limit on the number of occurrences that could be counted and scored, thus they could never have a ceiling effect.

Insert Table 4

Results

The main objective of this study was to determine whether men and women were equal in their perceptions of physical attributes of men and women. Overall, the results indicated that men and women subjects differed in their perceptions and recollections of men and women. Alpha was set at .05 for all analyses reported in this paper. Figure 1 illustrates a clear trend for women subjects to perceive suspect attributes more frequently than men. Women perceived suspect attributes of male and female suspects equally often (Fig. 1). None of these trends reached significance, $F(1,32) = 1.90, p = .178, \eta^2 = .05$ and $(1-\beta) = .267$. Figure 2 shows that women $(M = 33, SE = 1.72)$ recalled suspect attributes more frequently than men did $(M = 25, SE = 1.36)$ for male and for female suspects. This main effect of subject gender was significant according to the 95% confidence intervals and according to an ANOVA, $F(1,64) = 7.865, p = .007, \eta^2 = .109$ and $(1-\beta) = .789$.

Insert Figure 1

Insert Figure 2

In the following sections, I will describe the specific ways in which men and women subjects differed in their perceptions and recollections of men and women. This will reveal what specific suspect attributes men and women tend to perceive and which of these attributes are recalled most often at test.
Perception Results

Figure 3 shows that women subjects described 10 out of 12 suspect attributes more frequently than men did when they described female suspects. Although no suspect attributes reached significance according to the confidence intervals, Figure 3 clearly shows a trend for women to perceive attributes more frequently than men. The two exceptions to this trend were eyes/eyebrows descriptors and facial parts descriptors which men tended to describe more often than women.

Insert Figure 3

When describing photographs of men, women subjects described 11 out of 13 attributes more frequently than men did, though only two of these attributes reached significance (Fig. 4). Specifically, the 95% confidence intervals in Figure 4 show that women described clothing color and surroundings descriptors significantly more often than men. In addition, men described facial part descriptors more frequently than women did, though this trend was not significant (Fig. 4).

Insert Figure 4

Recall Results

Figure 5 indicates that women subjects recalled all 12 attributes more frequently than men did, when they recalled female suspects. In addition, Figure 5 illustrates that six of these twelve attributes attained significance with 95% confidence intervals (hair descriptors, eyes/eyebrows, clothes, clothes non-color, surroundings descriptors).

Insert Figure 5

In their test phase descriptions of male suspects, women recalled details about 11 of the 12 attributes more frequently than men did. Figure 6 shows this result and the 95% confidence intervals indicate that women recalled significantly more details about two attributes called clothing descriptors non-color and surroundings descriptors. The
attribute called clothing descriptors non-color focused on words and phrases that described the type of clothing such as type of material, brand name, the presence of stripes or buttons. The attribute called surroundings descriptors focused on words and phrases used to describe items such as stairs, backpacks or bags that make up the photographed suspect's surroundings.

Insert Figure 6

**Losses and gains between study and test**

To help determine whether the female superiority observed in recall of men and women suspects is due to gender differences in forgetting between study and test, I examined the gains and losses between study and test. Gains referred to the number of details that subjects included in their test descriptions but that were not mentioned at study. Gains were rare and no significant gender differences were present. Losses referred to the difference between the number of attributes that were described at study and the number of attributes described at both study and test. The 95% confidence intervals in Figure 7 indicated that overall, men and women do not differ in the number of details that they lose between study and test.

Insert Figure 7

**Study Task Manipulation**

To determine whether the study manipulation influenced subjects' test phase descriptions, I compared the test phase descriptions generated in the describe out loud condition with the test phase descriptions from the silent study condition. Figure 8 shows the means and 95% confidence intervals for the test phase descriptions. The effect of the study task was not significant.

Insert Figure 8
Discussion

This study investigated whether men and women would perceive and recall suspect attributes of men and women suspects equally often. The results indicated that men and women did not differ in their perceptions, however, they did differ in their recall of suspect attributes.

Women showed a trend to perceive more details than men about attributes of male and female suspects. Although this trend was not significant, it is unique. The fact that subjects in this study made virtually no errors in their study phase descriptions contrasts with Ebbeson and Rienick's (1998) subjects who described some suspect attributes incorrectly, despite the fact that the suspect was standing in front of them. This difference cannot be explained by time limits imposed on subjects since Ebbeson and Rienick's subjects had no time limit, while my subjects had a three minute limit.

The finding that women recalled significantly more details than men about attributes of male and female suspects does not support the same sex effect that frequently occurs in face recognition studies. However, this finding does agree with the less common finding in the face recognition literature, that women are better at recognizing all faces (McKelvie, 1981).

There are many possible explanations that may account for my results. These explanations include differences in men's and women's interests and experience with people, as well as differences in verbal abilities and episodic memory. I believe that my results can be attributed to a combination of these explanations. More research must be conducted to understand how these explanations may work together. For clarity, I will discuss these explanations separately.

According to the interest and the expertise accounts, women may show a trend towards better perception than men because men and women are interested in different
physical attributes and they differ in their experience at discriminating these attributes.

For example, the trend in this study for women to notice more details about stereotypically feminine attributes such as clothing and hair may stem from greater exposure to and interest in women's faces as well as clothing and hair styles.

Another possible explanation for the trend observed in perception is that men and women differ in verbal abilities. Recent studies have suggested that women outperform men on verbal tasks (Herlitz, Nilsson & Bäckman, 1997; Lewin, Wolgers & Herlitz, 2001). However, reasons for these differences are currently under debate in the gender differences literature (Kimura, 1999; Halpern, 2000). For example, Kimura (1999) found that women are better at verbal fluency tasks in which they are asked to generate as many words as possible that begin with a particular letter. If women do excel in verbal abilities, then this may have enabled the women in my study to describe more details than men during the study phase.

Based on intuition, it is also possible that men described attributes less often at study because they chose not to describe certain physical attributes. Despite instructing all subjects to describe each person as completely and accurately as possible, men may have felt that they shouldn't describe attributes that are stereotypically associated with women (e.g., hair, jewelry) because this would seem too feminine.

There are also many possible reasons why women recalled more details about suspect attributes compared with men. Based on the clear trend observed in perception, it is possible that women encoded more information about the suspects, and that this aided them at test. However, without significant results, I cannot be certain whether the recollections were influenced by perception or not.

Research in gender differences has suggested that women have better episodic memory than men. Recent studies have suggested that women outperform men in
episodic memory tests that require verbal processing or in tests where verbal responses are possible, though one recent study contradicted these findings (Herlitz et al., 1997; Lewin et al., 2001; cf. Ionescu, 2000). If women have better episodic memory than men, then they should forget fewer details between study and test. Examination of Figure 7 suggests that men and women do forget over time, but they do not differ in the amount of information that they forget. This suggests that either women do not have better episodic memory than men or that the retention interval was too short for significant forgetting to occur.

Herlitz et al. (1997) suggested that women may display better episodic memory than men because they have superior verbal abilities, not because they have better memory per se. Because a test of verbal abilities was not employed in my study, I cannot rule out the possibility that women may have recalled more details than men because they are better at verbalizing their memories.

I believe that the differences in recall were influenced by a combination of the explanations that I have discussed as well as additional factors yet to be explored. To determine how perception, verbal abilities, episodic memory or interests may work together to influence recall, more research is needed.
CHAPTER 3
EXPLORING THE INFLUENCE OF POST-EVENT INFORMATION ON MEMORY FOR SUSPECTS

Introduction

Witnesses receive information about suspects from many different sources: other witnesses, police, lawyers, the media, etc. (e.g., Loftus et al., 1979). The goal of this study was to investigate the influence of post-event information about a suspect's physical appearance on witnesses' memories of the suspect. Witnesses may receive factually accurate or inaccurate information about a suspect whom they witnessed. Factually accurate information can complement or reinforce a witness' memory, while inaccurate information may distort memory. For example, a witness may see a green car, but receive inaccurate information that the car was blue. When this happens, a witness' memory may be distorted and they may remember the car color as blue, or green or a greenish-blue (Loftus, 1977).

Most of what is known about the influence of post event information on memory comes from studies of the misinformation effect (e.g., Loftus, Miller & Burns, 1978). Misinformation, also called misleading information, is incorrect information about an event, object, person etc. The misinformation effect is the observation that subjects who are given misleading information after witnessing an event are more likely to choose this misleading information than the original accurate information on a forced choice recognition test. They are also more likely to recall the misleading information on a recall test. In contrast, subjects who are not given this misleading information will tend to choose the correct answer choice on a recognition test (Loftus & Hoffman, 1989).

In studies of the misinformation effect, subjects typically participate in three testing sessions, each held on a separate day. For the first session, subjects usually
Memory for Suspects

watch a video showing a mock crime or accident or they may see a live staged crime (e.g., Loftus, Miller & Burns, 1978). For the second session, subjects are given misleading information about one or more details of the event. This information may be delivered in written paragraphs, in question format, in the form of film slides or even composite drawings. For the third testing session, subjects are given a memory test for the original witnessed event. The test may be a forced choice recognition test, a free recall test, or it may require that the subject create a composite sketch of a suspect.

Several different kinds of misinformation have been shown to distort memory for an original event. Some studies provide subjects with misleading information that contradicts a detail that occurred in the original event. Alternatively, studies provide subjects with misleading information about details that did not exist in the original event. For example, Loftus (1975, Experiment 3) had subjects watch a film of a car driving down a road. Loftus suggested to subjects that a barn had been present in the film. Despite the fact that no barn had appeared, many subjects agreed with this misleading suggestion.

Research has suggested that subjects tend to accept subtle misinformation more frequently than blatant misinformation (Loftus, 1979). Loftus (1979) investigated whether subjects would accept misinformation that blatantly contradicted details of an original event as readily as they would accept more subtle contradictory misinformation. Subjects watched a film showing a man steal a wallet from a woman's bag, while the woman wasn't looking. Following a filler task, subjects were given a three choice recognition test for various details of the film. The following day, subjects read a written description containing misleading information about objects in the film. Some subjects received subtle misinformation, while others received subtle information plus a piece of information that blatantly contradicted an obvious object in the film. The subtle
misinformation concerned peripheral details in the film (e.g., a shirt in a store window, people passing the victim, an object carried by the victim’s friend, an overhanging sign). The blatant suggestion concerned a red wallet that was described as dark brown. Loftus considered dark brown to be blatantly false because subjects in an earlier experiment had all identified the wallet as red.

Results indicated that subjects who were accurate about an item on the initial accuracy test, before the misinformation was supplied, were more resistant to all misleading information about that item on the final recognition test. Subjects who received the blatant information about the wallet were more resistant to misinformation about all items on the final recognition test. In addition, an interaction appeared in which subjects who received blatant information and who were very accurate on the accuracy test were better able to resist misinformation than subjects who had not received blatant information. Loftus (1979) suggested that the blatant information may have made subjects more aware of all of the misinformation that they had read about and this may have helped them to pick the correct answer on the memory test.

This study had several problems that weaken Loftus’ results. One problem was that Loftus did not give examples of exactly what subtle information was provided about the four peripheral details in the film. Because I am interested in how misinformation influences memory for people, it would have been helpful to know whether Loftus gave subjects subtle misinformation about the physical appearance of any people in the film. Another potential problem concerned the timing of the first memory test. The first test was administered to subjects a few minutes after they watched the film, but a day before they read the misleading descriptions. Because the initial test was given so soon after the film, it may have given subjects the chance to consolidate their memory for the original event, making them more resistant to misleading information. Loftus (1979)
suggested that the blatant misinformation may have made subjects more resistant to all misinformation, but, perhaps the timing of the initial memory test also increased their resistance.

Most misinformation studies have focused on the influence of misinformation for objects (e.g., car, wallet) featured in events such as thefts or car accidents (e.g., Loftus, Miller & Burns, 1978). Although subjects are sometimes given misinformation about the suspects featured in these events, much less is known about the influence of misleading information about a suspect's physical appearance on memory for the suspect (see Loftus & Green, 1980; Jenkins & Davies, 1985). Loftus and Green conducted one of the few studies that have focused exclusively on the influence of misleading composite drawings of a suspect's face on subjects' memories for that face.

