The effects of sequential and simultaneous lineup procedures on confidence and accuracy of identity judgments

Julia Kheng Mei Chan

Submitted in partial fulfilment of the requirements of the degree of Doctor of Psychology (Forensic)

August 2017

School of Psychology
Faculty of Health
Deakin University
Australia

Produced on archival quality paper
I am the author of the thesis entitled

*The effects of sequential and simultaneous lineup procedures on confidence and accuracy of identity judgments*

submitted for the degree of Doctor of Psychology (Forensic)

This thesis may be made available for consultation, loan and limited copying in accordance with the Copyright Act 1968.

'I certify that I am the student named below and that the information provided in the form is correct'

**Full Name:** Julia Kheng Mei Chan

**Signed:** Signature Redacted by Library

**Date:** 22.02.2018
Declaration

DEAKIN UNIVERSITY

CANDIDATE DECLARATION

I certify the following about the thesis entitled:

The effect of sequential and simultaneous lineup procedures on confidence and accuracy of identity judgments

submitted as the research thesis for the degree of Doctor of Psychology (Forensic)

a) I am the creator of all or part of the whole work(s) (including content and layout) and that where reference is made to the work of others, due acknowledgment is given.

b) The work(s) are not in any way a violation or infringement of any copyright, trademark, patent, or other rights whatsoever of any person.

c) That if the work(s) have been commissioned, sponsored or supported by any organisation, I have fulfilled all of the obligations required by such contract or agreement.

d) That any material in the thesis which has been accepted for a degree or diploma by any university or institution is identified in the text.

e) All research integrity requirements have been complied with.

'I certify that I am the student named below and that the information provided in the form is correct'

Full Name: Julia Kheng Mei Chan

Signed: [Signature Redacted by Library]

Date: 2/08/2017
Acknowledgements

I would like to thank my primary supervisor Professor Don Thomson for all of his hard work and steady guidance throughout the completion of this research project. You have provided me with invaluable wisdom and infinite knowledge on this area of research throughout my Doctorate. I could not have completed this without you.

I would also like to thank my secondary supervisor Dr. Adrian Scott for his patience and help throughout data collection and in the later stages of my thesis write up. Your help with Qualtrics and the more practical aspects of this project has been amazing. Thank you for sticking with me until the end.

Also, thank you to Val Clarke for giving me precise and helpful advice in the final stages of writing up my thesis. Thank you for putting up with all of the things that I did not know and picking up all of my bad writing habits. You were consistent and specific with me, and you persisted to help me finish off this piece of work. Thank you so much for your perseverance and dependability! I will always be endlessly appreciative.

Finally, thank you to the Deakin School of Psychology, particularly Professor Greg Tooley, Lina Ricciardelli and Professor Jane McGillvray. You all provided me with much direction, particularly throughout the final stages of my thesis. It was amazing to receive support from so many of the faculty throughout the trials and tribulations that always litter the landscape of completing a doctoral thesis. Thank you for your unwavering guidance, assistance and encouragement through it all.

Julia Kheng Mei Chan
### Table of Contents

Access to Thesis ......................................................................................................................... i
Declaration ................................................................................................................................. iii
Acknowledgements ..................................................................................................................... v
Table of Contents ....................................................................................................................... vi
List of Tables ............................................................................................................................... xi
List of Figures .............................................................................................................................. xv
List of Abbreviations .................................................................................................................. xvi
Abstract ..................................................................................................................................... xvii

**CHAPTER 1 - The Problem of Eyewitness Misidentification** ......................... 1

1.1 Aims ................................................................................................................................. 1
1.2 Overview of the Thesis .................................................................................................... 1
1.3 The role of eyewitness identification evidence in the criminal justice system ................. 2
1.4 Eyewitness misidentification and wrongful convictions .............................................. 3
   1.4.1 Cases in the USA ........................................................................................................ 4
   1.4.2 Cases in the UK ......................................................................................................... 6
   1.4.3 Cases in Australia ...................................................................................................... 7
   1.4.4 Eyewitness identification evidence and interactions with other factors .................. 8
1.5 The process of eyewitness identification - the identification parade or lineup .......... 8
1.6 Legal rulings governing the process of identification lineups ........................................ 10
   1.6.1 Live versus photographic lineups ............................................................................ 11
   1.6.2 Number of foils ....................................................................................................... 11
   1.6.3 Type of Lineup ......................................................................................................... 12
1.7 Legal safeguards governing the use of eyewitness identification evidence in court cases . 12
   1.7.1 Case law in the USA ............................................................................................... 13
   1.7.2 Case law in the UK ................................................................................................. 14
   1.7.3 Case law in Australia .............................................................................................. 15
### Table of Contents

- 1.7.4 Identity versus similarity evidence ........................................... 16
- 1.8 Summary .......................................................................................... 17

**CHAPTER 2 - Factors Influencing Eyewitness Identification: The Psychological Literature ................................................................. 19**

- 2.1 Eyewitness memory ........................................................................... 19
- 2.2 System variables ............................................................................... 20
  - 2.2.1 Selection of foils ........................................................................ 21
  - 2.2.2 Number of foils ......................................................................... 24
  - 2.2.3 Instructions given to witnesses .................................................. 25
  - 2.2.4 Type of lineup .......................................................................... 26
    - 2.2.4.1 Explanations for the ‘sequential superiority’ effect .......... 29
- 2.2.5 Position of the suspect in the lineup ............................................. 37
- 2.3 Estimator variables .......................................................................... 38
  - 2.3.1 Duration of observation ............................................................. 39
  - 2.3.2 Obstacles to viewing, distance from the offender, and lighting conditions ................................................................................. 39
  - 2.3.3 Familiarity and visual similarity ................................................ 40
  - 2.3.4 Length of time between the offence and identification .......... 41
  - 2.3.5 Consistency between the description of the perpetrator and the appearance of the accused ......................................................... 43
  - 2.3.6 Eyewitness attention ................................................................... 44
- 2.4 Eyewitness confidence ...................................................................... 45
- 2.5 Legal safeguards: Are they supported by research? ....................... 50
  - 2.5.1 The distinction between identification and similarity evidence ............................................................................................................. 50
  - 2.5.2 System variables ........................................................................ 53
  - 2.5.3 Estimator variables ..................................................................... 54
    - 2.5.3.1 Duration of observation ....................................................... 55
    - 2.5.3.2 Obstacles to viewing, distance from the offender, and lighting conditions ......................................................................................... 55
    - 2.5.3.3 Length of time between the offence and identification ....... 56
    - 2.5.3.4 Consistency between the description of the perpetrator and the appearance of the accused ......................................................... 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.3.5 Eyewitness attention</td>
<td>57</td>
</tr>
<tr>
<td>2.5.4 Eyewitness confidence</td>
<td>58</td>
</tr>
<tr>
<td>2.5.4.1 Eyewitness confidence and jury decision making</td>
<td>61</td>
</tr>
<tr>
<td>2.6 Summary</td>
<td>62</td>
</tr>
<tr>
<td>CHAPTER 3 - Experiment 1: Target Present Lineups</td>
<td>64</td>
</tr>
<tr>
<td>3.1 Method</td>
<td>66</td>
</tr>
<tr>
<td>3.1.1 Participants</td>
<td>67</td>
</tr>
<tr>
<td>3.1.2 Design</td>
<td>67</td>
</tr>
<tr>
<td>3.1.3 Procedures</td>
<td>67</td>
</tr>
<tr>
<td>3.1.4 Materials and Apparatus</td>
<td>69</td>
</tr>
<tr>
<td>3.1.4.1 Photographs</td>
<td>69</td>
</tr>
<tr>
<td>3.1.4.2 Lineup composition</td>
<td>69</td>
</tr>
<tr>
<td>3.1.5 Data reduction &amp; statistical analyses relevant to the entire thesis</td>
<td>71</td>
</tr>
<tr>
<td>3.1.6 Data reduction &amp; statistical analyses for Experiment 1</td>
<td>71</td>
</tr>
<tr>
<td>3.1.6.1 Accuracy</td>
<td>71</td>
</tr>
<tr>
<td>3.1.6.2 Confidence</td>
<td>72</td>
</tr>
<tr>
<td>3.2 Results and discussion</td>
<td>73</td>
</tr>
<tr>
<td>3.2.1 Accuracy</td>
<td>73</td>
</tr>
<tr>
<td>3.2.1.1 Lineup procedure: Hits</td>
<td>73</td>
</tr>
<tr>
<td>3.2.1.2 Lineup procedure: False positives</td>
<td>78</td>
</tr>
<tr>
<td>3.2.1.3 Photograph condition (same versus different)</td>
<td>83</td>
</tr>
<tr>
<td>3.2.2 Confidence</td>
<td>85</td>
</tr>
<tr>
<td>3.2.2.1 Lineup procedure</td>
<td>85</td>
</tr>
<tr>
<td>3.2.2.2 Photograph</td>
<td>91</td>
</tr>
<tr>
<td>3.2.2.3 Target position</td>
<td>93</td>
</tr>
<tr>
<td>3.2.2.4 Confidence in correct versus false identifications</td>
<td>94</td>
</tr>
<tr>
<td>3.3 Summary</td>
<td>100</td>
</tr>
<tr>
<td>CHAPTER 4 - Experiment 2: Target Present and Absent</td>
<td>103</td>
</tr>
<tr>
<td>4.1 Method</td>
<td>105</td>
</tr>
<tr>
<td>4.1.1 Participants</td>
<td>105</td>
</tr>
<tr>
<td>4.1.2 Design</td>
<td>106</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Procedures</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Materials and apparatus</td>
</tr>
<tr>
<td>4.1.4.1</td>
<td>Photographs</td>
</tr>
<tr>
<td>4.1.4.2</td>
<td>Lineup composition</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Data reduction &amp; statistical analyses</td>
</tr>
<tr>
<td>4.1.5.1</td>
<td>Accuracy measures</td>
</tr>
<tr>
<td>4.1.5.2</td>
<td>Confidence ratings</td>
</tr>
<tr>
<td>4.2</td>
<td>Results and discussion</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Accuracy</td>
</tr>
<tr>
<td>4.2.1.1</td>
<td>Lineup procedure</td>
</tr>
<tr>
<td>4.2.1.2</td>
<td>Photograph condition (same versus different)</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Confidence</td>
</tr>
<tr>
<td>4.2.2.1</td>
<td>Lineup procedure</td>
</tr>
<tr>
<td>4.2.2.2</td>
<td>Photograph condition (same versus different)</td>
</tr>
<tr>
<td>4.2.2.3</td>
<td>Confidence for hit versus false positive responses</td>
</tr>
<tr>
<td>4.2.2.4</td>
<td>Confidence for miss versus correct rejection responses</td>
</tr>
<tr>
<td>4.3</td>
<td>Summary</td>
</tr>
</tbody>
</table>