Loftus and Greene (1980) conducted a series of experiments in which subjects were shown a set of male faces, a film, or a live male suspect. The next day, subjects read written descriptions of a suspect allegedly written by a professor. The authors did not specify whether the professor was meant to be a credible source, though subjects may have considered the professor to be credible. For the misleading condition, a piece of incorrect information about hairstyle, the presence of a moustache, or the suspect's teeth, was embedded in the professor's descriptions. Subjects in the misleading condition often included the incorrect detail in their written descriptions at test and in their identikit face reconstructions. They also recognized the target face less often than control subjects who had not been given the incorrect information.

To date, misinformation studies have shown that memory for people can be altered, however, they have focused on post-event information that is either accurate or inaccurate. Researchers use "misinformation" synonymously with the word "incorrect". By classifying post-event information as either right or wrong, researchers have not fully
taken into account the fact that there are many different words that can be used to describe a physical attribute.

Rather than classifying post-event information as accurate or inaccurate, I propose that it can be classified according to how likely people would be to use it to describe the physical appearance of another person. When asked to describe a suspect, witnesses may be more likely to use words that characterize a suspect attribute as being perceptually similar to what they saw. For example if a witness saw a dark haired suspect, they may consider black or dark brown to be better ways of describing the hair color than the word light brown because the black and dark brown are perceptually similar and closer to what they saw.

For this study, I am defining post-event information that people consider reasonable or better alternatives for describing a particular suspect attribute as *likely information*. Likely information is information “that seems as if it would happen, [that is] reasonably to be expected, [and] apparently true to facts” (Webster’s College Dictionary). I am defining post-event information that people consider less likely or poorer ways of describing a suspect attribute as *unlikely information*. Unlikely information would be less appropriate for describing a given attribute, however, it would not be considered implausible.

Likely and unlikely information can be conceptualized as falling along a scale. At one extreme of this scale are words that are very likely and plausible. At the opposite end of the scale are words that are very unlikely and seem less plausible. Likely and unlikely information, as I have defined them, fall between these two extremes as shown in Figure 9.

Insert Figure 9
My study was concerned with how far a witness’s memory for various physical attributes of a suspect could be pushed by likely and unlikely post-event information because this information may pose a serious threat a witness’s memory for a suspect. Whenever a witness accepts likely or unlikely information they may change their description of a suspect to include this information. It is possible that a witness will be more willing to accept likely information than unlikely information because, by definition, likely information is a better way of describing an attribute. Problems arise if a witness changes his description between giving a statement to police and testifying in court. Any discrepancies in the witness’s testimony, no matter how subtle, may convince the jury that the witness is unreliable. As a result, the jury may disregard testimony that is otherwise truthful.

The objective of my study was to examine whether likely information about physical attributes of a suspect would impair subjects’ memories for a suspect’s appearance (e.g., hair, height, clothing) more than unlikely information and neutral information. Because likely information should be considered a good way of describing an attribute, I expected that subjects who received likely post-event information would be more willing to accept this information than unlikely information. If likely information reduces the accuracy of subjects’ memories more than unlikely information, then subjects who received likely information would select the best answer on a recognition test less often than those subjects who received unlikely information or neutral information about a physical attribute.

Two phases were required to investigate my objective. Phase 1 was a preliminary step to the main study that was conducted in phase 2. The purpose of phase 1 was to empirically determine what words subjects would consider likely and unlikely ways of describing a particular suspect attribute. In phase 1, subjects were
asked to rate the likelihood of a series of statements that described physical attributes of two photographed women (e.g., hair color, clothes color, face shape). Once likely and unlikely were determined in this phase, I proceeded to phase 2. The purpose of phase 2 was to test my hypothesis that likely information would impair memory more than unlikely information or a neutral information control. For phase 2, different subjects studied the same photographs and were the given likely and unlikely information (collected during phase 1), as well as neutral information about the physical appearance of these suspects. Following this information, subjects were given a multiple choice memory test. For clarity, I will discuss the methods and results of phases 1 and 2 separately in the following sections.

Phase 1: Likelihood Ratings

Method

Subjects & Design

Subjects were twenty-five (21 women, 4 men) undergraduate students recruited from first and second year psychology courses at the University of British Columbia. They were given course credit in exchange for their voluntary participation. The purpose of phase 1 was to determine what kinds of descriptive information subjects would rate as likely and unlikely about certain physical attributes of photographed suspects (e.g., hair color, clothing color, height, facial expression).

Materials

Photographs. Three color photographs were used in this study. Each photograph showed a different woman. Two of these photographed women were used as the suspects in this study (see photos in Appendix C). The photograph of the third woman was used as a sample for a practice session. The two suspects were in their early twenties, while the third photographed woman was in her thirties. Caucasian
women were used to prevent suspect race from acting as a confound and to ensure that the results could be generalized to more than one photographed woman. Gender neutral names (Mel and Lee) were assigned to the suspect photographs in preparation for phase 2, in which these names would aid subjects in remembering the women.

_Likelihood questionnaires._ Two likelihood questionnaires were used in this study. The purpose of these questionnaires was to determine whether subjects considered a variety of descriptive words to be likely (i.e., possible, reasonable) descriptors of physical attributes of the photographed women. The questionnaires can be found in Appendix D. One questionnaire was created for each suspect. Each questionnaire contained 15 items. Each item consisted of a statement about the appearance of one of the two suspects. Five answer choices were listed below each statement. These answer choices ranged from descriptive words which, in the author's opinion, were very good descriptors, to words that were poor, but plausible descriptors. For example, item number one read, “Mel’s hair color could be described as . . . .?” This statement was followed by five answer choices: a) brown, b) dark brown, c) black, d) blonde, e) brown with red highlights. A five point rating scale appeared at the top of each questionnaire. For this scale, 1 meant very likely and 5 meant very unlikely. In addition, a practice item and accompanying answer choices appeared underneath the rating scale on each questionnaire. The practice item was used to give subjects practice with completing the questionnaire items.

_The North American Adult Reading Test (NAART)._ This test measured verbal intelligence and consisted of 61 words (Spreen & Strauss, 1998). These words were presented on two white 8.5” x 11” sheets of paper. Each sheet contained two columns of words. The NAART was intended to screen out subjects who were not proficient in English, however, the experimenter felt that all subjects were proficient enough to
complete the tasks they were given. A tape recorder was used to record subjects' responses for later scoring.

Procedure

Subjects were first given a practice session to familiarize them with the task of completing the likelihood questionnaires. In this practice session, subjects were shown a practice photograph and instructed that they would be shown photographs of two different women. They were told that they would be asked to decide how likely or unlikely they would be to describe various parts of a person in a particular way (e.g., height, hair). One of the likelihood questionnaires and a practice photograph of a woman was then placed in front of them. Subjects were told that they would be completing this questionnaire as well as another one. The order in which subjects received the questionnaires was counterbalanced. Subjects were then guided through a practice item on the questionnaire.

For the practice item, the experimenter placed the practice photograph beside the likelihood questionnaire. The experimenter then pointed to the practice statement at the top of the page and told subjects that “this statement asks you to rate the likelihood that this woman’s hair could be described in each of these different ways. So for example, how likely or possible is it that you would describe her hair as blonde or brown or maybe green?” (see Appendix D). The experimenter pointed to the five answer choices below the statement and explained how to use the 5 point likelihood rating scale to rate each answer choice. Subjects were then prompted to complete the practice statement by writing their ratings in the spaces beside the answer choices. If subjects seemed uncertain of the task, the experimenter read the answer choices out loud, prompting subjects to use the rating scale and to look at the practice photograph, repeating the task instructions as needed. After the subject had completed the practice
item, the experimenter suggested that “it will help if you imagine seeing these people in
different circumstances for example dim light, out of the corner of your eye as you walk
quickly past, under different lighting etc”. The experimenter told subjects that these were
just suggestions to help them think of all the possible ways they could describe each
suspect and that they should do whatever worked best for them. These instructions
were intended to encourage subjects to discriminate between the five answer choices
and to think about whether they would use each of the answer choices to describe the
suspect. The practice session concluded following these instructions, when subjects
indicated that they understood the task.

 Likelihood Rating Task. Immediately following the practice task, the likelihood
questionnaires were administered. One of the suspect photographs was placed on the
table beside the questionnaire for that photo. Subjects proceeded to complete the
likelihood questionnaire for the photograph while viewing the photograph. After
indicating that they had completed the first questionnaire, the experimenter removed the
questionnaire and photo and placed the next photo and corresponding questionnaire in
front of the subject. Subjects completed the questionnaires at their own pace.

 North American Adult Reading Test (NAART). The North American Adult
Reading Test (NAART) was administered immediately following the questionnaires. The
NAART was administered according to published guidelines (Spreen & Strauss, 1998).
Subjects’ responses were tape recorded for later scoring. The NAART took less than 5
minutes to complete.

 Results & Discussion

 The likelihood questionnaire items were analyzed separately for Suspect A and
Suspect B. For each suspect photograph, I calculated the mean likelihood rating for
each answer choice of each item. To calculate the mean likelihood rating for each
answer choice, I averaged all of the ratings provided by subjects for that particular answer. For example, item number one on the questionnaire for Suspect A read, “Mel’s hair color could be described as . . . ?”. This statement was followed by five answer choices: (a) brown, (b) dark brown, (c) black, (d) blonde, (e) brown with red highlights. The mean likelihood rating for these choices were as follows: (a) = 1.48, (b) = 1.28, (c) = 2.84, (d) = 4.64, (e) = 2.76 (see Appendix E for these data).

After calculating the mean likelihood ratings for each answer choice, these means were examined to determine which answer choice was rated as most likely, second most likely and so on, for each item. The means were examined separately for each photograph. As a reminder, the rating scale on the likelihood questionnaire ranged from 1, meaning very likely, to 5 meaning very unlikely. For each item, the answer choice with the highest likelihood rating (i.e., closest to 1) was considered to be the correct (i.e., the best) answer. The correct answer was considered to be the choice that best described the physical attribute featured in the item, although it was not necessarily factually correct. The answer choice that was rated as second most likely was considered to be a likely answer. The answer choice rated as least likely (i.e., closest to 5) was considered an unlikely answer. Due to computation errors, the answer choice rated as fourth most likely, rather than least likely was chosen as the unlikely information for two items (item 7 pants type for Suspect A and item 9 shirt type for Suspect B). However, the means for these answer choices did not differ significantly from the choices rated as least likely and thus this did not affect my data. Each of the items on the likelihood questionnaires had a correct choice, a likely choice and an unlikely choice. Appendix D shows the two likelihood questionnaires and which answer choices were used as the likely information and unlikely information. The correct or the best answer is also indicated by the letter “B” for each item on the questionnaires.
Phase 2: The Influence of Post-event Information on Memory for Suspects

Introduction

The purpose of phase 2 was to test the hypothesis, introduced earlier, that likely information regarding the physical appearance of photographed suspects would not impair memory more than unlikely or neutral information. If likely information impairs memory more than unlikely information, it was expected that subjects would not select the correct answer as often on the recognition tests of the photographed suspects when they were given likely information compared with unlikely information. The neutral information was used as a control condition.

Method

Subjects & Design

Twenty-five subjects (18 women, 6 men) were recruited from first and second year undergraduate psychology courses at the University of British Columbia. Subjects were given course credit for their voluntary participation.

The type of information regarding the physical appearance of the photographed suspects was a within-subjects factor. All subjects received likely information, unlikely information and neutral information about the physical appearance of the photographed suspects. The dependent variable was subjects’ accuracy on the multiple choice items on the recognition tests for the photographed suspects.

Materials

Photographs. The same photographs of the two Caucasian women used in phase 1 were studied by a different group of subjects in phase 2.

Filler Tasks. The purpose of the filler tasks was to distract subjects and to prevent floor and ceiling effects from occurring at test. The order of the filler tasks was as follows: Big Five Inventory, North American Adult Reading Test (NAART), height and
weight judgments of series of photographed people (not including the two suspects), Owl and Lark Questionnaire, Mystery Stories Recognition Test, and demographic information (Coleman, 1976; B. Meier, personal communication, Fall, 2001; Spreen & Strauss, 1998).

**Neutral Information.** A neutral piece of information was created for each item on each of the likelihood questionnaires. The purpose of creating a piece of neutral information for each item was to have a neutral information control condition for the present study. The likely and unlikely information collected in phase 1 (see phase 1 Results & Discussion) were pasted together with the neutral pieces of information to create written paragraph descriptions of each of the photographed suspects. The creation of these written descriptions will be described fully in the Written Descriptions section of this method.