**CHAPTER 5 - Experiment 3: Visual Similarity**

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Method</td>
<td>133</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Participants</td>
<td>133</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Design</td>
<td>133</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Procedures</td>
<td>134</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Materials and apparatus</td>
<td>134</td>
</tr>
<tr>
<td>5.1.4.1</td>
<td>Photographs</td>
<td>134</td>
</tr>
<tr>
<td>5.1.4.2</td>
<td>Lineup composition</td>
<td>135</td>
</tr>
<tr>
<td>5.1.5</td>
<td>Data reduction &amp; statistical analyses</td>
<td>136</td>
</tr>
<tr>
<td>5.1.5.1</td>
<td>Accuracy measures</td>
<td>136</td>
</tr>
<tr>
<td>5.1.5.2</td>
<td>Confidence ratings</td>
<td>137</td>
</tr>
<tr>
<td>5.2</td>
<td>Results and discussion</td>
<td>138</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Accuracy</td>
<td>138</td>
</tr>
<tr>
<td>5.2.1.1</td>
<td>Lineup procedure</td>
<td>140</td>
</tr>
<tr>
<td>5.2.1.2</td>
<td>Visual similarity of photographs and target position</td>
<td>143</td>
</tr>
<tr>
<td>5.2.1.3</td>
<td>Target replacement position</td>
<td>147</td>
</tr>
</tbody>
</table>
5.2.2 Confidence.................................................................148
  5.2.2.1 Lineup procedure..................................................150
  5.2.2.2 Visual similarity and target/target replacement position.......154
  5.2.2.3 Confidence in accurate versus inaccurate responses ..........154

5.3 Summary .........................................................................158

CHAPTER 6 - General Discussion ...........................................160

  6.1 Impact of lineup procedure on accuracy and confidence:
       Integration of findings, theoretical implications and practical
       application .........................................................................161

  6.2 The impact of visual similarity between the originally
       observed target and target in the lineup: Integration of
       findings, theoretical implications and practical application....
       .......................................................................................176

  6.3 The relationship between confidence and accuracy:
       Integration of findings, theoretical implications and practical
       application .........................................................................180

  6.4 Strengths, limitations and future directions.........................186

  6.5 Conclusions ....................................................................191

REFERENCES ...........................................................................193

APPENDIX A: Advertisement, PLSs and Consent Forms..............211

APPENDIX B: Ethics Approvals.................................................213
  Ethics Approval: Experiments 1 and 2....................................213
  Ethics Approval: Experiment 3..............................................215

APPENDIX C: Example of the Online Survey............................217
List of Tables

Table 3.1  Mean (SD) number of responses and sample sizes for hit, miss and false positive responses for sequential and simultaneous lineups in the same or different photograph condition for early and late target presentations. ........................................................................................................... 74

Table 3.2  Mean (SD) confidence levels for hit responses for sequential and simultaneous lineups in the same or different photograph condition for early and late target. ............................................................................................................... 86

Table 3.3  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had miss responses in the same and different photograph condition. ................................................................. 87

Table 3.4  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had miss responses for early and late target presentations. ......................................................................................... 87

Table 3.5  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who falsely identified a lineup foil in the same and different photograph condition. ......................................................... 88

Table 3.6  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who falsely identified a lineup foil for early and late target presentations. ................................................................. 89

Table 3.7  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had both hit and false positive responses for both same and different photograph conditions. .............................................. 94

Table 3.8  Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had both hit and false positive responses when the target was presented early and when it was presented late. .......... 96
Table 3.9  Mean (SD) confidence levels and sample sizes (n) for hit and false positive responses (FPs) in sequential and simultaneous lineups when the target was presented early and when it was presented late in same and different photograph conditions. ..........................................................98

Table 4.1  Participants were randomly allocated to conditions. They either participated in simultaneous or sequential lineups. Black font corresponds to target present lineups, whereas grey font corresponds to target absent lineups. Note, that the variables of photograph (same vs different photographs of the target or same vs different backgrounds behind the target replacement) and target position (early presentation at position 2 or 3 or late position at position 6 or 7), are perfectly balanced for each participant, such that they represent between-subjects factors when considering target present and target absent lineups in isolation. NB: TP= target present, TA= target absent. .....110

Table 4.2  Sample sizes for hit, miss, correct rejection and false positive confidence ratings for target absent and present sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: N/A refers to conditions where the relevant response type was not possible. .................................................112

Table 4.3  Mean (SD) numbers of responses and sample sizes for hits, misses, and false positive responses in lineups where the target was present for sequential and simultaneous lineups in the same or different photograph condition for early and late target presentations. .................................113

Table 4.4  Mean (SD) numbers of responses and sample sizes for correct rejection and false positive responses where either a lineup foil or target replacement was selected in lineups where the target was absent for sequential and simultaneous lineups in the same or different photograph condition for early and late target replacement presentations. ...............114
**Table 4.5** Mean (SD) confidence levels for hit, miss and false positive responses when the target was present in sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: n/a refers to conditions where there was no standard deviation due to a sample size of 1; Nil refers to conditions in which no responses were given so no confidence ratings existed. .......................... 119

**Table 4.6** Mean (SD) confidence levels for correct rejection and false positive responses where either the target replacement or a different lineup foil was chosen and when the target was absent in sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: Nil refers to conditions in which no responses were given so no confidence ratings existed...... 120

**Table 4.7** Mean (SD) confidence levels for hit and false positive responses in simultaneous and sequential lineups where the target or target replacement was presented early or late in same or different photograph conditions. ................................................................. 123

**Table 4.8** Mean (SD) confidence levels for miss and correct rejection responses in simultaneous and sequential lineups where the target replacement was presented early or late in same or different photograph conditions. NB: n/a refers to conditions where there was no standard deviation due to a sample size of 1. ......................................................................................... 126

**Table 5.1** Mean (SD) numbers of responses and sample sizes for hits, misses, and false positive responses in lineups where the target was present for sequential and simultaneous lineups in highly similar (SS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late target presentations................................................................. 139
Table 5.2  Mean (SD) frequencies and sample sizes for correct rejection and false positive responses where either a lineup foil or target replacement was selected in lineups where the target was absent for sequential and simultaneous lineups for early and late target replacement presentations..................................................140

Table 5.3  Sample sizes for hit, miss, correct rejection and false positive confidence ratings for target absent and present sequential and simultaneous lineups in highly similar (SS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late photograph presentations ..................................................149

Table 5.4  Mean (SD) confidence levels for hit, miss and false positive responses when the target was present in sequential and simultaneous lineups in highly similar (SS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late photograph presentations. .150

Table 5.5  Mean (SD) confidence levels in for correct rejection and false positive responses where either the target replacement or a different lineup foil was chosen and when the target was absent in sequential and simultaneous lineups for early and late photograph presentations.....151

Table 5.6  Mean (SD) confidence levels for miss and correct rejection responses in simultaneous and sequential lineups where the target replacement was presented early or late in the lineup. ..................................................156
List of Figures

Figure 3.1. Example of photograph displays for simultaneous (top pane) and sequential (bottom pane) identification decisions. Photographs taken from Gross and colleagues (2010). ................................................................. 70

Figure 4.1. Examples of photographic lineups including target photographs (left-most column), for the same (top two panes) and different (bottom two panes) photograph conditions when the target was present (first and third panes) and absent (second and fourth panes). Different photographs contain a different background of the target person....... 108

Figure 5.1. Example of a lineup photograph with high backed chair background as was used for all lineup photographs in experiment four. Taken from Gross (2010). ................................................................. 135
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>dpi</td>
<td>Dots per inch</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>HS</td>
<td>Highly similar</td>
</tr>
<tr>
<td>LS</td>
<td>Least similar</td>
</tr>
<tr>
<td>M</td>
<td>Mean</td>
</tr>
<tr>
<td>MS</td>
<td>Moderately similar</td>
</tr>
<tr>
<td>n</td>
<td>Sample size</td>
</tr>
<tr>
<td>NSW</td>
<td>New South Wales</td>
</tr>
<tr>
<td>r</td>
<td>Effect size</td>
</tr>
<tr>
<td>R</td>
<td>Regina</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>TP</td>
<td>Target present</td>
</tr>
<tr>
<td>TA</td>
<td>Target absent</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>v/vs</td>
<td>Versus</td>
</tr>
</tbody>
</table>
Abstract

Eyewitness misidentification is a significant problem in the criminal justice system, and has been identified as a leading factor in wrongful convictions in various countries throughout the world. A number of variables, both controlled by the criminal justice system (system variables) and not under control of the criminal justice system (estimator variables) are known to affect the accuracy of identifications. In particular, lineup procedure is suggested to affect accuracy with many researchers claiming that sequential lineups are superior to simultaneous lineups. However, this view is contentious and some researchers suggest that the superiority of sequential lineups is limited only to false identifications, whereas simultaneous lineups are superior in terms of correct identifications. As such, there is a current need to clarify the pattern of accuracy, in terms of hits, false positives, misses and correct rejections under simultaneous and sequential conditions in order to evaluate current theoretical explanations and the practical benefits of these different lineup procedures. Further, the relationship between witnesses’ confidence in their identification and the accuracy of their identification is controversial with courts in the USA asserting that eyewitness confidence can be used as a predictor of identification accuracy but courts in the UK warning jurors that confidence is a poor predictor. Similarly, findings of research examining the relationship of confidence and identification accuracy are mixed. Little research has examined the way in which different lineup procedures may affect eyewitness confidence and its relationship with patterns of accuracy. Therefore the current thesis had two primary aims: 1) to investigate patterns of accuracy under different lineup procedures; and 2) to examine the relationship between this pattern of accuracy and confidence.