To create the neutral information for each item on the likelihood questionnaires, I wrote a phrase that mentioned the key physical attribute referred to in that item (e.g., hair color). However, the neutral information did not inform the reader about the physical appearance of the attribute. For example, item number one from the likelihood questionnaire for Suspect A stated that "Mel's hair color can be described as ____"? The neutral piece of information for this item was, "I noticed the color [of Mel's hair]" (Appendix F). This neutral information mentioned the key idea of hair color but did not say what color it was. In contrast, the likely information for item number one clearly described the color of Mel's hair as being brown (See Appendix D for likely and unlikely information).

**Written Descriptions.** Six written descriptions were created for each of the two photographed suspects. Each description was between 59 and 62 words long and described the physical appearance (e.g., height, hair, clothing) of one of the
photographed women. The descriptions were used to communicate the likely, unlikely and neutral information about the women’s appearances to subjects. The following description was used in this study and is shown here as an example. This description contains likely, unlikely and neutral information about the photographed woman named Mel (see Appendix G for descriptions).

Target A Form 5

Mel weighed about 110-120 pounds. She wore a navy blue knitted sweater. Her facial expression was clearly visible, and her hair was blonde, and the style was noticeable. Her face was square and she wore dress pants that were navy blue with white. I noticed Mel’s age, but wasn’t sure about the eyebrows. She was about 5’6” – 5’8” tall with a normal bodybuild (Appendix G).

To create the descriptions for each photographed woman, the 15 items from the likelihood questionnaires were randomly divided into three groups of items. One of the 15 items was not used because the experimenter felt that subjects would not find it plausible. As a result, each group contained 4 or 5 items. The likely, unlikely and neutral types of information were counterbalanced across these three groups of items to ensure that all subjects received likely, unlikely and neutral information in each written description. For example, one subject received a description of Suspect A that contained likely information about hair color, unlikely information about height and neutral information about shirt color. This same subject may have received likely information about Suspect B’s height, unlikely information about shirt color and neutral information about hair color.

Due to errors in assigning information to written descriptions, one description for Suspect A contained the wrong piece of likely information about weight (i.e., Item 13),
while one description for Suspect B contained the wrong piece of unlikely information about height (i.e., Item 12). However, the means for these answer choices did not differ significantly from the choices rated as likely or unlikely, respectively, on the likelihood questionnaires. These errors did not affect my results.

Table 5 shows how the likely, unlikely and neutral information were counterbalanced across the three groups of items. In Table 5, the types of information are labeled L = likely, U = unlikely and N = neutral. The groups of items are labeled Group 1, Group 2 and Group 3 because the items were placed in these groups at random. The codes A1, A2, A6, etc. identify each of the descriptions that subjects read. For example, the code A1 indicates that the description was written about Suspect A and that it was description number one out of a total of six descriptions written for that suspect.

Memory Tests. A multiple choice recognition test was constructed for each of the photographed women (shown in Appendix H). The purpose of the tests was to determine whether subjects who received likely, unlikely or neutral information about a particular suspect attribute (e.g., hair color, shirt color) would choose the correct answers as opposed to other answer choices on the test items. Each memory test consisted of the same items and answer choices that were used in the likelihood questionnaire from phase 1. Each test contained a 5 point confidence rating scale printed at the top of the page (1 = very confident, 5 = not confident at all) which subjects used to rate their confidence in their answer choice for each item.

Procedure

Study Session. Subjects were tested in one hour-long session. Subjects were given the two photographs, one at a time, and asked to imagine that they were a
witness to a crime and that these photographed women were suspects. The order of presentation for these photographs was counterbalanced (see Table 6). Subjects were instructed to study the photographs for 15 seconds each. This time length was chosen to prevent ceiling effects from occurring at test. The experimenter told subjects that they should try to create a mental picture of the suspects to help remember them. The experimenter informed subjects that they would be given a memory test for these women later in the experiment. The instructions to subjects were intentionally brief because piloting work indicated that giving subjects lengthier instructions that pushed them to study the photographs very carefully may lead to ceiling effects at test. A series of filler activities was administered immediately following this task.

Insert Table 6

**Filler tasks.** The order of the filler tasks was as follows: Big Five Inventory, North American Adult Reading Test (NAART), height and weight judgments of a new series of photographed suspects, Owl and Lark questionnaire, mystery stories recognition test, demographic information (Coleman, 1976, B. Meier, personal communication, Fall, 2001; Spreen & Strauss, 1998). The NAART was administered according to published guidelines (Spreen & Strauss, 1998). For the height and weight judgment task, subjects judged the height and weight of photographed people (not including the suspects) while viewing the photos. The filler tasks took approximately 35 minutes to complete.

**Written Descriptions.** Immediately following the filler tasks, subjects were told that they would be reading descriptions of the suspects they had studied earlier. The experimenter told subjects a cover story to prevent them from focusing on whether the descriptions were accurate or inaccurate. For the cover story, the experimenter explained that the descriptions were provided by other students who had studied each photograph for three minutes and then had described the women from memory
immediately following their study session. The experimenter added that these students' descriptions were “quite accurate”. Subjects were instructed to rate the clarity of each written description. They were told to focus on the sentence flow, grammar and writing quality of the paragraph, when making their clarity ratings, rather than on the accuracy. The clarity ratings were used to motivate subjects to read the descriptions carefully and to discourage them from focusing on whether the descriptions were accurate. Subjects were told that they would have 45 seconds to read and to rate the clarity of each description. This time limit was chosen so that all subjects would have enough time to read and understand the information.

Immediately following these instructions, the written description of Suspect A and a copy of the five point clarity rating scale was placed in front of subjects. The experimenter started the stopwatch as soon as the description was in front of the subject. After 45 seconds, the description was removed from the table and subjects were prompted to orally rate the clarity of the description if they had not done this already. The same procedure was followed for the description of Suspect B.

Table 6 indicates how subjects were assigned to receive the written descriptions for Suspects A and B. The codes A1 to A6 represent Suspect A and the six different descriptions numbered 1 to 6. The codes B1 to B6 represent the six written descriptions for Suspect B.

The descriptions were delivered immediately before the memory tests, rather than immediately after studying the photographs because research has shown that giving post-event information immediately prior to testing, will increase the chance that subjects will accept the information and that they may experience confusion over the source of the information (Loftus, Miller & Burns, 1978; Jenkins & Davies, 1985). The description of Suspect A was always delivered before the description of Suspect B.
Memory Tests. Immediately following the written descriptions, subjects were given memory tests for the two photographed women. These tests were administered approximately 35 minutes after viewing the suspect photographs. For the memory tests, subjects were instructed to try to remember the photographed suspects that they had studied earlier. They were instructed to read the answer choices carefully and then to circle the answer choice that they felt was correct. They were also asked to rate their confidence in the answer that they had selected as correct. Confidence ratings were made using a five point rating scale (1 = very confident, 5 = not confident at all). An extra copy of the rating scale was placed in front of subjects so that they would not have to flip back to the first page of the test to see the rating scale. No time limit was imposed and subjects completed the tests in about five minutes. The experiment ended when subjects indicated that they had completed the tests.

Results

To assess whether likely information would impair memory more than unlikely information, the mean proportion of correct responses were compared for the likely, unlikely, and neutral information conditions. Results indicated that likely, unlikely and neutral information did not differ in their influence on the proportion of correct responses to test items. Thus, likely information did not impair memory for suspect attributes more than unlikely or neutral information and my null hypothesis was retained, $F(2,46) = 1.274$, $p = .290$, $\eta^2 = .052$, (1-\(\beta\)) = .263.

After determining that there was no significant effect of information type, I conducted several exploratory analyses to determine whether information type would show any interactions with factors such as subject gender or type of test item (e.g., face, clothing, demographics). I also examined the relationship between confidence and accuracy on the memory tests.
To explore whether information type would interact with various types of test items (e.g., face items, clothing items, demographic items) or with subject gender, I divided the test items into three groups according to items that concerned the face, clothing and demographics of the suspect. These three groups were chosen because the raw data in Appendix I revealed that accuracy on face items and some clothing items was lower when subjects received likely information compared with unlikely information.

Results revealed that information type did not interact with subject gender or with any particular type of test item (i.e., face items, clothing items, demographic items). However, an interaction between test items and subject gender was found as well as a main effect of subject gender. This interaction was significant according to an ANOVA, $F(2,44) = 3.751, p = .031, \eta^2 = .146, (1-\beta) = .656$. When this interaction was broken down, women more accurately remembered clothing items than face items or demographic items on the memory tests. This finding was confirmed with a post hoc Tukey's test, Tukey $a(3, 44) = .2342, p < .05$. The significant effect of subject gender was also confirmed with an ANOVA, $F(1,22) = 8.042, p = .010, \eta^2 = .268, (1-\beta) = .773$.

I also examined subjects' confidence ratings on the memory tests to determine whether the type of information that subjects received had in any way affected their confidence ratings. No significant effects of information type were found, $F(2,46) = 1.58, p = .217, \eta^2 = .064$. In addition, I examined the correlation between subjects' accuracy on the memory tests and their confidence ratings. Subjects were more confident when responding correctly to test items as revealed by Pearson's $r = .474, p = .019$. 
Discussion

The results indicated that likely information did not impair memory for people more than unlikely or neutral information. There are several possible reasons why this result occurred. In the next few sections, I will discuss some of these reasons.

It is possible that a significant effect of information type was not found because power was low, \((1-\beta) = .263\). I suspect that the low power may be attributed to the small number of subjects in this study.

The choice of information that was selected for the unlikely information type may have also affected my results. It is possible that significant results were not obtained because subjects found some of the unlikely information to be too implausible. In the present study, unlikely information was not intended to be implausible, rather, it was intended to be less reasonable than the likely information. My suspicions were supported by informal interviews of subjects that had been conducted at the end of their testing session. Some subjects said that they thought some of the written descriptions that they had read were wrong, or that parts of the descriptions were wrong.

It is also possible that using a 35 minute retention interval allowed for less forgetting or less confusion regarding the source of the memory. Because subjects' memories for the suspects were not weakened over time, they may not have been influenced by the different types of information in the written descriptions. This is supported by research which suggests that longer retention intervals of at least 24 hours result in the misinformation effect more often than shorter intervals of 30 minutes or less (Payne, Toglia & Anastasi, 1994). Longer intervals may lead to misinformation effect because they can help to prevent ceiling effects and they enable researchers to take advantage of subjects' weakened memory trace. A short retention interval was used in my study because successful competition for subjects in a limited subject pool
required an experiment with a short and convenient time commitment that would attract subjects.

The results also indicated that subjects were more confident for correct answers on the memory tests than for incorrect answers. This result supports the common assumption that confident witnesses are accurate (Sporer, Penrod, Read & Cutler, 1995). Perhaps subjects in this study had a fairly strong memory trace for the original photographs because of the short retention interval. The strong memory trace may have helped them to make their confidence ratings.

Taken together, the interaction between item type and subject gender as well as the main effect of subject gender suggest that that women remember female suspects more accurately than men, and that women are better at remembering clothing colors or styles than facial features, hair, or demographics. These results support the results reported in Chapter 2 of this paper. In Chapter 2, women perceived and recalled more details than men and they seemed to perceive and recall more details about clothing than about other physical attributes.
CHAPTER 4

CONCLUSION

Together the studies reported in Chapters 2 and 3 suggest that women are superior to men at remembering suspects, especially women. The study reported in Chapter 2 revealed that men and women perceive and recall suspects differently. The study reported in Chapter 3 did not find an effect of information type, however, it revealed that women remember suspects more accurately than men and that women are especially accurate at remembering a women's clothing. These studies are the first step towards understanding what an eyewitness remembers about suspects.

Because the results of these studies can be explained in many different ways, more work will be needed to determine how interest, experience, verbal ability, information type and additional factors may influence men and women and lead to the differences found in memory for suspects.
References


*Law and Human Behavior, 4, 323-334.*


*Journal of Experimental Psychology: General, 114, 1-16.*


*Bulletin of the Psychonomic Society, 31, 447-448.*


RONNIE
Appendix B

Scoring Key

A copy of the complete scoring key used for the experiment presented in Chapter 2 can be found below.

Memory for Persons: Scoring Key for Picture Protocols

June 13, 2001 Version

General Instructions

There are four pictures, each labeled with a name: Ronnie, Kelly, Pat, Jamie. Each subject has one description protocol and one recall protocol which are scored individually for each of the items below.

This scoring key has thirty (30) items. Most items are scored twice, once for description and once for recall. Matches between description and recall for each of these items are also scored and recorded. Intrusions or errors in subjects' descriptions and recollections are also recorded.