Three experiments were conducted. The first experiment (N= 546) investigated patterns of accuracy and confidence in simultaneous and sequential lineups when a target person was present in the lineup. The second experiment (N= 297) investigated patterns of accuracy and confidence in both target present and target absent simultaneous and sequential lineups. The final experiment (N= 138) investigated accuracy and confidence when visual similarity between the originally observed target and target observed in the lineup was manipulated.
Participants in all experiments of the current investigation were volunteers aged 18 years and over who participated in an online survey and completed several lineup identification tasks (eight, four, and 12 lineups for the three experiments respectively). For all lineups, participants viewed a target photograph for five seconds, then viewed a distractor photograph of a person and were asked to estimate the age of the distractor person. Participants viewed a lineup and were asked to indicate whether or not they could identify the target person in that lineup. Subjective confidence ratings on a 5-point Likert scale for all identification decisions were obtained. For all experiments, lineup procedure was manipulated as a between-subjects variable, and different photograph conditions were viewed by all participants. In Experiments 1 and 2 in target present lineups, either the same or a different photograph of the target person as was originally viewed was presented in the lineup. For target absent lineups, the background of the target replacement was either the same as, or different from, the background originally presented behind the target person. In the final experiment, photograph similarity between the originally observed target and target presented in the lineup for target present lineups was varied across three levels: highly similar, moderately similar and least similar. Early versus late target or target replacement position in the lineup was also included as an independent variable in all experiments. Accuracy, as measured by the number of correct identifications, false identifications, incorrect rejections and correct rejections of the lineup as well as participant reported confidence in all lineup decisions were dependent variables. The relationship between confidence and accuracy was assessed by comparing levels of confidence of correct and false identifications, and levels of confidence of correct and incorrect rejections of the lineup.

Consistently, across the three experiments, there were more correct identifications, less false identifications and no differences in the number of incorrect rejections of the lineup in simultaneous compared with sequential lineups. In addition, there was a main effect of photograph similarity across the three experiments for target present lineups. The higher the visual similarity between the target person originally observed and the target person in the lineup, the more correct identifications. A significant effect of photograph similarity for both simultaneous and sequential lineups where the target was absent was also observed in the second experiment whereby viewing a different background behind the target replacement as was originally presented behind the target person increased correct rejections of the
The comparison of confidence of correct versus false identifications consistently indicated that confidence for correct identifications was higher on average than that of false identifications. This finding was not, however, observed for correct versus incorrect rejections of the lineup. Comparison of patterns of confidence in the different response types suggested that the relationship between confidence and accuracy was more reliable under some conditions than others. Results indicated that confidence in false identifications was higher in sequential than simultaneous lineups, particularly when a target replacement was identified and that more false identifications occurred in sequential than simultaneous lineups in general.

The findings of the three experiments are inconsistent with previous findings and with current explanations of the way eyewitness lineup judgments are made. There are two major explanations accounting for the sequential superiority effect. First, it has been argued that the simultaneous lineup procedure encourages individuals to make relative judgments rather than absolute judgments. Second, it has been argued that sequential lineups encourage witnesses to be more conservative, adopting a higher threshold for choosing. The pattern of current results indicated no evidence of a sequential superiority effect and were not consistent with either a more conservative threshold or more frequent use of absolute judgments in sequential lineups. The present findings are novel within the literature and suggest that alternative theoretical explanations for the process underlying eyewitness judgments should be considered. For example, identification decisions may be highly based on a judgment of similarity, and comparison between lineup members may occur across both simultaneous and sequential lineups, rather than only in simultaneous lineups as has been previously suggested. Furthermore, the present experiments provide evidence that while a person who is confident is perhaps more likely to be accurate, individual variation dictates that accuracy cannot be predicted reliably from confidence. These results support the view that confidence should not be presumed to be a reliable indicator of accuracy when considering eyewitness identification evidence in criminal trials.
CHAPTER 1
The Problem of Eyewitness Misidentification

1.1 Aims
The present thesis had two primary aims: to clarify patterns of accuracy across different eyewitness lineup conditions; and to characterise how patterns of confidence in lineup judgments related to patterns of accuracy of those judgments. These aims are pursued due to many criminal cases and much empirical psychological research demonstrating that eyewitness misidentification is a serious and systemic problem within the criminal justice system that has the devastating consequence of wrongful convictions. Indeed, many wrongful convictions have occurred in the United States of America (USA; Huff, 2004), the United Kingdom (UK; Lutz, Lutz, & Ulmschneider, 2002), and Australia (Colvin, 2009). The single biggest contributor to these convictions is mistaken eyewitness identification (Colvin, 2009; Huff, 2004).

1.2 Overview of the Thesis
The introduction to the current thesis will begin by considering the role of eyewitness identification evidence in the criminal justice system and how this relates to the process of criminal conviction. The term eyewitness will be defined as a witness who observes the perpetrator of a crime but who is not the victim of the crime. The aim of Chapter 1 is to draw on legal cases and statutes to demonstrate the extent to which misidentification of an alleged suspect can lead to a wrongful conviction. Eyewitness identification parades, or lineups, play a crucial role in the eyewitness identification process within the criminal justice system. Therefore, the second aim of this chapter is to describe the process by which eyewitness identification occurs, which is relevant to the focus of the current research on lineup procedures and conditions. Thus, Chapter 1 will begin with a discussion of the powerful role of eyewitness identification evidence within the criminal justice system. It will then move to a discussion of notable cases of wrongful conviction in the USA, the UK, and Australia and the ways in which several variables may interact with eyewitness identification to produce wrongful convictions within the criminal justice system. The examination of the cases in the USA, the UK, and Australia is intended to give a comprehensive overview of the common rulings within major
western legal systems, which is relevant to the current research which aims to examine the support for these rulings in psychological research. Finally, this chapter will shift to a discussion of the legal process and safeguards involved in eyewitness identification, with particular focus on those processes within the region in which the current research was conducted, namely Australia. This discussion focuses on the ways in which the process of eyewitness identification may impact on eyewitness accuracy, which pertains directly to the issues investigated by the current research.

The focus of Chapter 2 will be an examination of the variables in the psychological literature which have been found to influence eyewitness identification accuracy, including the type of lineup. Legal rulings and guidelines where no relevant statute exists, are practicably based in common law, or on the decisions of judges in response to cases being contested in a courtroom (Beazer, 2002). The weakness of this approach is that variables are not considered if they do not arise in a specific legal case, and the outcomes of legal cases are dependent on the views of single judges who draw from findings of previous judges. Thus, the extent to which legal safeguards and rules to prevent eyewitness misidentification are supported by the psychological literature will also be canvassed in Chapter 2. The two major explanations for the ‘sequential superiority effect’, a common finding within this psychological literature, will be discussed. These explanations posit that sequential lineups lead to superior accuracy either through a more stringent threshold for choosing, or through more frequent use of absolute judgments. Further, the relationship between confidence and identification accuracy, which has been studied through psychological experiments in response to legal rulings, will be discussed. The three experiments undertaken in the current investigation and the findings are described and discussed in Chapters 3, 4 and 5. Finally, the conclusions to be drawn from the findings of the three experiments are discussed in Chapter 6.

1.3 The role of eyewitness identification evidence in the criminal justice system

Eyewitness identification evidence has a unique status within the legal system. It is considered to be direct, rather than circumstantial evidence, meaning that it is thought to directly attest to the presence of the identified person at the crime scene, without inference or presumption, and if true, conclusively establishes that fact (Festa v R, 2001; Wells, 1984). Very few other forms of evidence are
considered direct evidence. Even scientific evidence, such as fingerprints or DNA, is considered circumstantial evidence and must be evaluated within the framework of the supporting contextual factors. As such eyewitness identification evidence is particularly powerful and persuasive evidence within the context of a criminal trial. The criminal investigative process reflects the importance placed on eyewitness identification evidence, and often revolves around obtaining such evidence. Indeed, when a crime occurs, a standard criminal investigation will begin by looking for witnesses to the event. Witnesses are interviewed and asked for a description of the perpetrator and the investigation often proceeds on the basis of this description. In Victoria, Australia, in the event that there is no specific suspect, for identification purposes witnesses are shown photograph books containing photographs of people who are deemed to resemble the perpetrator described by witnesses (Victoria Police, 1998). If no witnesses are immediately located, the events leading up to the crime are examined in detail often in order to identify whether any witnesses may exist, or to gather enough circumstantial evidence to identify a suspect or suspects. Thus, it would appear that eyewitness identification evidence can play a crucial role in the outcome of criminal trials.

1.4 Eyewitness misidentification and wrongful convictions

Perhaps the most salient demonstration of the impact of eyewitness identification evidence within the criminal justice system is the literature concerning wrongful convictions. Defining a ‘wrongful conviction’ is somewhat problematic, and there is no reliable method to determine the true incidence of wrongful conviction to date (Huff, 2004). Indeed, no objective evidence or self-report survey would be reliable under these circumstances and while some estimate the incidence of yearly wrongful convictions in the USA to be 0.5% (Huff, 2004), other estimates are closer to 5% (Rattner, 1988), or 25% in cases where no biological evidence is available (Huff, 2004). For these reasons, the term ‘wrongful conviction’ in the literature has been used to refer to those cases where an accused person has been convicted of a crime but was subsequently exonerated by evidence that was available, but not sufficiently utilised, at the time of conviction; new evidence that was not previously available exonerates the accused; the confession of the actual culprit exonerates the accused; or the conviction was overturned on appeal, or by some other legal body as an unjust or unfair conviction (Colvin, 2009; Huff, 2004;
Chapter 1 | The Problem of Eyewitness Misidentification

Rattner, 1988). For the purposes of the current research, this is how the term ‘wrongful conviction’ will be defined.