For each subject's description and recall protocol, you need to count the occurrence of each of the items that are described below. These items are scored by counting words and phrases that meet the criteria for the item. For all items, the first occurrence of any information unit (i.e., word or phrase) is given 1 point. Repetitions, either verbatim or similar in meaning, not counted (except for the Item dealing with repetitions). Any word or phrase may be counted for only one item; no word or phrase should be counted twice, that is, for two different items.

When scoring these protocols, it is very important to consider the context in which the subject describes something. For example, if a subject mentions the word "shoulder", determine if they are describing what the shoulders look like (e.g., "broad shoulders") or whether they are actually describing a bag that is "slung over the
shoulder". In the first case, the subject is describing the body part (shoulder). In the second case, they are not describing the "shoulder", instead, they are describing the "bag". They mention the shoulder simply to help describe how the bag is positioned.

Language

Total Word Count.

DO NOT DO THIS BY HAND!! This count can be obtained from MS Word. A count of the total word length of each protocol. The total word count does include repetitions. However, this count does not include the non-word noises that are often used to fill pauses. This count does not include words spoken by the experimenter. Before using MS Word to count the words, all words spoken by the experimenter must be deleted. All filled pauses and titles must also be deleted prior to the word count.

Total Repetitions. (RPs)

A count of all instances, words and/or phrases, when participants repeated themselves (e.g., he has... he has green eyes). The focus is on idea units, and the count includes only units that were used to convey the same idea in the same context.

For example, if the subject describes the photographed person's body as being thin, then the second (and third, etc.) time the idea of “thin” is mentioned, this is considered a repetition. The underlined phrases are repetitions (e.g., He's a thin uh uh a very slim guy. He has blue eyes, brown hair and he is really really skinny. 2 points)

This count does not include idea units that are repeated (that have to be repeated) as part of normal language use (i.e., The red jacket ... it had green buttons). Do not include repetitions of filled pauses in this count.

Filled Pauses. (FPs)

A count of all the instances of pauses that were filled with some non-word noise, such as um, ah or er. Count every filled pause. A pause is counted only once, whether it
Memory for Suspects

is filled with only one noise (um) or with many different noises (um um uh). The total length of the pause does not matter. You may want to score this item first because it is very easy to miss some of the filled pauses.

Inferences

Assumptions & Evaluations.

A count of words or phrases that refer to what a depicted person is or is assumed/inferred to be. This count focuses on persons' emotional state (e.g. he is sad), on character (e.g., a likeable person) and on possible professional or social roles (e.g. She seems like a student/professor).

This category also includes words and phrases that give an overall impression or evaluation of the person and/or their appearance (e.g., she is elegant). Also includes words or phrases used to refer to the general style (e.g., fashionable, frumpy) or function (e.g., outdoor) of clothing worn by the person in the picture. Also include words/phrases that refer to assumed nationality or country of origin of the depicted person.

Do not include age and ethnicity in this count. Do not include facial expressions, such as "she is smiling" or "he is frowning" unless the subject is expressing their opinion (e.g., "He has a nice smile").

Examples receiving a score of 1: "She looks happy..." "He looks like a janitor." "She looks like a teacher." "He is handsome." "He/She is scary." "He/She seems to be Chinese." "He's dressed casually."

Action or State.

A count of words or phrases that focus on possible actions (passive or active) or action states that are inferred from the depicted person. (Because the pictures are "stills", all action descriptions are inferential.) Do not include phrases like "holding a
cup”; statements like this are not inferred, they are descriptive and will be counted separately.

Examples receiving a score of 1: "He is eating something." "She's looking in her bag." "He seems to be relaxing." "He is drinking coffee."

Age.

The maximum score for this item is 1. A count of words or phrases that refer to a person’s age. Do not count any repetitions of age.

Examples receiving a score of 1: "He is in his early twenties." "A middle-aged woman." "She's young . . . maybe in her mid to late teens or early twenties"

Body Things

Hair.

The maximum score for this item is 1. A count of the occurrence of mentioning a person's hair. Do not count words and phrases that refer only indirectly to hair; indirect statements are counted separately as part of Hair Descriptors.

Examples receiving a score of 1: "She had short hair" "She had curly hair."

Examples that are not counted for this item: "He is balding... has a receding hairline." "She is a brunette."

Hair Descriptors.

A count of the number of words and phrases used for describing hair. For this count, include reference to hair length, hair loss, hairstyle or hairdo (e.g., bangs, ponytail, buzz cut) and hair color. Also include references to where the hair is and/or how it hangs or is parted etc. DO NOT include references to facial hair or lack of facial hair. DO NOT count references to hair clips or pins. These will be counted in Jewelry/Accessories.
Examples: "He has blonde hair." "He is balding" 2 points. "He has black hair, balding at the front..." 3 points. "She has long, dark, straight hair, hanging about three inches below her shoulder." 4 points

Facial Hair.

This category includes words and phrases that identify and describe facial hair only (e.g., beard, moustache, stubble). This category also includes references to lack of facial hair (e.g., clean shaven) and where the facial hair (or lack of) is located on the face. Do not score references to eyebrows, hair or lack of hair on the top of the head in this category; these are scored in other categories (e.g. "He has stubble" "It looks like he hasn't shaved" ". . . a beard").

Eyes/Eyebrows.

The maximum score for this item is 2. A count of the occurrence of the words eyes and eyebrows. Do not count words or phrases that refer to eyes or eyebrows only indirectly ("she is squinting"); these are counted separately.

Example receiving a score of 1: "He has blue or grey eyes."

Example receiving a score of 2: "She had big eyes and curved eyebrows..... she has dark eyes.

Eye/Eyebrow Descriptors.

A count of the number of words and phrases (or distinct ideas) used to describe eyes and eyebrows. Example receiving a score of 1: "... he had light eyes." Example receiving a score of 3: "... she had big dark eyes and really curved eyebrows."

Face/Facial Features.

A count of words and phrases that identify the face itself and any internal facial features (i.e., other than eyes, eyebrows, hair). In this count, include specific mention of the face, nose, cheekbones, cheeks, chin, forehead, lips, teeth.
Do not count descriptors of facial features; they are counted separately.

Do not count "head" or "ears" in this category. They are not facial features, they are body parts. Example receiving a score of 1: "She has high cheekbones..." Example getting a score of 2: "His nose is round, and he has a square jaw ...."

**Face/Facial Features Descriptors.**

A count of words or phrases used to describe all facial parts (or the face as a whole (i.e., other than eyes, eyebrows, hair). For this count, include descriptions of overall face shape (e.g., "...he had a long face."). in addition to descriptors of nose, cheekbones, cheeks, chin, forehead, lips, teeth. Also include descriptions of a person's skin and whether they are wearing/not wearing makeup. Count each distinctive descriptive idea separately (i.e., the nose is big and upturned).

Do not count facial expressions. Do not count the word "face".

Examples receiving a score of 1: "She's wearing makeup." "She has freckles." "She has high cheekbones," "he has dark skin" "he has pimples". Example receiving a score of 3: "His nose is round and upturned, and he has a square jaw "

**Facial Expressions.**

A count of words and phrases used to describe specific facial expressions. For this count, do not include inferences about a person's internal emotional state.

Examples receiving a score of 1: "She is smiling" "He is frowning".

**Body build.**

A count of the words or phrases used to describe the person's body build. This item includes general descriptions of body build, height and weight (e.g., stocky, tall, small-boned). This item also includes specific references to size, height and weight (e.g., 6 feet tall, weighs 135 pounds, etc.) Do not include impression statements, like "she is sexy", "he looks like a gangster"; these descriptors are counted separately.
Examples: "She's medium size." "He is fat". (1 point each). "He's heavy-set and probably weighs a lot. About 160-170 pounds." (3 points). "He is tall and thin." (2 points) "He is tall, probably about 6 foot 2." (2 points)

Body Parts.

A count of the words or phrases that identify body parts other than the person's face. This item includes body parts that the subject actually describes and focuses on. Only body parts that the subject describes and focuses on should be included in this category. For example, if the subject says that the "bag is slung over the shoulder", then they are not describing the shoulder itself.

Examples receiving a score of 1: "He has a thick neck." "He has broad shoulders." "She has strong legs."

Body-part Descriptors.

A count of the words or phrases that describe or characterize body parts other than the person's face. Each piece of descriptive information is counted individually.

Examples receiving a score of 1:"He has a thick neck." "He has broad shoulders."

Examples receiving a score of 2: "Her fingernails are painted black" (2 points). "She's wearing red nailpolish "

Body & Body-Part Posture.

A count of the number of words or phrases used to describe a person's posture. Specifically, this includes how a person's entire body is positioned (e.g., standing up) or what they are doing (e.g., holding something). This item also captures words and phrases that describe how arms and legs are positioned or whether the head is turned etc. For example, the subject may identify the person's arms and describe how (or
where) the arms are positioned (e.g., the arms are crossed; arms are on the table). In this example, the idea of “crossed” is given 1 point and “on the table is” given 1 point.

Example getting a score of 1: "He is standing up." "She has her legs crossed." "He is holding a cup." "His hands in his back pockets." "He is resting his arm on the table."

Non-Body Things

Clothes.

A count of all words and phrases that name specific items of clothing that the person is wearing. Each piece of clothing is counted separately: shirts, pants, shorts, jackets, T-shirts, turtlenecks, etc.. Do not count parts of clothing items (e.g., lapel, collar, sleeve, cuff, etc.); they are counted together with clothing descriptors.

Do not count any mention of clothing items which are correctly identified as not being present in the photo; this belongs in the Absent clothing/jewelry category

Example receiving a score of 3: "...she is wearing a yellow, purple and black jacket, black jeans and a green turtle-neck."

Clothing Descriptors: Non-Color.

A count of words and phrases used to describe or characterize specific pieces of clothing. For this count, include words and phrases that refer to parts of specific clothing items (e.g., lapel, cuff, sleeve collar, buttons, etc.), type of clothing items (e.g., jean jacket), style of clothing items (e.g., bell-bottom, collarless, Cargo), brand names of clothing items (e.g., CK, Levis). Also count any mention of signs, marks, patterns (e.g., stripes, lined), and writing appearing on clothing items. Do not count any mention of colors; color is counted separately. Do not count words and phrases that refer to the arrangement or position of clothing items or to the position of any designs, lines,
patterns, labels, stripes, etc on the clothing (e.g., one appears on top of the other, the shirt is open); this information is included in Object Position and Arrangement.

Example receiving a score of 4: "He was wearing a jean jacket with brown collar and cuffs. And he was wearing Cargo shorts."

Note: The word jeans can be used in two different ways, for describing the material of a clothing items (e.g., "he was wearing a jean jacket") -- in which case it will be counted for this item, or for identifying an article of clothing (e.g., "He is wearing light blue jeans") - in which case the word will be counted as part of the Clothes Item.

**Clothing Descriptors: Color.**

A count of all words and phrases used to describe and characterize references to the colors of each piece of clothing or specific parts of clothing articles. Each piece of color information is counted separately.

Example receiving a score of 2: "she is wearing a blue and green jacket", "his shirt has yellow stripes and a red collar".

**Jewelry/Accessories.**

A count of words and phrases that identify jewelry and accessories on the person. Jewelry/Accessories include: watches, rings, necklaces, earrings, hair clips, and belts, etc. Do not include shoes, keys, makeup or nailpolish.

Example receiving a score of 2: "She is wearing a gold chain and pearl earrings."

**Jewelry/Accessories Descriptors.**

A count of the words and phrases used to describe jewelry and accessories. For this count, include any descriptive information that refers to specific jewelry or accessory items (e.g., watchband). Each word used to describe a piece of jewelry/accessory is counted individually. Do not count the position or arrangement of jewelry or accessory items; this information is counted separately.
Examples receiving a score of 2: "She's wearing pearl earrings and a gold chain ....". "He is wearing a watch with a black watchband."

**Object Position & Arrangement.**

A count of words or phrases that describe where clothing, parts of clothing and jewelry or accessories are located (e.g., around the neck, over the t-shirt, on the left side, etc.). This also includes the location of smaller details on clothing such as stripes, wrinkles etc. This item also captures words and phrases that describe how clothing, parts of clothing and jewelry or accessories are arranged (e.g., the shirt is open, the sleeves are rolled up).

Examples receiving a score of 1: "There is a small rip in the right thigh of the jeans" "He's wearing a white shirt underneath the jacket." "The shirt is open." "There's a gold necklace around her neck."