Analyses of the wrongful convictions of over 1000 people in the USA have led to the identification of eyewitness error as the single biggest factor leading to wrongful convictions (Wells, 1993). Consistent with this conclusion, in large cohort analyses of criminal cases, eyewitness misidentification has been found to be a leading factor in wrongful convictions in the UK (Walker & Starmer, 1999), Canada (MacFarlane, 2005), and the USA (Huff, Rattner, & Sagarin, 1996), and a significant factor in wrongful convictions in Australia (Colvin, 2009; Langdon & Wilson, 2005; Wilson, 1989). That there are so many wrongful convictions based wholly or partly on eyewitness misidentification would suggest that judges and jurors strongly believe in the accuracy of eyewitness identification evidence, despite evidence of its role in miscarriages of justice.

1.4.1 Cases in the USA

Much of the research on wrongful convictions has emanated from the USA. Anecdotal cases of wrongful executions in the USA date back to as early as the 1700s (MacNamara, 1969). Discussions of wrongful convictions and executions have heightened since a series of wrongful convictions in the mid-1900s (Borchard & Lutz, 1932; MacNamara, 1969). In particular, Edwin Borchard (1932) documented 65 cases of wrongful conviction in the early 1900s. Among these cases, many involved eyewitness misidentification, including one in which 17 eyewitnesses identified Herbert Andrewes as the man who wrote dozens of bad cheques, only for the true culprit to be found one year later (Borchard & Lutz, 1932). Indeed, in the conclusion to this work, Borchard noted that eyewitness misidentification was responsible practically alone for 29 of the 65 cases of wrongful conviction (Borchard & Lutz, 1932). Subsequently, many similar reviews have documented individual cases of wrongful convictions relating to eyewitness misidentification (Rattner, 1988). Upon examination of the literature relating to the wrongful convictions in the USA, most studies have found that the factor most often associated with wrongful conviction is eyewitness error (Colvin, 2009; Huff, 2004; Rattner, 1988; Wells, 1993). Huff, Rattner, Sagarin and MacNamara (2003) reported that of their database of 500 cases of wrongful conviction, 60% involved mistaken eyewitness identification. Similarly, Rattner (1988) studied 205 cases eventuating in wrongful
convictions, most of which involved serious criminal offences, and 38.5% for which the sentence was death or life imprisonment. Of these cases, eyewitness misidentification contributed to wrongful convictions in 48.8% of cases. Further, Bedau and Radelet (1987) conducted an analysis of 350 cases of wrongful conviction and found that witness error was involved in 193 of these cases, with sincere eyewitness misidentification being involved in 56 of these.

In the past three decades, the introduction of DNA testing in criminal cases has given rise to the ‘innocence movement’ in the USA (Leo & Gould, 2009). In 1989, David Vasquez, originally convicted of murder, was exonerated on the basis of DNA evidence, and shortly following this, Gary Dotson was also exonerated on the basis of DNA evidence after serving 10 years for a rape he did not commit. Since this time, post-conviction DNA testing has led to the release of over 300 additional wrongly convicted prisoners in the USA and eyewitness misidentification was found to be involved in as many as 75% of these cases (Leo & Gould, 2009; Wells & Olson, 2003). A consistent observation among the cases relating to wrongful convictions is that eyewitness misidentification is a significant contributor to miscarriages of justice.

However, despite these findings, USA case law has consistently afforded eyewitness identification evidence a powerful status within the legal system. Judges in the USA hold the primary responsibility for preventing and minimising the effects of potential eyewitness errors (Wise & Safer, 2004). For example, because judges have the power to exclude evidence, they can specify how lineups are conducted and if an accused has the right to have an attorney present during a lineup (United States v Ash, 1973; Kirby v. Illinois, 1972; United States v. Wade, 1967). Some instances of caution exercised by judges concerning eyewitness evidence can be noted. In New Jersey, for example, a report was prepared by a special master during a remand proceeding in the case of New Jersey v Hendersen (2011). This report comprehensively researched published literature and heard expert testimony with respect to eyewitness identification (Gaulkin, 2010). In 2011, based on the report, the court issued a direction requiring close examination of eyewitness identification evidence by trial courts in New Jersey. However, subsequently in the case of Perry v. New Hampshire (2012), the USA Supreme Court held that judicial examination of eyewitness testimony was only required in the case of police misconduct. Further, expert testimony on the reliability of eyewitness evidence is only sometimes allowed.
Chapter 1 | The Problem of Eyewitness Misidentification

and only in some jurisdictions. Additionally, instruction to the jury about the reliability of eyewitness evidence is generally prohibited in the USA. Indeed, it would seem that key players within the legal system, such as judges, place a great deal of faith in the utility of eyewitness evidence in the courtroom. Research on the knowledge and attitudes of USA judges would indicate that judges most often believe that eyewitness error occurs less often than the actual incidence, and that the average juror would already be aware of the issues relating to eyewitness identification (Wise & Safer, 2004). A frequent reason for the refusal of judges to hear expert testimony relating to eyewitness identification is the belief that the subject matter is already within the knowledge of jurors (United States v Hall, 1999; Wise & Safer, 2004). This refusal indicates the possibility that important figures in the legal system may overestimate juror knowledge about the issues relating to eyewitness identification evidence, which may eventuate in less preventative measures against miscarriages of justice being taken.

1.4.2 Cases in the UK

There are significant differences between the legal system in the USA and those in the UK. For example, unlike the USA system, in English-based systems, judges can and commonly do assess the strength of evidence for the jury, although the ultimate decision still rests with the jury (Colvin, 2009).

In the UK, miscarriages of justice gained a high profile after controversial cases in the 1970s where individuals were convicted for being involved in IRA bombings, based on faulty evidence and/or coerced confessions (Lutz et al., 2002). In 1976, the Devlin Inquiry (Report of the Committee on Evidence of Identification in Criminal Cases, or Devlin Report; Devlin, 1976) was established to investigate wrongful convictions and the aspects of law and procedure relating to evidence of identification in criminal cases. Subsequently in 1993, the Royal Commission on Criminal Justice was formulated to assess the effectiveness of the criminal justice system in England and Wales in convicting the guilty and acquitting the innocent (Walker & Starmer, 1999). Similarly, in 1997, the formulation of the Criminal Cases Review Commission in the UK marked the acknowledgment of miscarriages of justice in the form of wrongful convictions and has since facilitated the quashing of 272 wrongful convictions (Langdon & Wilson, 2005; Weathered & Blewer, 2009). Research on these wrongful convictions has identified eyewitness misidentification
as a leading factor in these miscarriages of justice (Colvin, 2009; Walker & Starmer, 1999).

1.4.3 Cases in Australia

The Australian legal system is an English-based system and therefore places a similar emphasis on judges’ directions as legal systems in the UK. In Australia, wrongful convictions have gained comparatively less attention, as no review, commissions, or major reforms have taken place to address the problem. However, instances of wrongful conviction involving eyewitness misidentification in Australia can certainly be noted, with the quashing of Gordon Wood’s conviction in 2012, for killing his then girlfriend an Australian model, Caroline Byrne in 2008, being one such example (Wood v R, 2012). Despite the spate of post-conviction DNA exoneration in the USA, only one appellate level DNA exoneration has occurred in Australia in the case of Frank Button who was convicted of rape in 2000 based on erroneous eyewitness identification evidence, but this conviction was quashed on appeal due to DNA evidence that had not been tested in the original conviction (Weathered & Blewer, 2009). The lack of DNA exonerations is likely due to the lack of Australian legislation allowing the post-conviction testing of DNA evidence (Weathered, 2003; Weathered & Blewer, 2009). The exception to this, is New South Wales (NSW), which is the only Australian state to enact such legislation with the Crimes (Appeal and Review) Amendment (DNA Review Panel) Act (2006; Weathered & Blewer, 2009). As such, it can be assumed that studies of wrongful convictions in Australia are limited by the absence of legislation allowing the post-conviction testing of DNA, and thus official documentation of many wrongful convictions. Given this limitation, the factors leading to wrongful convictions in Australia have not been well studied and thus the number of cases in which eyewitness misidentification has led to wrongful conviction has not been elucidated. There is little reason to suspect, however, that these factors would differ substantially from findings in the UK, as the Australian legal system closely follows the English-based systems in the UK (Macken & Dupuche, 2011). Thus, it appears likely that unreliable eyewitness identification evidence is a significant contributor to wrongful convictions in Australia, with one study suggesting that unreliable eyewitness identification was responsible for miscarriages of justice in 16% of cases in an Australian sample of 32 (Langdon & Wilson, 2005).
It is argued that English-based legal systems, such as those in Australia and parts of the UK, have additional safeguards against the admission of unreliable eyewitness identification evidence compared with the USA (Colvin, 2009). This is because judges are allowed, and in some cases obliged, to evaluate evidence for, and warn juries of, the specific dangers present in the identification evidence before the court (Kirby, 1991). However, judges do not always choose to warn jurors adequately, or even exclude highly unreliable identification evidence. For example, in the case of *R v Dupas* (2011), the court decided to allow eyewitness evidence of identifications that were only made after witnesses had seen media portrayals of Dupas connecting him with another murder. Indeed, Australian Justice Michael Kirby laments “Some of the most shocking cases of miscarriage of justice have occurred as a result of mistaken identity” (Kirby, 1991, p. 1043).

### 1.4.4 Eyewitness identification evidence and interactions with other factors

It must be noted that certain authors also argue that the current research on eyewitness identification is too focussed on singular factors (Colvin, 2009). It may be that while eyewitness misidentification is a factor contributing to many miscarriages of justice, there may also be factors interacting with eyewitness identification which amplify this effect. For example, it may be that rather than misidentification itself, it is the tendency for police officers to rely on unreliable eyewitness identification evidence to secure a conviction, or the tendency for lawyers to rely on this evidence in the courtroom. These factors may contribute just as much to the rate of involvement of eyewitness misidentification in wrongful convictions as the misidentification itself. Similarly, in English-based systems, there may be an interaction with a tendency of judges to believe that eyewitness identification evidence is reliable and a decreased tendency to warn jurors. As such, the research on eyewitness misidentification must be understood not only in isolation, but also in light of these various factors, which potentially contribute to the relationship between misidentifications and wrongful convictions.