**Absent Clothing, Jewelry & Accessories.**

A count of words and phrases that describe a particular article of clothing, jewelry, or accessory which the subject correctly identifies as NOT being present in the photo (e.g. He is not wearing glasses). Do not include words and phrases that describe a lack of hair, lack of height/weight, lack of makeup etc. (e.g. Her hair is not very long, he is not very tall, she doesn't weigh very much). These will be scored in other categories. Descriptions of lack of tattoos or scars may also be included here, though these are very rare.

Examples receiving a score of 1: "He is not wearing a watch." "He doesn't have any tattoos."
Demographic Information

Ethnicity.

The maximum score for this item is 1. A count of words and phrases that identify the person's ethnic and/or cultural background.

Examples receiving a score of 1: "She's Asian." "This is a black person." "He is white."

DO NOT count words or phrases that seem intended only to describe skin color, such as "he has dark skin" or "she is fair skinned". Skin color descriptors are counted separately. Do not count phrases such as "She looks Korean" or "He looks like he's from China." These will be scored under the category Assumptions & Evaluations because they are inferences.

Gender.

The maximum score for this item is 1. A count of words and phrases that identify the person's gender. In view of the fact that the encoding task is to describe pictures of persons, the may identify gender indirectly, by means of pronouns only (e.g., he, she). Include such words and phrases in the count. Do not count repetitions, such as when a person is first identified as a woman and then is referenced by means of the pronoun she.

Example getting a score of 1: "He is wearing a hat". "The woman is Asian". "It's a man". "She's an Asian girl..."

Name.

The maximum score for this item is 1. A count of proper names (shown on each photo) used to identify each person (e.g., "Ronnie is an Asian male." "Pat is Caucasian...").
Context Information

Surroundings.

A count of words and phrases that identify other objects in the photograph (other than the person). This can include objects that the person is holding or carrying and anything else (e.g., other people and objects) in the background setting of the photo.

Example receiving a score of 1: "He is carrying a bag." "She has a coffee cup." "He has some keys on his belt."

Surrounding Descriptors.

A count of words and phrases that describe the surroundings. Surrounding descriptors may also characterize the setting in which a photo was taken. Count each piece of descriptive information individually.

Examples receiving a score of 1: "She is sitting on a white chair." "He has some keys on his belt." "The dog has a chain on his neck." "It's a patio picture." "The bag has one strap." 2 points (because there are 2 ideas)

Order of Recall

On the scoring sheet:

1. Record the order in which the subject describes the photographed people at study.

2. Record the order in which the subject recalls the photographed people at test.

Matches

A count of the similarity between description and recall protocols. Calculate this separately for each of the items in the scoring key. Count the total number ideas that were given both in the description protocol and in the recall protocol. For this count, the focus is on the idea or concept overlap (i.e., overlap in meaning) between description and recall, not on the overlap on exact words and phrases.
Matches include only accurate information that is mentioned in both the
description and the recall protocols. Record matches under "M" on the scoring sheet.

Examples receiving a score of 1:

Description: "He is wearing a sports jacket"
Recall: "He was wearing a jacket kinda like sporty."

Description: "She had curly hair"
Recall: "Her hair was curly."

Errors.

A count of words and phrases in the recall protocols that identify objects or
object attributes that were not present in any picture, that are remembered
inappropriately, or that were confused between pictures.

Errors include only information that is inaccurate in the recall protocols. Record
errors under "I" on the scoring sheet. Do not include any inaccurately recalled
information in the TEST "T" column on the scoring sheet.

Examples receiving a score of 1:

Description: "She was wearing a gold chain."
Recall: "She was wearing a silver chain."

Description: "He's holding his dog on a leash."
Recall: "He's walking his dog."

Description: "He has a blue and green sweater"
Recall: "He is wearing a black sweater"

Gains.

This will be computed in Excel. A count of the amount of information that is
present at recall but not in the encoding protocol.

Example receiving a score of 1:
Memory for Suspects 65

Encoding:  "He has short black hair."

Recall:  "He had short black hair and his hairline was receding."

Losses.

This will be computed in Excel. A count of the amount of information that is present in the encoding protocol but not in the recall protocol.

Example receiving a score of 1:

Encoding:  "Her hair is black and is tied up... it looks kinda out of place."

Recall:  "She had black hair."

These photographs are shown smaller than actual size to conserve space.

Suspect A "Mel"

Suspect B "Lee"
Appendix D

Likelihood Questionnaires

The following two likelihood questionnaires were used in Phase 1 of Chapter 3.

The original formatting and headings of the questionnaires have been retained for clarity. The answer choices that were used as the likely and unlikely information for the written descriptions are marked “L” for likely and “U” for unlikely beside each of the items. “B” indicates the best answer that was used as the correct answer choice on the memory test for Phase 2.

Subject ID: ______ Subject Gender: ______ Date: _____________ Target Order: ______

Likelihood Judgments Questionnaire (Target A)

Read each statement carefully. For each statement, rate all of the answers according to their likelihood. Write your rating in the space beside each answer choice. Use the photographs to help you make your ratings.

Use the following scale to make your likelihood ratings:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Very Likely</td>
<td>Likely</td>
<td>Neutral</td>
<td>Unlikely</td>
<td>Very Unlikely</td>
</tr>
</tbody>
</table>

Example
The woman’s hair color could be described as ______.
brown ______
dark blonde ______
green ______
blonde ______
light blonde ______
e) brown with red highlights ______

1). Mel’s hair color could be described as ______.
a) brown ______ L
b) dark brown ______ B
 c) black ______
d) blonde ______ U

e) brown with red highlights ______

2). Mel’s hair could be described as ______.
a) very curly ______
b) slightly wavy ______ B
 c) straight ______ L
 d) frizzy ______ U
e) stringy ______

3). Mel’s face could be described as ______ shaped.
a) round ______
b) oval ______ B
Mel's eyebrow color could be described as _______.

4. Mel's ethnicity could be described as _______.

5. Mel could be described as wearing ________.

6. Mel could be described as ________.

7. Mel could be described as ________.

8. The color of Mel's lower-body clothing could be described as ________.

9. Mel could be described as wearing a ________.

10. Mel's ethnicity could be described as ________.

11. Mel could be described as ________.

12. Mel could be described as being ________ tall.

13. Mel could be described as ________.
14). Mel's body-build could be described as ________.
   a) slender    B
   b) big-boned   
   c) stocky     U
   d) small-boned L
   e) muscular   

15). Mel could be described as being about ________ years old.
   a) 15-19     L
   b) 20-25     B
   c) 26-30     
   d) 31-35     
   e) 36-40     U
Likelihood Judgments Questionnaire (Target B)

Read each statement carefully. For each statement, rate all of the answers according to their likelihood. Write your rating in the space beside each answer choice. Use the photographs to help you make your ratings.

Use the following scale to make your likelihood ratings:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Very Likely</td>
<td>Likely</td>
<td>Neutral</td>
<td>Unlikely</td>
<td>Very Unlikely</td>
</tr>
</tbody>
</table>

Example: The woman's hair color could be described as _______.

a. brown _____
b. dark blonde _____
c. green _____
d. blonde _____
e. light blonde _____

1). Lee's hair color could be described as _______.
   a) brown   _____
   b) red _____
   c) dark brown _____
   d) light blonde   _____
   e) reddish-brown   _____

2). Lee's hair could be described as _______.
   a) very curly   _____
   b) slightly wavy   _____
   c) straight   _____
   d) stringy _____
   e) frizzy _____

3). Lee's face could be described as _______ shaped.
   a) round   _____
   b) oval   _____
   c) pear _____
   d) square   _____
   e) egg _____

4). Lee's eyebrow color could be described as _______.
   a) dark brown   _____
   b) black   _____
   c) red _____
   d) light blonde _____
   e) reddish-brown   _____

5). Lee could be described as wearing _______.
   a) bright red lipstick   _____
   b) pale pink lipstick _____
   c) clear lipgloss   _____
   d) red lipstick _____
   e) no lipstick   _____

6). The color of Lee's upper-body clothing could be described as _______.
   a) light purple   _____
   b) red   _____
   c) maroon _____
   d) dark purple _____
   e) pink   _____
7). Lee could be described as wearing
   a) a short skirt
   b) a mini skirt _U_
   c) an ankle-length skirt _L_
   d) a long skirt _B_
   e) a dress

8). The color of Lee's lower-body clothing could be described as
   a) black _B_
   b) dark grey _L_
   c) light grey _U_
   d) deep purple __
   e) navy blue _L_

9). Lee could be described as wearing a
   a) t-shirt _B_
   b) tank top __
   c) short-sleeve blouse _L_
   d) pullover _U_
   e) sweater __

10). Lee's ethnicity could be described as
    a) Caucasian _B_
    b) Caucasian/Asian _L_
    c) Native Indian __
    d) Asian _U_
    e) Hispanic __

11). Lee could be described as
     a) smiling _B_
     b) frowning _U_
     c) grinning __
     d) smirking __
     e) pouting __

12). Lee could be described as being about tall.
    a) 5'0"-5'3" _L_
    b) 5'4"-5'6" _B_
    c) 5'7"-5'9" __
    d) 6'0"-6'2" _U_

13). Lee could be described as weighing about_________ pounds.
    a) 110-120 _U_
    b) 120-130 __
    c) 130-140 _L_
    d) 140-150 _B_
    e) 150-160 __

14). Lee's body-build could be described as
     a) average __
     b) chubby _B_
     c) thin _U_
     d) stocky _L_
     e) muscular __

15). Lee could be described as being about_________ years old.
    a) 15-19 _L_
    b) 20-25 _B_
    c) 26-30 __
    d) 31-35 __
    e) 36-40 _U_
Appendix E

Likelihood Questionnaire Data

The following tables display the means and standard deviations from the likelihood questionnaires for each photographed suspect. The column marked Info Type indicates which answer choices were selected as likely, unlikely and correct pieces of information. The information types are represented by the following letters: L = likely information, U = unlikely information. B indicates the correct answer for each item.

Table E1
Likelihood Questionnaire Data for Suspect A

Photograph A "Mel"

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean</th>
<th>SD</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hair Color</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a brown</td>
<td>1.48</td>
<td>0.77</td>
<td>L</td>
</tr>
<tr>
<td>1b dark brown</td>
<td>1.28</td>
<td>0.98</td>
<td>B</td>
</tr>
<tr>
<td>1c black</td>
<td>2.84</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>1d blonde</td>
<td>4.64</td>
<td>0.70</td>
<td>U</td>
</tr>
<tr>
<td>1e brown with red highlights</td>
<td>2.76</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>2 Hair Style</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a very curly</td>
<td>3.88</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>2b slightly wavy</td>
<td>1.44</td>
<td>0.58</td>
<td>B</td>
</tr>
<tr>
<td>2c straight</td>
<td>2.76</td>
<td>1.16</td>
<td>L</td>
</tr>
<tr>
<td>2d frizzy</td>
<td>3.88</td>
<td>0.88</td>
<td>U</td>
</tr>
<tr>
<td>2e</td>
<td>stringy</td>
<td>3.60 0.96</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Face Shape</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a</td>
<td>round</td>
<td>2.72 0.94</td>
<td></td>
</tr>
<tr>
<td>3b</td>
<td>oval</td>
<td>1.88 0.93 B</td>
<td></td>
</tr>
<tr>
<td>3c</td>
<td>pear</td>
<td>3.56 1.19</td>
<td></td>
</tr>
<tr>
<td>3d</td>
<td>square</td>
<td>4.36 0.81 U</td>
<td></td>
</tr>
<tr>
<td>3e</td>
<td>egg</td>
<td>2.24 1.30 L</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Eyebrow color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>light brown</td>
<td>2.40 1.12 L</td>
<td></td>
</tr>
<tr>
<td>4b</td>
<td>black</td>
<td>3.00 1.22</td>
<td></td>
</tr>
<tr>
<td>4c</td>
<td>dark blonde</td>
<td>3.88 0.97</td>
<td></td>
</tr>
<tr>
<td>4d</td>
<td>blonde</td>
<td>4.64 0.70 U</td>
<td></td>
</tr>
<tr>
<td>4e</td>
<td>reddish-brown</td>
<td>2.00 1.19 B</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Lipstick</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5a</td>
<td>bright red lipstick</td>
<td>4.32 1.07 U</td>
<td></td>
</tr>
<tr>
<td>5b</td>
<td>dark pink lipstick</td>
<td>2.92 1.32</td>
<td></td>
</tr>
<tr>
<td>5c</td>
<td>clear lip gloss</td>
<td>1.84 0.99 L</td>
<td></td>
</tr>
<tr>
<td>5d</td>
<td>brownish lipstick</td>
<td>3.52 1.36</td>
<td></td>
</tr>
<tr>
<td>5e</td>
<td>no lipstick</td>
<td>1.68 0.80 B</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Shirt Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6a</td>
<td>black</td>
<td>1.40 0.65 B</td>
<td></td>
</tr>
<tr>
<td>6b</td>
<td>charcoal grey</td>
<td>3.08 1.19</td>
<td></td>
</tr>
<tr>
<td>6c</td>
<td>navy blue</td>
<td>1.56 0.87 L</td>
<td></td>
</tr>
<tr>
<td>6d</td>
<td>deep green</td>
<td>4.00 0.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>6e</td>
<td>brown</td>
<td>4.44</td>
<td>0.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7 Pants/Skirt Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>7a track pants</td>
</tr>
<tr>
<td>7b sweat pants</td>
</tr>
<tr>
<td>7c tights</td>
</tr>
<tr>
<td>7d dress pants</td>
</tr>
<tr>
<td>7e khakis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8 Pants/Skirt Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>8a Black with white</td>
</tr>
<tr>
<td>8b navy blue with white</td>
</tr>
<tr>
<td>8c black with grey</td>
</tr>
<tr>
<td>8d dark green</td>
</tr>
<tr>
<td>8e very dark grey</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9 Shirt Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>9a sweatshirt</td>
</tr>
<tr>
<td>9b hooded sweatshirt</td>
</tr>
<tr>
<td>9c pullover</td>
</tr>
<tr>
<td>9d hooded jacket</td>
</tr>
<tr>
<td>9e knitted sweater</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10 Ethnicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10a Caucasian</td>
</tr>
<tr>
<td>10b Asian</td>
</tr>
<tr>
<td>10c Caucasian/Asian</td>
</tr>
<tr>
<td>10d Hispanic</td>
</tr>
<tr>
<td>10e Native Indian</td>
</tr>
<tr>
<td>------------------</td>
</tr>
</tbody>
</table>