### 1.5 The process of eyewitness identification - the identification parade or lineup

Eyewitness identification requires two elements: at least one witness to a crime and a suspect. The procedures of eyewitness identification are different in
different jurisdictions although there are some procedural similarities across many western jurisdictions. In Australia these procedures also differ slightly between different state jurisdictions, however it is outside the scope of the current research to provide a comprehensive overview of such differences. Given that the present research was conducted using mainly participants from Australia, for brevity, the process will be described here in general terms based on the procedures in Victoria, Australia (Victoria Police, 1998).

The process of eyewitness identification most often starts with the police obtaining a description of the offender from the witness, witnesses, or from a closed circuit video. This forms part of the evidence available to police when identifying a suspect. Other forms of evidence are also collected, such as physical evidence or biological evidence, for example fingerprints or DNA. On the basis of this evidence, the police ideally obtain a suspect, or a person the police consider is likely to have committed the offence. In the case that no suspect can be found, witnesses may be shown photograph books for identification purposes in order to obtain a suspect (Victoria Police, 1998). In the case that the police obtain a potential suspect, they will attempt to strengthen their case against the suspect through several avenues, most commonly a police interrogation, or interview, of the suspect, gathering circumstantial or indirect evidence from other areas of the suspect’s life, and/or a police identification parade, or lineup. The police interrogation is a process in which the police attempt to confirm their suspicion that the suspect committed the offence through eliciting a confession or other incriminating information from the suspect. [See Kassin & Gudjonsson (2004) for a review.] The police interrogation, or interview, sometimes, although not always, succeeds in eliciting a confession from the suspect. In the case that no confession is elicited, the only piece of direct evidence that the police can obtain becomes the witness’s identification of the suspect as the person they saw commit the act in question (Wells, 1984; Wells et al., 1998).

The police lineup, or identification parade, is the gold standard procedure for eyewitness identification of a perpetrator and is considered reliable identification evidence by the courts (Alexander v R, 1981). The police lineup is a procedure in which a witness views a suspect alongside several foils, or people who are known to be innocent, and the witness is asked if they see the person who they saw commit the crime (Wells et al., 1998; Wogalter, Malpass, & Mcquiston, 2004). Foils for lineups
Chapter 1 | The Problem of Eyewitness Misidentification

are selected based on visual similarity to the suspect. In the case of a live identification parade, or lineup, the witness is asked to attend the police station to view the lineup. Traditionally, in most jurisdictions the police have used the simultaneous lineup procedure. This procedure is still routinely used in Victoria, Australia. Simultaneous lineups involve all lineup members being presented to the witness at the same time and the witness is asked whether the person who they saw commit the relevant crime is present in the lineup (Steblay, Dysart, Fulero, & Lindsay, 2001). In contrast, in sequential lineups witnesses view lineup members one at a time and are asked to make a judgment about whether the lineup member being presented was the person they saw commit the offence. If the witness indicates that the lineup member being presented is not the person who committed the crime, that lineup member is replaced by another lineup member. This process continues until either all the members of the lineup have been presented or a lineup member is identified as the person who committed the offence, and when that occurs no further lineup members are presented.

The police officers involved in the investigation typically construct the lineup, asking the suspect along with several other foils if they will give consent to participate in the identification parade. It is not usual that double blind procedures are used in police investigations in Victoria. In other words, the officers conducting the investigation also commonly conduct the lineup (Victoria Police, 1998). The witness views the lineup members, and then makes a decision about whether the perpetrator of the crime is in the lineup. If the witness indicates that they see the person, they are asked to indicate which lineup member they think is the perpetrator, usually by indicating the number, or position of the person in the lineup. The witness is not given any identifying information about the members of the identification parade. The witness will also commonly be asked how sure, or certain, they are of their identification decision (Victoria Police, 1998). If an identification is made, the result is recorded and this can be used in court as evidence. If the witness indicates that the person who they saw commit the crime in question is not in the lineup, then they may be asked to come back at a later date to view another lineup.

1.6 Legal rulings governing the process of identification lineups

The procedure of the identification lineup is complex and involves several variables, some of which are under the control of the police within the criminal
justice system. Unless a legal statute stipulating specific guidelines exists, police procedures are guided by common law. In other words, police procedures not given specific guidelines in legal statutes are only scrutinised if a criminal case arises in which the use of these procedures is contested in a courtroom (Beazer, 2002). As such, there are many variables inherent in lineup identification procedures that have been given no clear legal guidelines, such as selection of foils, instructions given to witnesses, the position of the suspect in the lineup, and the type of lineup administered (Evidence Act, 2008). There have been several rulings which govern and restrict the way in which the police in different jurisdictions conduct lineups. These rulings give guidelines on several procedures that have been identified in the criminal justice system as potentially impacting upon the course of justice.

### 1.6.1 Live versus photographic lineups

Although a live identification parade is considered by judges setting legal precedent to be the gold standard procedure for obtaining an eyewitness identification, in the event that this is not possible, photographic identification may take place (Alexander v R, 1981). Legal rulings have been made, however, that photographs must be presented among other photographs (as in a lineup parade) and that the photographs cannot be suggestive of guilt (for example, a suspect cannot be viewed in prison clothes or handcuffs; Alexander v R, 1981; R v Williams, 1983). A photographic lineup procedure is similar to a live identification parade, however instead of live people the witness views several photographs and makes their identification decision based on these (Victoria Police, 1998). In this case, the police suspect is not present when the photographic identification takes place. The use of photographic lineups is common in the USA (Wogalter, Malpass, & Burger, 1993), the UK (Valentine & Davis, 2015) and Australia (Bates, 1977; Evidence Act, 2008), and is often preferred to the live identification parade for logistical reasons, such as physical availability of the witness or foils. Thus, the current research will use photographic lineups in its examination of identification confidence and accuracy.

### 1.6.2 Number of foils

Many cases have ruled that identification of a suspect when the witness views the suspect without any foils (also called a show up) is not an acceptable identification procedure (Alexander v R, 1981; Davies and Cody v R, 1937; R v
Williams, 1983). This issue arose within the context of legal cases in which police officers carrying out identification procedures included a single suspect in the identification lineup and asked if the witness could identify the suspect as the perpetrator. It was decided by judges in these cases that using identification procedures in which only one suspect is viewed by a witness was unduly suggestive. In other words, the context in which the witness views the suspect (in a police station or in a courtroom dock) suggested or implied their guilt and therefore this lineup procedure was considered unfair to the suspect. Although there have been many rulings relating to show ups, there are no clear guidelines surrounding how many lineup members should be included in a police lineup. Under Commonwealth law, identification parades are required to include at least nine people including the suspect in question (Crimes Act, 1914). However, criminal justice is not governed by Commonwealth law and there is no consensus among the different Australian state jurisdictions, who conduct lineups under criminal law, as to how many foils should be included in a lineup (Evidence Act, 2008).

1.6.3 Type of Lineup

Although no legal rulings exist in Australia stipulating the type of lineup that should be administered, there have been some guidelines suggested for jurisdictions in the USA (Wells et al., 2000; Wells et al., 1998). In the USA some psychological research has been taken into account when outlining legal rules and guidelines for the conduct of identification lineups (National Institute of Justice, 1999). Although the researchers involved in helping to define these legal rules have suggested that sequential lineups be recommended over simultaneous lineups (Wells et al., 2000; Wells et al., 1998), this recommendation has not yet eventuated in a legal rule or guideline. Thus, the decision as to the type of lineup administered is left up to the discretion of the administering officers.

1.7 Legal safeguards governing the use of eyewitness identification evidence in court cases

Many countries have attempted to put in place legal safeguards to prevent miscarriages of justice. These safeguards have generally been formed through the outcomes of court cases and have addressed both system and estimator variables. As such, precedents have been set in many countries as to the aspects of eyewitness
identification evidence that judges and juries should consider, when assessing the reliability of the testimony before them. These precedents are reviewed in the following sections.

### 1.7.1 Case law in the USA

USA case law, like that in Australia and the UK, reveals an attempt to address the dangers of eyewitness identification testimony. Differences between the UK and American legal systems mean that the judge is less involved in these safeguards. As such, in the USA, safeguards are less strict than is traditional in the UK, rely much less on the judge’s instructions to the jury, and much more on the facts of the individual case.

The most powerful safeguard in the USA, is the case law regarding admissibility of evidence. The federal standard on this matter is set forth in *Manson v. Braithwaite* (1977) and dictates that reliability of eyewitness identification testimony must be evaluated only if an identification procedure is found to be unnecessarily suggestive. In this case, the court must then determine whether, under the totality of circumstances, the identification appears reliable, and if not, the identification evidence must be excluded from trial. As such, the case law governing admissibility in the USA depends not only on how suggestive the identification procedure was, but also if the identification evidence appears reliable or not. Both of these conditions must be violated for the evidence to be inadmissible from court. This is further reflected in the standard regarding admissibility of an in court identification, articulated in *Simmons v United States* (1968, p. 390) "pretrial identification will be set aside on the ground of prejudice only if the pretrial identification procedure was so impermissibly suggestive as to give rise to a very substantial likelihood of irreparable misidentification".

The more commonly utilised safeguard in the USA revolves around changing the weight given to eyewitness identification evidence on a case by case basis. This safeguard was articulated by the United States Supreme Court in *Neil v Biggers* (1972) in which five factors were identified to be considered by courts in assessing the weight to be given to eyewitness identification evidence. These were: “opportunity of the witness to view the criminal at the time of the crime, the witness's degree of attention, the accuracy of the witness's prior description of the criminal, the level of certainty demonstrated by the witness at the confrontation, and
the length of time between the crime and the confrontation” (p. 409). Although these factors bear some resemblance to those of *R v Turnbull* (1976) outlined below, two factors represent a contrast to those considered in *R v Turnbull* (1976), namely the witness’s degree of attention and certainty. Of particular interest is the emphasis placed on certainty of the witness, given that the Devlin Inquiry, reported in the Devlin Report, or Report of the Committee on Evidence of Identification in Criminal Cases (Devlin, 1976) and conducted by a UK based committee, looked at a number of criminal cases in order to investigate wrongful convictions and came to the conclusion that juries should be careful not to put too much emphasis on eyewitness confidence when assessing identification evidence.