11 Facial Expression

| 11a smiling      | 1.92 0.76 L |
| 11b frowning     | 3.68 1.14   |
| 11c squinting    | 1.40 0.71 B |
| 11d scowling     | 3.92 1.00   |
| 11e pouting      | 4.04 0.98 U |

12 Height

| 12a 5'0"-5'2"   | 3.00 1.12   |
| 12b 5'3"-5'5"   | 1.56 0.65 B |
| 12c 5'6"-5'8"   | 2.12 1.13 L |
| 12d 5'9"-6'1"   | 3.84 1.28   |
| 12e 6'2"-6'4"   | 4.60 0.71 U |

13 Weight

| 13a 110-120     | 1.96 0.84 L |
| 13b 120-130     | 1.64 0.64 B |
| 13c 130-140     | 2.84 1.25   |
| 13d 140-150     | 3.96 1.02   |
| 13e 150-160     | 4.48 0.77 U |

14 Bodybuild

<p>| 14a slender      | 1.60 0.65 B |
| 14b big-boned    | 3.68 1.22   |
| 14c stocky       | 4.28 0.89 U |
| 14d small-boned  | 2.56 1.08 L |</p>
<table>
<thead>
<tr>
<th>Age Group</th>
<th>Memory</th>
<th>Error</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>15a 15-19</td>
<td>1.64</td>
<td>0.91</td>
<td>L</td>
</tr>
<tr>
<td>15b 20-25</td>
<td>1.60</td>
<td>0.71</td>
<td>B</td>
</tr>
<tr>
<td>15c 26-30</td>
<td>3.32</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>15d 31-35</td>
<td>4.56</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>15e 36-40</td>
<td>4.76</td>
<td>0.52</td>
<td>U</td>
</tr>
</tbody>
</table>

14e muscular, 3.24, 1.30
Table E2

Likelihood Questionnaire Data for Suspect B

Suspect B "Lee"

<table>
<thead>
<tr>
<th>Items</th>
<th>Means</th>
<th>SD</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hair Color</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a brown</td>
<td>1.44</td>
<td>0.71</td>
<td>B</td>
</tr>
<tr>
<td>1b red</td>
<td>3.16</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>1c dark brown</td>
<td>2.68</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>1d light blonde</td>
<td>4.16</td>
<td>1.11</td>
<td>U</td>
</tr>
<tr>
<td>1e reddish-brown</td>
<td>2.00</td>
<td>1.04</td>
<td>L</td>
</tr>
<tr>
<td>2 Hair Style</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a very curly</td>
<td>4.40</td>
<td>0.82</td>
<td>U</td>
</tr>
<tr>
<td>2b slightly vavy</td>
<td>2.68</td>
<td>1.18</td>
<td>L</td>
</tr>
<tr>
<td>2c straight</td>
<td>1.72</td>
<td>0.89</td>
<td>B</td>
</tr>
<tr>
<td>2d stringy</td>
<td>3.16</td>
<td>1.25</td>
<td></td>
</tr>
<tr>
<td>2e frizzy</td>
<td>3.80</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>3 Face Shape</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a round</td>
<td>1.68</td>
<td>1.07</td>
<td>B</td>
</tr>
<tr>
<td>3b oval</td>
<td>2.08</td>
<td>1.04</td>
<td>L</td>
</tr>
<tr>
<td>3c pear</td>
<td>3.68</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>3d square</td>
<td>4.00</td>
<td>1.08</td>
<td>U</td>
</tr>
<tr>
<td>3e egg</td>
<td>2.76</td>
<td>1.45</td>
<td></td>
</tr>
</tbody>
</table>
### 4 Eyebrow Color

<table>
<thead>
<tr>
<th>Color</th>
<th>Score</th>
<th>Error</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark brown</td>
<td>2.48</td>
<td>1.00</td>
<td>L</td>
</tr>
<tr>
<td>black</td>
<td>4.52</td>
<td>0.77</td>
<td>U</td>
</tr>
<tr>
<td>red</td>
<td>2.84</td>
<td>1.21</td>
<td></td>
</tr>
<tr>
<td>light blonde</td>
<td>3.56</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>reddish-brown</td>
<td>1.84</td>
<td>1.07</td>
<td>B</td>
</tr>
</tbody>
</table>

### 5 Lipstick

<table>
<thead>
<tr>
<th>Lipstick Type</th>
<th>Score</th>
<th>Error</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>bright red lipstick</td>
<td>4.88</td>
<td>0.33</td>
<td>U</td>
</tr>
<tr>
<td>pale pink lipstick</td>
<td>2.20</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>clear lipgloss</td>
<td>1.68</td>
<td>0.63</td>
<td>L</td>
</tr>
<tr>
<td>red lipstick</td>
<td>4.16</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>no lipstick</td>
<td>1.08</td>
<td>0.28</td>
<td>B</td>
</tr>
</tbody>
</table>

### 6 Shirt Color

<table>
<thead>
<tr>
<th>Shirt Color</th>
<th>Score</th>
<th>Error</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>light purple</td>
<td>1.00</td>
<td>0.00</td>
<td>B</td>
</tr>
<tr>
<td>red</td>
<td>4.80</td>
<td>0.50</td>
<td>U</td>
</tr>
<tr>
<td>maroon</td>
<td>4.48</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>dark purple</td>
<td>3.84</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>pink</td>
<td>3.68</td>
<td>1.11</td>
<td>L</td>
</tr>
</tbody>
</table>

### 7 Pants/Skirt Type

<table>
<thead>
<tr>
<th>Skirt Type</th>
<th>Score</th>
<th>Error</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>a short skirt</td>
<td>4.68</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>a mini skirt</td>
<td>4.80</td>
<td>0.82</td>
<td>U</td>
</tr>
<tr>
<td>an ankle-length skirt</td>
<td>1.52</td>
<td>1.16</td>
<td>L</td>
</tr>
<tr>
<td>a long skirt</td>
<td>1.20</td>
<td>0.82</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>7e</td>
<td>a dress</td>
<td>3.32</td>
<td>1.22</td>
</tr>
<tr>
<td>8</td>
<td>Pants/Skirt Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8a</td>
<td>black</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>8b</td>
<td>dark grey</td>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>8c</td>
<td>light grey</td>
<td>4.64</td>
<td>0.57</td>
</tr>
<tr>
<td>8d</td>
<td>deep purple</td>
<td>3.16</td>
<td>1.25</td>
</tr>
<tr>
<td>8e</td>
<td>navy blue</td>
<td>2.32</td>
<td>0.99</td>
</tr>
<tr>
<td>9</td>
<td>Shirt Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9a</td>
<td>t-shirt</td>
<td>1.16</td>
<td>0.37</td>
</tr>
<tr>
<td>9b</td>
<td>tank top</td>
<td>4.28</td>
<td>1.14</td>
</tr>
<tr>
<td>9c</td>
<td>short-sleeve blouse</td>
<td>3.04</td>
<td>1.24</td>
</tr>
<tr>
<td>9d</td>
<td>pullover</td>
<td>3.96</td>
<td>0.98</td>
</tr>
<tr>
<td>9e</td>
<td>sweater</td>
<td>3.96</td>
<td>1.27</td>
</tr>
<tr>
<td>10</td>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10a</td>
<td>Caucasian</td>
<td>1.52</td>
<td>0.87</td>
</tr>
<tr>
<td>10b</td>
<td>Caucasian/Asian</td>
<td>2.88</td>
<td>1.39</td>
</tr>
<tr>
<td>10c</td>
<td>Native Indian</td>
<td>3.64</td>
<td>1.11</td>
</tr>
<tr>
<td>10d</td>
<td>Asian</td>
<td>4.28</td>
<td>0.94</td>
</tr>
<tr>
<td>10e</td>
<td>Hispanic</td>
<td>3.96</td>
<td>1.10</td>
</tr>
<tr>
<td>11</td>
<td>Facial Expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11a</td>
<td>smiling</td>
<td>1.12</td>
<td>0.33</td>
</tr>
<tr>
<td>11b</td>
<td>frowning</td>
<td>4.60</td>
<td>0.71</td>
</tr>
<tr>
<td>11c</td>
<td>grinning</td>
<td>2.64</td>
<td>1.38</td>
</tr>
</tbody>
</table>
### 11 Smirking and Pouting

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11d</td>
<td>smirking</td>
<td>3.52</td>
</tr>
<tr>
<td>11e</td>
<td>pouting</td>
<td>4.44</td>
</tr>
</tbody>
</table>

### 12 Height

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12a</td>
<td>5'0&quot;-5'3&quot;</td>
<td>2.84</td>
</tr>
<tr>
<td>12b</td>
<td>5'4&quot;-5'6&quot;</td>
<td>1.88</td>
</tr>
<tr>
<td>12c</td>
<td>5'7&quot;-5'9&quot;</td>
<td>2.68</td>
</tr>
<tr>
<td>12d</td>
<td>6'0&quot;-6'2&quot;</td>
<td>4.36</td>
</tr>
<tr>
<td>12e</td>
<td>6'3&quot;-6'5&quot;</td>
<td>4.72</td>
</tr>
</tbody>
</table>

### 13 Weight

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13a</td>
<td>110-120</td>
<td>4.52</td>
</tr>
<tr>
<td>13b</td>
<td>120-130</td>
<td>3.68</td>
</tr>
<tr>
<td>13c</td>
<td>130-140</td>
<td>2.68</td>
</tr>
<tr>
<td>13d</td>
<td>140-150</td>
<td>2.24</td>
</tr>
<tr>
<td>13e</td>
<td>150-160</td>
<td>2.72</td>
</tr>
</tbody>
</table>

### 14 Bodybuild

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>14a</td>
<td>average</td>
<td>3.44</td>
</tr>
<tr>
<td>14b</td>
<td>chubby</td>
<td>1.84</td>
</tr>
<tr>
<td>14c</td>
<td>thin</td>
<td>4.72</td>
</tr>
<tr>
<td>14d</td>
<td>stocky</td>
<td>2.08</td>
</tr>
<tr>
<td>14e</td>
<td>muscular</td>
<td>4.08</td>
</tr>
</tbody>
</table>