1.7.2 *Case law in the UK*

The safeguards in legal systems in the UK generally rest with the trial judge to exclude poor, unfair or prejudicial identification evidence. In English law, there is authority to exclude dock identification, where the accused is identified in the dock at trial, as well as photographic identification, and these forms of evidence are excluded unless the accused directly or indirectly raises the matter in order to argue for its inclusion (*R v Wainwright*, 1925 as cited in *Alexander v R*, 1981). Following the Devlin Inquiry, the courts identified aspects of identification evidence that the jury should take into account when the case against the accused depends substantially on eyewitness identifications which are disputed by the defence. These were set out in *R v Turnbull* (1976) and included seven factors for the jury to consider when evaluating the reliability of eyewitness identification evidence (Thomson, 2003). The seven factors were: 1) the length of time the witness observed the offender; 2) the distance the witness was from the offender upon observation; 3) lighting conditions prevailing upon observation; 4) the presence of any obstructions when the witness observed the offender; 5) how familiar the offender was to the accused; 6) the length of elapsed time between the original observation and the identification of the accused; and 7) the consistency of the description given by the witness of the offender and the appearance of the accused. Also, the court held that where the case substantially depended upon a disputed identification, the jury must be warned to be cautious before deciding on the identification evidence alone. The Court of Appeal further stated that the warning should include the reason for warning, the possibility of a mistaken witness being a convincing one and a caution
that several witnesses may all be mistaken (R v Turnbull, 1976). It was also stated that the judge must indicate any specific weaknesses in the prosecution evidence, such that where the evidence is of good quality, it can be put to the jury without more than a warning, however where in the opinion of the judge the identification evidence is of poor quality, the judge should instruct the jury to acquit unless there is other evidence which goes to support the correctness of the identification (R v Turnbull, 1976). However, this ruling assumes that judges are cognisant of all the factors which impact upon accuracy in eyewitness identification, which has not always been found to be the case (Wise & Safer, 2004).

1.7.3 Case law in Australia

The case law in Australia has generally followed the approach in the UK, particularly that of English law. Indeed, the propositions of R v Turnbull (1976) have been approved in appeals decisions in most jurisdictions of Australia (Bennett v R, 1982; McCusker v R, 1977; R v Burchielli, 1981; Sutton v R, 1977). However some disparities in emphasis exist. In terms of exclusion of evidence, there is the discretion to exclude photograph identifications and to exclude photographs where, for example, the accused is depicted in prison clothes or handcuffed (Alexander v R, 1981; R v Williams, 1983). In Alexander v R (1981), it was stated that wherever possible where there is a firm suspect, a live identification parade should take place in preference to photographic identification. Additionally, Australian courts law has adopted English case law concerning identifications of individuals shown alone to witnesses, and is set out in the following passage from the judgment in Davies and Cody v R (1937):

“We think the view accepted in England, and, as far as we know, elsewhere in the Dominions, where the provisions of the Criminal Appeal Act have been adopted, should be applied in Victoria. That view, as we understand it, is that, if a witness whose previous knowledge of the accused man has not made him familiar with his appearance has been shown the accused alone as a suspect and has on that occasion first identified him, the liability to mistake is so increased as to make it unsafe to convict the accused unless his identity is further proved by other evidence direct or circumstantial” (p. 182).
However, a slight difference is that while the tendency in the courts of the UK is to exclude evidence unless the matter is explicitly raised by a party asking the court to consider its inclusion, the tendency in Australia is to allow the evidence as long as a suitable warning to the jury has been given (Alexander v R, 1981). Indeed, in Australian case law (Kelleher v R, 1974), it is stipulated that the identification warning must be given where circumstances necessitate it, but is not compulsory, although the High Court has recognised that frequently such cases involve dangers. Further, the High Court of Australia has stipulated that a general warning as to the reliability of eyewitness identification is not sufficient, and the jury must be warned as to the specific flaws or weaknesses in the prosecution’s identification evidence (Domican v R, 1992; Kelleher v R, 1974). As such, depending on the facts of the case, the jury should be warned: to be especially cautious before conviction in a case relying substantially on identification evidence, even if there is other evidence (Evidence Act, 2008, s. 116; R v Turnbull, 1976); that mistakes in identification do occur and two unsafe identifications do not necessarily support one another (R v Burchielli, 1981); that as a result of honest but mistaken identification, innocent people have been convicted (R v Clarke, 1997); and that people close to the suspect, honest and convincing witnesses, and multiple witnesses can all be mistaken (R v Turnbull, 1976). Also in contrast to the ruling under English law, there is some divergence between states in Australia as to whether the trial judge has the power to instruct the jury to acquit where the prosecution’s evidence, if accepted, would raise a prima facie case, but in the judge’s view, it would be unsafe to convict on the basis of it.

### 1.7.4 Identity versus similarity evidence

Australian courts distinguish between evidence of identity, such as positive identification in a police lineup, and similarity evidence (Pitkin v R, 1995; R v Morgan, 2009). Identification evidence is considered a direct form of evidence and therefore, as long as the evidence that a crime has been committed is satisfactory, no other evidence against an accused suspect is necessary to obtain a conviction. As such, certain safeguards have been put in place to guard against wrongful conviction due to eyewitness misidentification. However, evidence of visual similarity, that is when the witness says the suspect looks like the perpetrator, has a different status within the system (Pitkin v R, 1995; R v Morgan, 2009). Indeed, similarity evidence
Chapter 1 | The Problem of Eyewitness Misidentification

is considered circumstantial evidence, and as such a suspect cannot be convicted on similarity evidence alone \((\text{Pitkin v R}, 1995)\). However, such evidence may be admitted as circumstantial evidence along with other evidence which may point to the accused \((\text{R v Adams}, 2004)\). In the case that similarity evidence is admitted, the safeguards in place to protect against eyewitness misidentification, do not apply to similarity evidence as it is not considered identification evidence, for example the judge is not required to warn the jury as to weaknesses of the similarity evidence \((\text{Evidence Act}, 2008)\). It is only where it is unclear whether the evidence is similarity or identification evidence that the jury must be warned to first consider whether the evidence presented to them is that of positive identification, rather than similarity, in order to base their decision on this evidence \((\text{R v Morgan}, 2009)\).

The distinction between similarity and positive identification evidence is intended to protect innocent suspects by requiring a very high threshold of certainty from the witness. For eyewitness evidence to be considered direct evidence, on the grounds of which an individual can be convicted, the witness must recognise the suspect as the same person as the offender, not merely similar to the alleged offender \((\text{Pitkin v R}, 1995)\). Prominent High Court of Australia judge, Justice Michael Kirby, however, questioned the utility and logic of discriminating between similarity and identification evidence in his statement, “An objection to the admissibility of identification evidence cannot be met by simply categorising it as just another piece of the circumstantial evidence” \((\text{Festa v R}, 2001, p.51)\). As implied in his statement, it appears questionable whether the distinction between similarity and identification evidence is likely to protect innocent suspects. Rather, it may instead allow unreliable identification evidence that could be detrimental to the innocent suspect to be presented in court in a different form.

1.8 Summary

Eyewitness misidentification has been found to be a significant contributor to wrongful convictions in legal cases from several western jurisdictions, including the USA, the UK and Australia. Since legal cases illustrating the dangers of eyewitness misidentification have come to light, legal statutes and common law decisions have outlined safeguards to guard against these dangers. These safeguards include restrictions placed on police lineup procedures as well as identifying factors that may reduce the weight that eyewitness identification evidence carries in legal decision
making. However, these safeguards have been derived in response to common law cases that have been contested. This approach, while practical, relies on the subjective judgment of judges and juries in specific legal cases, and is not objective in its assessment of the efficacy of these safeguards. Furthermore, safeguards from some jurisdictions conflict with those from others. For example, while seminal rulings in the USA suggest that eyewitness confidence is a reliable measure of eyewitness accuracy, legal rulings on this issue in the UK have suggested the opposite. Thus, psychological research has been conducted in the area of eyewitness identification in order to clarify which variables may impact upon eyewitness accuracy. An examination of psychological theory explaining the processes underlying eyewitness identification, and of the extent to which legal safeguards to prevent eyewitness misidentification are supported by the psychological literature, will be presented in Chapter 2.
CHAPTER 2
Factors Influencing Eyewitness Identification: The Psychological Literature

Since the examination of legal cases in which wrongful conviction has occurred, the issue of eyewitness misidentification has become a prominent area of research within the empirical psychological literature. Although the procedure of a police lineup appears simple, there have been several issues identified within the psychological literature concerning the conduct of such lineups. The following sections will review the findings of psychological literature pertinent to the process of eyewitness identification. Subsequently, the extent to which legal safeguards put in place in the USA, UK, and Australia are supported by empirical psychological literature is considered.

2.1 Eyewitness memory

Memory is essential to the process of eyewitness identification. In order to make a lineup decision, witnesses must compare their memory of the perpetrator of a crime to a suspect in a lineup. In the simplest sense, the process of memory is the acquisition, storage, and retrieval of information. Acquisition of information will happen upon viewing or otherwise experiencing stimuli, and this depends on successful encoding of specific information into the memory system. If there is insufficient attention or exposure to the task or stimuli to be encoded there will be no memory for that task or stimuli (Cowan, 1995; Reynolds & Pezdek, 1992). Further, disruptions to memory consolidation, the transference of memories to longer term storage, or distortions of the memory once acquired can lead to inaccuracies in memory. Many factors such as alcohol consumption (Birnbaum, Parker, Hartley, & Noble, 1978; Goodwin, 1995), impaired executive functioning (Cowan, 1988), inadequate sleep (Carskadon, 2011), distracting stimuli, and attentional difficulties (Cowan, 1988; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996) may disrupt memory consolidation and storage. Some studies have also found that age may disrupt eyewitness performance with children generally giving more false identifications than adults in eyewitness identification tasks - this may be due to developmental differences in memory encoding, storage or retrieval (Humphries & Flowe, 2015; Humphries, Holliday, & Flowe, 2012). A similar effect has been
observed in seniors aged 60 to 80 with seniors making more errors and falsely recognising more faces in lineup identification tasks (Searcy, Bartlett, & Memon, 1999). This latter finding may be due to the cross-age effects whereby seniors are better at identifying same-aged faces compared to other-aged faces (Bryce and Dodson, 2013). Some recent research has also suggested that configural processing, or processing of how faces are configured as opposed to processing of the components, or features, of faces, is important to the process lineup identification, which may indicate that this kind of processing is important for memory retrieval under these circumstances (Flowe, Smith, Karoğlu, Onwuegbusi, & Rai, 2015).