### 15 Age

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15a</td>
<td>15-19</td>
<td>2.08</td>
</tr>
<tr>
<td>15b</td>
<td>20-25</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Memory for Suspects 81</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>-----------------------</td>
</tr>
<tr>
<td>15c</td>
<td>26-30</td>
<td>3.04 1.17</td>
</tr>
<tr>
<td>15d</td>
<td>31-35</td>
<td>4.24 0.93</td>
</tr>
<tr>
<td>15e</td>
<td>36-40</td>
<td>4.88 0.33 U</td>
</tr>
</tbody>
</table>
## Neutral Information for the Written Descriptions

### Table F1
Neutral Information For Suspect A “Mel”

<table>
<thead>
<tr>
<th>Item #</th>
<th>Neutral Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I noticed the color [of her hair]</td>
</tr>
<tr>
<td>2</td>
<td>the [hair] style was noticeable.</td>
</tr>
<tr>
<td>3</td>
<td>[Mel’s] face was normal</td>
</tr>
<tr>
<td>4</td>
<td>Not sure about her eyebrows</td>
</tr>
<tr>
<td>5</td>
<td>I noticed her lips.</td>
</tr>
<tr>
<td>6</td>
<td>[I noticed the color of her] shirt</td>
</tr>
<tr>
<td>7</td>
<td>[Mel was wearing] pants</td>
</tr>
<tr>
<td>8</td>
<td>[Mel was wearing] pants</td>
</tr>
<tr>
<td>9</td>
<td>[Mel was wearing a] shirt.</td>
</tr>
<tr>
<td>11</td>
<td>[Mel’s] facial expression was clearly visible</td>
</tr>
<tr>
<td>12</td>
<td>Not sure about her height</td>
</tr>
<tr>
<td>13</td>
<td>Mel’s weight suited her body.</td>
</tr>
<tr>
<td>14</td>
<td>Mel’s bodybuild was normal.</td>
</tr>
<tr>
<td>15</td>
<td>Mel’s age was noticeable/I noticed Mel’s age</td>
</tr>
</tbody>
</table>

Note: [ . . . ] indicates information that was not included in the written descriptions.
Table F2

Neutral Information for Suspect B "Lee"

<table>
<thead>
<tr>
<th>Item #</th>
<th>Neutral Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I noticed the color [of her hair]</td>
</tr>
<tr>
<td>2</td>
<td>the [hair] style was noticeable.</td>
</tr>
<tr>
<td>3</td>
<td>Lee's face was normal.</td>
</tr>
<tr>
<td>4</td>
<td>Her eyebrows were noticeable/l noticed her eyebrows</td>
</tr>
<tr>
<td>5</td>
<td>I noticed her lips</td>
</tr>
<tr>
<td>6</td>
<td>I noticed the color of her upper clothing /color was noticeable</td>
</tr>
<tr>
<td>7</td>
<td>Lee was wearing a skirt</td>
</tr>
<tr>
<td>8</td>
<td>I noticed the color [of her lower clothing]</td>
</tr>
<tr>
<td>9</td>
<td>Lee was wearing a shirt</td>
</tr>
<tr>
<td>11</td>
<td>Lee's facial expression was noticeable</td>
</tr>
<tr>
<td>12</td>
<td>not sure about her height</td>
</tr>
<tr>
<td>13</td>
<td>Lee weight suited her body</td>
</tr>
<tr>
<td>14</td>
<td>Lee's body build was normal.</td>
</tr>
<tr>
<td>15</td>
<td>I'm not sure about her age.</td>
</tr>
</tbody>
</table>

Note: [ ... ] indicates information that was not included in the written descriptions.

The following description shows the neutral information for Suspect A underlined.

Target A: Form 2

Mel wore a navy blue shirt. She had frizzy hair and I noticed the color. She was about 5'6"-5'8" tall and weighed about 120-130 pounds. Her eyebrows were blonde and her face was normal. Mel was about 36-40 years old and her pants
were navy blue with white. She had a stocky build. I noticed her lips. She was pouting. (see Appendix G).
Appendix G

Written Descriptions

Written Descriptions for Suspect A “Mel”

Target A: Form 1

Mel was wearing pants -- sweatpants. I'm not sure about her height, but her hair was brown and frizzy. Her face was egg shaped. She was pouting and had clear lipgloss on. She was wearing a sweatshirt, and I noticed the color. Her weight suited her body. Her eyebrows were blonde. She looked about 36-40 years old and had a stocky build.

Target A: Form 2

Mel wore a navy blue shirt. She had frizzy hair and I noticed the color. She was about 5'6”-5'8” tall and weighed about 120-130 pounds. Her eyebrows were blonde and her face was normal. Mel was about 36-40 years old and her pants were navy blue with white. She had a stocky build. I noticed her lips. She was pouting.

Target A: Form 3

Mel was smiling and her eyebrows were light brown. Her weight suited her body but I'm not sure about her height. She wore a knitted sweater, and the color was noticeable. She was about 15-19 years old and had straight blonde hair. She had dress pants and had bright red lipstick. She was small-boned and had a square face.

Target A: Form 4

Mel was about 6'2”-6'4” tall and had a brown shirt. I noticed the color of her hair and I remember it was straight. Her eyebrows were light brown. She was smiling and her face was normal. I noticed her lips. She looked about 15-19 years old. Her pants were dark green and she weighed about 150-160 pounds. She was small-boned.
Target A: Form 5

Mel weighed about 110-120 pounds. She wore a navy blue knitted sweater. Her facial expression was clearly visible, and her hair was blonde, and the style was noticeable. Her face was square and she wore dress pants that were navy blue with white. I noticed Mel’s age, but wasn’t sure about the eyebrows. She was about 5’6” – 5’8” tall with a normal bodybuild.

Target A: Form 6

Mel was 6’2”-6’4” tall. Her hair was brown and I noticed the style. Mel’s sweatpants were dark green. Her bodybuild was normal and her facial expression was clearly visible. She wore clear lip gloss and her face was egg shaped. She weighed about 150-160 pounds, but I’m not sure about her eyebrows. She was wearing a brown sweatshirt and I noticed her age.

Written Descriptions for Suspect B “Lee”

Target B: Form 1

Lee was wearing a mini skirt. It was navy blue. Her sweater was pink. Her hair was light blonde and the style was noticeable. Her face was square and her facial expression and eyebrows were noticeable. She wore bright red lipstick. I’m not sure about her age. She looked about 5’0”-5’3”. She weighed about 130-140 pounds and her bodybuild was normal.

Target B: Form 2

Lee weighed about 110-120 pounds. She was wearing an ankle-length light grey skirt. She had reddish-brown hair, and I noticed the style. Her face was oval and I noticed her eyebrows. Her short-sleeve blouse was red and she was about 6’0”-6’2” tall. Her facial expression was noticeable but I’m not sure about her age. Her bodybuild was normal and she wore clear lipgloss.
Target B: Form 3

Lee weighed about 130-140 pounds. She wore a pink shirt. She had a thin build. I noticed her lips. Her skirt was navy blue. She was about 5'0"-5'3" tall. I noticed the color of her hair. It was very curly. She was frowning but her face was normal. Her eyebrows were black and she was about 36-40 years old.

Target B: Form 4

Lee had a thin build. Her weight suited her body. Her face was oval, but I'm not sure about her height. She wore an ankle-length skirt, and I did notice the color of it. Her hair was reddish-brown and it was very curly. She had a short-sleeve blouse on and clear lipgloss. Lee was frowning, her eyebrows were black and she looked about 31-35 years old.

Target B: Form 5

Lee had a stocky build. I noticed the color of her hair. It was slightly wavy. I noticed her lips. Her face was normal and she was grinning. She was wearing a red shirt and she was about 6'0"-6'2" tall. She had a light grey skirt on and weighed about 110-120 pounds. Her eyebrows were dark brown and she was about 15-19 years old.

Target B: Form 6

Lee was wearing a skirt. It was short and I noticed the color of it. Her weight suited her body, but I'm not sure about her height. Her hair was light blonde and slightly wavy. She wore a sweater. I noticed its color. Her bodybuild was chubby. Her face was square shaped and she was grinning. She had bright red lipstick, reddish brown eyebrows and was about 15-19 years old.
Appendix H

Memory Tests for Suspects A and B

Version: Feb 23, 2002
Subject ID: _____  Subject Gender: _____  Date: _________  Condition: _______
Start Time: ______  End Time: ______

Memory Test (Suspect A)

Read each question carefully. For each question, circle the correct answer. Circle only one answer for each question. After you have circled your answer, rate your level of confidence in your answer. To make your confidence rating, use the 5 point scale below. 1 means very confident. 5 means not confident at all. Write the number of your confidence rating in the space marked “confidence” beside each question.

<table>
<thead>
<tr>
<th></th>
<th>Very Confident</th>
<th>Confident</th>
<th>Neutral</th>
<th>Not Confident</th>
<th>Not Confident At All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1). Mel’s hair color was _____.  Confidence: ____
   
a) brown

b) dark brown

c) black

d) blonde

e) brown with red highlights

2). Mel’s hair was _____.  Confidence: ____
   
a) very curly
b) slightly wavy

c) straight

d) frizzy

e) stringy

3). Mel's face was ____ shaped.  Confidence: ____

a) round

b) oval

c) pear

d) square

e) egg

4). Mel's eyebrow color was ____  Confidence: ____

a) light brown

b) black

c) dark blonde

d) blonde

e) reddish-brown

5). Mel was wearing ____  Confidence: ____

a) bright red lipstick

b) dark pink lipstick

c) clear lip gloss lipstick

d) brownish lipstick

e) no lipstick

6). The color of Mel's upper-body clothing was ____  Confidence: ____

a) black
b) charcoal grey

c) navy blue

d) deep green

e) brown

7). Mel was wearing ______. Confidence: ____

a) track pants

b) sweatpants

c) tights

d) dress pants

e) khakis

8). The color of Mel's lower-body clothing was ______. Confidence: ____

a) black with white

b) navy blue with white

c) black with grey

d) dark green

e) very dark grey

9). Mel was wearing a ______. Confidence: ____

a) sweatshirt

b) hooded sweatshirt

c) pullover

d) hooded jacket

e) knitted sweater

10). Mel was ______. Confidence: ____

a) smiling
b) frowning

c) squinting

d) scowling

e) pouting

11). Mel was about ____ tall.  Confidence: ____
   a) 5'0"-5'2"
   b) 5'3"-5'5"
   c) 5'6"-5'8"
   d) 5'9"-6'1"
   e) 6'2"-6'4"

12). Mel weighed about ____ pounds.  Confidence: ____
   a) 110-120
   b) 120-130
   c) 130-140
   d) 140-150
   e) 150-160

13). Mel's body-build was _____.  Confidence: ____
   a) slender
   b) big-boned
   c) stocky
   d) small-boned
   e) muscular

14). Mel was about ____ years old.  Confidence: ____
   a) 15-19
   b) 20-25
c) 26-30

d) 31-35

e) 36-40
Memory Test (Suspect B)

Read each question carefully. For each question, circle the answer that you believe is correct. Circle only one answer for each question. After you have circled your answer, rate your level of confidence in your answer. To make your confidence rating, use the 5 point scale below. 1 means that you are very confident in your answer. 5 means that you are not confident at all. Write the number of your confidence rating in the space marked “confidence” beside each question.