As such, although the process of eyewitness identification may appear simple, the human memory system required for such a judgment is complex. Such a judgment relies not only on successful encoding of the perpetrator in the witness’ mind, but also accurate storage of that information, successful transference of that memory into long term memory, and accurate retrieval of that information at the time of lineup, possibly eventuating in recognition of the perpetrator. Interruptions to any one of these processes has the potential to disrupt the ability of a witness to make an accurate identification judgment.

### 2.2 System variables

Factors identified within the psychological literature influencing the accuracy of eyewitness memory and identification have generally been classified into two categories: *system variables* and *estimator variables* (Wells, 1993; Wells, Memon, & Penrod, 2006). System variables are variables that can be manipulated and controlled by the criminal justice system and include the form and conduct of the police lineup. System variables have been researched extensively in the psychological literature and in some jurisdictions lineup procedures have been changed in response to this research. Estimator variables are variables that are not controlled by the criminal justice system and usually have to do with the characteristics of the offender, the witness, or the circumstances of the crime. The distinction between system and estimator variables has been seen as important within the psychological literature on eyewitness identification because research targeting system variables can lead to changes in process and procedure within the criminal justice system. Knowledge of estimator variables, on the other hand, informs the capacity of a particular witness to provide reliable eyewitness identification.
Evidence. Several system variables relevant to the current examination, including selection and number of foils, instructions given to witnesses, the type of lineup, and the position of the suspect in the lineup, that have been found to affect eyewitness accuracy are discussed in the following sections (Wells, 1993).

2.2.1 Selection of foils

Foils for lineups are selected based on visual similarity to the suspect. This is logically intended to protect an innocent suspect from being erroneously identified by the witness due to superficial similarities shared by the suspect and the perpetrator in the case that a witness feels pressured to provide an identification for authority figures (Brewer & Palmer, 2010; Lindsay & Wells, 1980; Wells, 1984, 1993; Wells et al., 2000). Consistent with this logic, is psychological research demonstrating that low-similarity foils result in significantly more false identifications of an innocent suspect than high-similarity foils (Fitzgerald, Price, Oriet & Charman, 2013; Lindsay & Wells, 1980). Lindsay and Wells (1980) used a staged crime paradigm including 96 unsuspecting witnesses to examine the costs and benefits of high and low foil similarity. Witnesses to the staged crime were asked to identify a criminal (the person they saw commit the staged crime) from six-picture arrays that included foils that were either high or low in similarity to the criminal. One high-similarity and one low-similarity lineup contained a picture of the criminal while another two lineups (one high-similarity and one low-similarity lineup) contained a picture of an innocent suspect who resembled the criminal. The results indicated that high-similarity lineups produced less identifications of the criminal and of the innocent suspect than low-similarity lineups but that the reduction of the identification of the criminal was much less dramatic than the reduction in the identification of the innocent suspect. These findings suggested that using high-similarity foils protected against false identifications of the innocent suspect at low cost to identifications of guilty suspects. However, the process of selecting foils based on visual similarity to the suspect has been criticised (Wogalter et al., 2004). Laughery, Jensen and Wogalter (1988) found that there was a response bias whereby if one face was prototypic of the others in the set, then the prototype is more likely to be identified as the target face. The argument was that the prototype/suspect was only one step away from the foils in similarity, but the foils were two steps away from each other, making the prototype, or suspect in the case of eyewitness lineups, more likely to be chosen.
Subsequently, Wogalter, Marwitz and Leonard (1992) conducted four experiments examining whether target-based lineups (lineups where foils were selected on visual similarity to the target) were more suggestive than alternative lineups. In their first experiment they had 10 undergraduate students construct 10 six-person photographic lineups where the foils were selected based on visual similarity to a target. Photographs utilised in the experiment were of white male senior students cut out of student yearbooks and were homogenous in composition. The students were given 10 photographs of faces to serve as targets along with 25 photographs of faces that generally resembled the target. This preliminary resemblance was made by experimenters based on hair and face shape in order to withhold implausible foils (i.e. a brunette curly-haired thin-faced foil was not grouped with a blonde straight-haired broad-faced target). In order to construct the lineups, the students, working in pairs, were instructed to choose the photographs most resembling the targets followed by the next most similar and so on until they had chosen the five most similar faces for each of the two targets they were assigned. Subsequently, 82 different undergraduate participants were asked to study each lineup carefully and guess which face the target was without ever previously having seen the target. They were told that the lineups were constructed based on the suspect’s appearance, and were asked to ignore all factors other than appearance (e.g. ‘guilty’ looks). They found that the target was selected significantly more frequently than at chance rates (22.4% rather than 16.7%), which indicated that participants were biased towards choosing the target under these conditions.

In their three subsequent experiments, Wogalter et al. (1992) used the same procedure as the first experiment with different yearbook photographs, but used 40 lineups in total to compare this procedure with three alternative lineup constructions (i.e. 10 lineups constructed based on similarity to the target and 10 lineups constructed using each of three alternative constructions). Alternative lineups were constructed based on either: visual similarity to the target as well as visual similarity to the foil selected as most similar to the target (Experiment 2); visual similarity to the target as well as visual similarity to a randomly selected foil (Experiment 3); and all lineup members being equally similar to one another (Experiment 4). In these three experiments, 20 students participated in constructing the lineups for Experiment 2; 10 students participated in constructing the lineups for each of Experiments 3 and 4; and 82 different students participated in viewing the lineups.
and guessing which face was the target in all three experiments. No student participated in more than one experiment. The findings of these experiments supported the finding of the first experiment that when the target-based lineup procedure was used, targets were selected significantly more frequently than chance. However, this effect disappeared in all of the alternative lineup construction conditions (lineups constructed based on visual similarity to the target as well as visual similarity to another foil; visual similarity to the target as well as visual similarity to two other foils; and all lineup members being equally similar to one another). In other words, using the alternative construction methods, the target person was not chosen significantly more than at chance levels (13.3%, 16.9% and 17.2% for Experiments 2, 3, and 4) and was chosen less than using the target-based construction method. Given that all other procedures among the experiments were identical, these findings suggest that changing the lineup construction method to being based on visual similarity with foils as well as with the target reduced the bias associated with balancing visual similarity only in reference to the target. These findings demonstrate that the way in which visual similarity is balanced across the lineup can bias eyewitnesses towards choosing a given target, which may impact on accuracy in real life identification situations.

It must be noted, however, that participants in this experiment were told that lineups were constructed based on visual similarity to the suspect, whereas in police lineup situations, the witness would not be told how the lineup was constructed. It is possible that the bias seen in these experiments may have been as a result of participants looking for shared features as a result of this instruction, and that this may not occur in real lineup situations where information on how the lineup is constructed is not given. However, the results do suggest that lineups constructed based on similarity to the target can lead to a bias towards selecting the target, and this may be due to the foils sharing more similarities with the target than they do with each other (Wogalter et al., 2004).

Thus the findings relating to lineup construction based on visual similarity to the target are conflicting and difficult to reconcile (Lindsay & Wells, 1980; Wogalter et al., 1993; Wogalter et al., 2004). On the one hand, foils with high visual similarity to the suspect have been found to protect against biased identifications seen where the suspect is the only one to match the witness’s description, and on the other hand, high visual similarity foils have been found to lead to bias due to foils sharing more
features with the suspect than they do with each other. It is possible that the bias created by the foils sharing more features with the suspect than each other is smaller than that created by low visual similarity lineups. Thus, in studies including low visual similarity foils, the bias related to high visual similarity foils is outweighed in size by the low visual similarity foil bias. Nevertheless, this conflict presents practical difficulties when considering the selection of a method for the construction of fair lineups - a conflict that has yet to be resolved.

Luus and Wells (1991) have also criticised the target-based lineup construction method, and argued that foils should be selected based on visual similarity to the witness’s description of the target, rather than visual similarity to the suspect. However, this has not been widely adopted as a practice by policing bodies, possibly due to the practical constraint that eyewitness descriptions are often vague and unspecific, providing a significant challenge in trying to match foils with such descriptions (Wogalter et al., 2004). Thus, the foils in the current investigation will not be chosen based on similarity with the target, but will be selected randomly excluding implausible foils.

2.2.2 Number of foils

The issue of the number of foils that should be included in a lineup is an important one given that the number of foils is directly related to the chances of false identification (Brewer & Palmer, 2010; Wells et al., 1998). Increasing the number of foils should logically decrease the chances of false identification. Data from the USA suggest that the average number of lineup members for live lineups is about six (Wogalter et al., 1993; Wogalter et al., 2004). Assuming that the lineup is fairly constructed, the chances that an innocent suspect in a lineup would more closely resemble the perpetrator than the other lineup members and be falsely identified in a six-person lineup would be 1/6 if the witness always guesses. Further, it has been suggested that when the estimated real world identification accuracy is taken into account, the chances that an innocent suspect will be identified in a six-person lineup may be as high as 10% (Wells et al., 1998). Given the ultimate consequences of this scenario, it has been argued that a 10% chance is too high (Wells et al., 1998). However there has been no real consensus in the psychological literature as to how many foils there should be. The implications of the suggestion given in Wells et al. (1998) would be that there must be more than five foils, however how many more
may logically be dictated by the practical considerations of the police and the strength of law enforcement’s desire to protect the innocent suspect. Thus, the investigations of the current research will use seven lineup foils due to the suggestion in research that more than five foils should be included and that more foils increases the fairness of lineups, as well as practical constraints of including more than seven foils in the current research given that lineup photographs were drawn from a database and were thus not unlimited.