1) Lee’s hair color was ____.
   - a) brown
   - b) red
   - c) dark brown
   - d) light blonde
   - e) reddish-brown

   Confidence: ____

2) Lee’s hair was ____.
   - a) very curly
   - b) slightly wavy

   Confidence: ____
c) straight
d) stringy
e) frizzy

3). Lee's face was ____ shaped.  Confidence: ____
   a) round
   b) oval
   c) pear
   d) square
   e) egg

4). Lee's eyebrow color was ____.  Confidence: ____
   a) dark brown
   b) black
   c) red
   d) light blonde
   e) reddish-brown

5). Lee was wearing ____.  Confidence: ____
   a) bright red lipstick
   b) pale pink lipstick
   c) clear lip gloss
   d) red lipstick
   e) no lipstick

6). The color of Lee's upper-body clothing was ____.  Confidence: ____
   a) light purple
   b) red
c) maroon  
d) dark purple  
e) pink

7). Lee was wearing _____.  Confident: _____.  
   a) a short skirt  
   b) a mini skirt  
   c) an ankle-length skirt  
   d) a long skirt  
   e) a dress

8). The color of Lee's lower-body clothing was _____.  Confident: _____.  
   a) black  
   b) dark grey  
   c) light grey  
   d) deep purple  
   e) navy blue

9). Lee was wearing a _____.  Confident: _____.  
   a) T-shirt  
   b) tank top  
   c) short sleeve blouse  
   d) pullover  
   e) sweater

10). Lee was _____.  Confident: _____.  
    a) smiling  
    b) frowning
c) grinning

d) smirking

e) pouting

11). Lee was about ____ tall.  
Confidence: ____

a) 5'0"-5'3"

b) 5'4"-5'6"

c) 5'7"-5'9"

d) 6'0"-6'2"

e) 6'3"-6'5"

12). Lee weighed about ____ pounds.  
Confidence: ____

a) 110-120

b) 120-130

c) 130-140

d) 140-150

e) 150-160

13). Lee's body-build was ____.  
Confidence: ____

a) average

b) chubby

c) thin

d) stocky

e) muscular

14). Lee was about ____ years old.  
Confidence: ____

a) 15-19

b) 20-25

c) 26-30
d) 30-35  
e) 36-40
Appendix I

Number of Correct Responses for Each Memory Test Item Across Suspect Photographs

<table>
<thead>
<tr>
<th>Test Items</th>
<th>Likely</th>
<th>Unlikely</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hair color</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2 Hair style</td>
<td>9</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>3 Face shape</td>
<td>9</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>4 Eyebrow color</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>5 Lipstick</td>
<td>8</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>6 Shirt color</td>
<td>12</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>7 Pants/skirt type</td>
<td>9</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>8 Pants/skirt color</td>
<td>13</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>9 Shirt type</td>
<td>11</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>10 Face expression</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>11 Height</td>
<td>8</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>12 Weight</td>
<td>6</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>13 Bodybuild</td>
<td>13</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>14 Age</td>
<td>8</td>
<td>12</td>
<td>6</td>
</tr>
</tbody>
</table>

Note: Maximum number of correct responses for each item = 16
Table 1

Experiment Design for Chapter 2

<table>
<thead>
<tr>
<th>Subject Gender</th>
<th>Method used to study photos</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>Describe photos out loud</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silently study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>Describe photos out loud</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Silently study</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The letters A and B represent photographs of men. The letters C and D represent photographs of women.*
Table 2
Assignment of Subjects and Counterbalancing of Materials

<table>
<thead>
<tr>
<th>Subject Gender</th>
<th>Study Task</th>
<th>Order of Study for Photographs</th>
<th>Subject Gender</th>
<th>Study Task</th>
<th>Order of Study for Photographs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Women Subjects</td>
<td>Describe</td>
<td>1 2</td>
<td>3 4</td>
<td>1 2</td>
<td>3 4</td>
</tr>
<tr>
<td></td>
<td>out loud</td>
<td>2 1</td>
<td>4 3</td>
<td>2 1</td>
<td>4 3</td>
</tr>
<tr>
<td></td>
<td>Silent Study</td>
<td>4 3</td>
<td>2 1</td>
<td>4 3</td>
<td>2 1</td>
</tr>
<tr>
<td>Men Subjects</td>
<td>Describe</td>
<td>3 4</td>
<td>1 2</td>
<td>3 4</td>
<td>1 2</td>
</tr>
<tr>
<td></td>
<td>out loud</td>
<td>2 1</td>
<td>4 3</td>
<td>2 1</td>
<td>4 3</td>
</tr>
<tr>
<td></td>
<td>Silent Study</td>
<td>4 3</td>
<td>2 1</td>
<td>4 3</td>
<td>2 1</td>
</tr>
</tbody>
</table>
### Table 3
Order of Tasks

<table>
<thead>
<tr>
<th>Order #</th>
<th>Tasks</th>
<th>Time (mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Consent</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Study Phase Instructions</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Study Phase Practice</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Study Phase</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>NAART</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Picture Differences</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>Test Phase</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>Debriefing</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 4

Summary of Floor and Ceiling Effects for Study Phase and Test Phase Descriptions

<table>
<thead>
<tr>
<th>Floor Effects</th>
<th>Ceiling Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study Phase</strong></td>
<td><strong>Test Phase</strong></td>
</tr>
<tr>
<td>Assumption/Evaluation</td>
<td>Assumption/Evaluation</td>
</tr>
<tr>
<td>Action State</td>
<td>Action State</td>
</tr>
<tr>
<td>Facial Expressions</td>
<td>Facial Hair</td>
</tr>
<tr>
<td>Bodyparts</td>
<td>Facial Expressions</td>
</tr>
<tr>
<td>Bodyparts Descriptors</td>
<td>Bodyparts</td>
</tr>
<tr>
<td>Body &amp; B.-Part</td>
<td>Bodyparts Descriptors</td>
</tr>
<tr>
<td>Posture</td>
<td></td>
</tr>
<tr>
<td>Jewelry/Accessory</td>
<td>Body &amp; B.-Part</td>
</tr>
<tr>
<td>Descr</td>
<td>Posture</td>
</tr>
<tr>
<td>Absent</td>
<td>Jewelry/Accessory</td>
</tr>
<tr>
<td>Clothes/Jewelry</td>
<td>Descriptors</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
</tr>
<tr>
<td></td>
<td>Clothes/Jewelry</td>
</tr>
</tbody>
</table>

Note. The attributes age, hair, eyes/eyebrows, ethnicity, gender and name all have a limit to the number of occurrences that can be scored as defined in the scoring key.
Table 5

Counterbalancing of Information Types Across Groups of Items

<table>
<thead>
<tr>
<th>Description</th>
<th>Group</th>
<th>Group</th>
<th>Group</th>
<th>Description</th>
<th>Group</th>
<th>Group</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td># 1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td># 1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>A1°</td>
<td>U°</td>
<td>L°</td>
<td>N°</td>
<td>B1°</td>
<td>N</td>
<td>U</td>
<td>L</td>
</tr>
<tr>
<td>A2</td>
<td>U</td>
<td>N</td>
<td>L</td>
<td>B2</td>
<td>N</td>
<td>L</td>
<td>U</td>
</tr>
<tr>
<td>A3</td>
<td>L</td>
<td>U</td>
<td>N</td>
<td>B3</td>
<td>U</td>
<td>N</td>
<td>L</td>
</tr>
<tr>
<td>A4</td>
<td>L</td>
<td>N</td>
<td>U</td>
<td>B4</td>
<td>U</td>
<td>L</td>
<td>N</td>
</tr>
<tr>
<td>A5</td>
<td>N</td>
<td>U</td>
<td>L</td>
<td>B5</td>
<td>L</td>
<td>N</td>
<td>U</td>
</tr>
<tr>
<td>A6</td>
<td>N</td>
<td>L</td>
<td>U</td>
<td>B6</td>
<td>L</td>
<td>U</td>
<td>N</td>
</tr>
</tbody>
</table>

°Groups 1 – 3 are groups of items from the likelihood questionnaire.

°The codes A1 – A6 indicate Suspect A and the number of the description for Suspect A.

°U = unlikely information

°L = likely information

°N = neutral information

°The codes B1 – B6 indicate Suspect B and the number of the description for Suspect B.
Table 6

Subject Assignment and Presentation Sequence of Photographs and Written Descriptions

<table>
<thead>
<tr>
<th>Subject #</th>
<th>First</th>
<th>Second</th>
<th>Suspect A (Mel)</th>
<th>Suspect B (Lee)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>B&lt;sup&gt;b&lt;/sup&gt;</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>B</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>B</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>6</td>
<td>A</td>
<td>B</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>7</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>9</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>10</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>11</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>12</td>
<td>B</td>
<td>A</td>
<td>A&lt;sup&gt;c&lt;/sup&gt;</td>
<td>B&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

12 x 2 = 24 subjects total

<sup>a</sup>A = photograph of Suspect A
<sup>b</sup>B = photograph of Suspect B
<sup>c</sup>A1-A6 = descriptions of Suspect A and B1 – B6 = descriptions of Suspect B
Figure Captions

Figure 1. Subjects' study phase descriptions of men and women. Error bars represent 95% confidence intervals.

Figure 2. Subjects' test phase descriptions of men and women. Error bars represent 95% confidence intervals.

Figure 3. Subjects' study phase descriptions of female suspects. Error bars represent 95% confidence intervals.

Figure 4. Subjects' study phase descriptions of male suspects. Error bars represent 95% confidence intervals.

Figure 5. Subjects' test phase descriptions of female suspects. Error bars represent 95% confidence intervals.

Figure 6. Subjects' test phase descriptions of male suspects. Error bars represent 95% confidence intervals.

Figure 7. Losses of information between study and test descriptions. Mean losses indicate the mean differences between subjects' study phase scores and their matches between study and test. The error bars represent 95% confidence intervals.

Figure 8. This figure shows the absolute mean frequencies of attributes recalled by subjects who described photos out loud or silently studied them in the study session. The method of studying photos did not influence recall. Error bars indicate 95% confidence intervals.

Figure 9. The likelihood scale ranges from very likely to very unlikely. Likely and unlikely information lie between the extreme ends of the scale.
Figure 1: Subjects' Study Phase Descriptions of Men and Women

- Photos of Women
- Photos of Men

Women Subjects N = 24
Men Subjects N = 10

Subjects & Photographs
Figure 2: Subjects' Test Phase Descriptions of Men and Women

<table>
<thead>
<tr>
<th></th>
<th>Women Subjects N = 48</th>
<th>Men Subjects N = 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photos of Women</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photos of Men</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subjects & Photographs
Figure 3: Subjects' Study Phase Descriptions of Female Suspects

- Women Subjects N = 24
- Men Subjects N = 10

Suspect Attributes:

- Hair Descriptors
- Eyes/Eyebrows Desc.
- Facial Parts
- Facial Parts Descriptors
- Bodybuild
- Clothes
- Clothes Desc. Non-Color
- Clothes Desc. Color
- Jewelry/Accessories
- Object Position & Art.
- Surroundings
- Surroundings Desc.

Absolute Mean Frequency

[Graph showing the frequency of descriptions for different attributes among women and men subjects.]
Figure 4: Subjects' Study Phase Descriptions of Male Suspects

<table>
<thead>
<tr>
<th>Suspect Attributes</th>
<th>Women Subjects N = 24</th>
<th>Men Subjects N = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hair Descriptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Hair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes/Eyebrows Desc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Parts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Parts Descriptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bodybuild</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes Desc. Non-Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes Desc. Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewelry/Accessories</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object position &amp; Arr.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surroundings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surroundings Desc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Absolute Mean Frequencies
Figure 5: Subjects Test Phase Descriptions of Female Suspects

- Women Subjects N = 48
- Men Subjects N = 18

<table>
<thead>
<tr>
<th>Suspect Attributes</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hair Descriptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes/Eyebrows Desc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Parts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Parts Descriptors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bodybuild</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes Des. Non-Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothes Desc. Color</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jewelry/Accessories</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Object position &amp; Arr.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surroundings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surroundings Desc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6: Subjects’ Test Phase Descriptions of Male Suspects

- Women Subjects N = 48
- Men Subjects N = 18

Suspect Attributes

Absolute Mean Frequency

Hair Descriptors
Eyes/Eyebrows Desc.
Facial Parts
Facial Parts Descriptors
Bodybuild
Clothes
Clothes Des. Non-Color
Clothes Des. Color
Jewelry/Accessories
Object position & Arr.
Surroundings
Surroundings Desc.
Figure 7: Mean Losses Between Study & Test

Subject Gender & Suspect Gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>Suspect Gender</th>
<th>Mean Differences (Study Minus Matches)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women Subjects</td>
<td>Female Suspects</td>
<td>1.5 (±0.5)</td>
</tr>
<tr>
<td></td>
<td>Male Suspects</td>
<td>1.5 (±0.5)</td>
</tr>
<tr>
<td>Men Subjects</td>
<td>Female Suspects</td>
<td>1.0 (±0.5)</td>
</tr>
<tr>
<td></td>
<td>Male Suspects</td>
<td>1.5 (±0.5)</td>
</tr>
</tbody>
</table>
Figure 8: Influence of Describing Out Loud and Silent Study on Recollections of Suspects

- Describe out loud \( N = 136 \)
- Silent Study \( N = 128 \)

Suspect Attributes:
- Hair Descriptors
- Eyes/Eyebrows Desc.
- Facial Parts
- Facial Parts Descriptors
- Bodybuild
- Clothes
- Clothes Des.Nonn-Color
- Clothes Desc. Color
- Jewelry/Accessories
- Object position & Arr.
- Surroundings
- Surroundings Desc.