2.2.3 Instructions given to witnesses

The susceptibility of human perceptual processes to outside intrusions and external cues, especially of a social nature, has long been recognised in many realms of memory research, but particularly within the empirical literature on eyewitness memory (Brewer & Palmer, 2010; Loftus, 2005; Steblay, 1997). It has been noted for some years that instructions given to witnesses by lineup administrators can have an effect on witness decision making (Malpass & Devine, 1981a). Malpass and Devine (1981a) conducted an experiment in which 100 student witnesses to a staged vandalism underwent lineup identification procedures under biased and unbiased instruction conditions. Biased instructions led students to believe that the perpetrator was in the lineup, whereas unbiased instructions stated that the perpetrator may or may not be present. They found that unbiased instructions reduced false identifications without decreasing correct identifications. This finding has been replicated in several subsequent studies (Brewer & Wells, 2006; Steblay, 1997) and suggests that when witnesses are led to believe that the perpetrator is in the lineup, they are more likely to make a false identification. This effect, however, has been found to be weaker than originally suggested by Malpass and Devine’s (1981a) research (Köhnken & Maass, 1988). The effect of biased instructions may be a product of social influence, as much psychological research suggests that people are susceptible to normative social influence, particularly when authority figures are involved (Steblay, 1997). Witnesses who are uncertain may feel normative pressure from the police to identify someone from the lineup. In this way, instructions to witnesses may affect witnesses’ decision making strategies and readiness to choose, rather than affecting their memory or perception of the perpetrator.

Much of the research on lineup instructions has focussed on whether the witness is pressured to make a decision, or directly or indirectly led to believe that
the suspected perpetrator is in the lineup (Steblay, 1997). However, more subtle differences in administrator wording, the witness’s knowledge or misinterpretation of administrator actions, and the implications of making an identification in general may also affect witness responding. For example, a witness may be likely to assume that the police would not arrange an identification lineup if they did not believe they had detected the perpetrator, which may cause the witness to feel some pressure to identify a person from the lineup as the perpetrator. Many of these factors are yet to be investigated and as such their effects remain unknown. Indeed, in their recommendations for eyewitness identification procedures, Wells et al. (1998) suggested that the lineup administrator should never know the identity of the suspect in the lineup - a recognition of the possibility for very subtle external cues to be unintentionally transmitted to the witness by a knowing lineup administrator. Thus the current investigation used an online survey for administration of lineups in order to remove possible experimenter biases.

2.2.4 Type of lineup

With the growing public and researcher awareness of the problem of eyewitness misidentification, the type of lineup used in an eyewitness identification has become a contentious issue (Brewer & Palmer, 2010; McQuiston-Surrett, Malpass, & Tredoux, 2006; Steblay et al., 2001; Steblay, Dietrich, Ryan, Raczynski, & James, 2011; Steblay, Dysart, & Wells, 2011; Zimmerman, Malpass, & MacLin, 2006). A considerable amount of research has been conducted into comparing the outcomes of simultaneous and sequential lineup procedures. Many of the reported findings show that while simultaneous lineups may result in marginally more offenders being correctly identified than sequential lineups, many less false identifications are made with sequential lineups than simultaneous lineups (Clark & Davey, 2005; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Pozzulo et al., 2008; Steblay et al., 2001; Steblay, Dysart & Wells, 2011). For example Lindsay and Wells (1985) found that sequential lineups eventuated in less false positive identifications of the suspect in the lineup without reducing the number of hits when the target was present. Thus, generally, it has been concluded that sequential lineups are superior to simultaneous lineups (Kassin, Tubb, Hosch, & Memon, 2001; Steblay et al., 2001; Steblay, Dysart & Wells, 2011). There have, however, been some findings contrary to this conclusion within the literature (Carlson, Gronlund, &
Clark, 2008; Dobolyi & Dodson, 2013; Ebbesen & Flowe, 2002; Flowe & Ebbesen, 2007; Gronlund, 2005; Gronlund, Carlson, Dailey, & Goodsell, 2009; Mickes, Flowe, & Wixted, 2012; Zimmerman et al., 2006). For example, one study which sought to replicate the original findings of a sequential advantage reported more false identifications of the innocent suspect under sequential lineup procedures (Carlson et al., 2008). They also found that the sequential advantage only occurred for unfair lineups where the suspect stood out and was enhanced when the target was placed late in the lineup (Carlson et al., 2008). Another large study on differences in accuracy between sequential and simultaneous lineup procedures also found that there was no evidence for an advantage of either sequential or simultaneous lineup procedures under real world conditions where other factors, such as photograph quality, vary randomly (Gronlund et al., 2009). Further, research has found that discrimination of whether a guilty target is present or absent in the lineup is inferior in sequential lineups (Mickes et al., 2012). Some researchers criticise more traditional methods of analysis used to investigate the difference in accuracy between lineup types, contending the receiver operating characteristic (ROC) analysis should be used to measure discriminability (Rotello & Chen, 2016; Wixted & Mickes, 2014; Wixted & Mickes, 2015). However, these arguments have been contentious, with others arguing that the benefits of ROC analysis have been overstated (Lampinen, 2016). Regardless of the analysis technique used, inconsistencies in findings between studies investigating differences in accuracy between lineup types have continued to exist (Dobolyi & Dodson, 2013; Mickes et al., 2012). The reasons for the inconsistency in findings are thus far unclear and as such variables contributing to differences in findings require further investigation.

One noteworthy observation of the eyewitness literature examining differences between sequential and simultaneous lineups as a whole is the diversity of paradigms used across different experiments (Sporer, Penrod, Read, & Cutler, 1995; Steblay et al., 2001; Steblay, Dysart & Wells, 2011). Most paradigms stage live crimes and therefore variables such as the type of crime committed, the length of time the perpetrator is observed, and the delay before participation in the lineup vary between investigations (Clark & Davey, 2005; Lindsay, Read, & Sharma, 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; Wells, Lindsay, & Ferguson, 1979). Most previous studies on eyewitness identification, have included some delay between observation of the target or
suspect, and commencement of the lineup (Clark & Davey, 2005; Lindsay et al., 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; Wells et al., 1979). Such delays most often occur in studies where live crimes have been staged or viewed and then some time is taken to debrief participants and obtain consent to participate in the lineup identification task (Lindsay et al., 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; Wells et al., 1979). Sometimes the time delay is 5 to 20 minutes (Clark & Davey, 2005; Kneller, Memon, & Stevenage, 2001; Lindsay et al., 1998; Lindsay & Wells, 1985; Meissner, Tredoux, Parker, & MacLin, 2005; Wells et al., 1979), whereas in other protocols, several hours or days may elapse before a lineup identification takes place (Malpass & Devine, 1981a). Although it is common in the real world for there to be a delay of several days or weeks before an eyewitness identification takes place, the impact of the varying delays on outcomes in aforementioned psychological experiments is not clear.

Furthermore, generally lineup procedures include a single lineup and each participant makes one lineup decision. Thus, variables are generally investigated between-subjects. Due to practical difficulties with recruitment, this has often led to few variables being manipulated or investigated at one time. Particularly, McQuiston-Surrett and colleagues (2006) note that much of this literature does not investigate the effect of target position and does not counterbalance target position in the lineup, which may be important especially for sequential lineups. As such, it is difficult to ascertain whether variables that differ between investigations, such as target position, may have some interaction with lineup procedure which may contribute to results. Thus, the current investigation aims to provide a thorough examination of several variables that may interact with lineup type, including target position, using a within-subjects design which allows participants to contribute to a number of different conditions. This design will allow for investigation of several variables as well as tighter experimental controls, in terms of controlling for differences in lineup characteristics between lineups and will increase the power of the current investigation’s findings considerably.

Nevertheless, if any variables have had an impact on the influence of lineup type, their effects have not been sufficiently strong to extinguish the observation of a clear benefit of sequential over simultaneous lineups, at least in terms of false positive responding, within the literature (Clark & Davey, 2005; Lindsay et al., 1998;
Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; McQuiston-Surrett et al., 2006; Sporer, 1993; Steblay et al., 2001; Steblay, Dysart & Wells, 2011; Wells et al., 1979). The theoretical explanations behind the sequential superiority effect observed in the literature, however, have been less widely agreed upon. The current research aims to add to the current understanding of the theoretical mechanisms underlying eyewitness identifications.

2.2.4.1 Explanations for the ‘sequential superiority effect’

Two explanations have been proposed to account for the sequential superiority effect, absolute versus relative judgments and higher decision criterion.

Absolute versus relative judgments

It has been argued that the simultaneous lineup procedure encourages individuals to rely on a relative judgment process when making their identification. Wells (1984) proposed that in simultaneous lineups, rather than comparing each individual in the lineup to their memory of the perpetrator, the witness compares each individual in the lineup to one another and chooses the person who is most similar to the perpetrator. Wells (1984) proposed that the way in which foil similarity in simultaneous lineups protects the suspect, is to spread the choices among other members of the lineup, making it less likely that the witness will choose the suspect. This proposition stemmed from Lindsay and Wells’ (1980) findings that high visual similarity foils produced less false identifications of an innocent suspect than low visual similarity foils. This was interpreted as high-similarity foils resulting in a spread of choices from the innocent suspect to the known innocent foils. Further, they, and others, found that fewer false positives could also be obtained in simultaneous lineups when witnesses were told that the perpetrator may, or may not be in the lineup (Brewer & Palmer, 2010; Malpass & Devine, 1981a; Wells, 1984). It was argued that these findings could be understood to result from a tendency of witnesses to make their decisions based on relative judgments, and thus telling witnesses that the perpetrator may not be present makes apparent the erroneous use of relative judgment processes in perpetrator-absent lineups (Wells, 1984). The impact of instructions and foil similarity in sequential lineups was not investigated, although the underlying assumption appeared to be that these factors would have a differing, or lesser, impact in sequential lineups. Further, it was not made clear
whether witnesses would then avoid the use of relative judgments under these circumstances, and if so, what strategy witnesses would thereafter use to make identification judgments. Further, in the case that a witness makes a relative judgment in an identification situation, it is not clear, according to the explanation of relative judgments in the literature, when in the identification process the witness compares the lineup members with the image of the perpetrator in their mind. Logically, this comparison must occur in order for a witness to decide which lineup member most resembles the perpetrator, however Wells (1984) does not describe how the relative judgment occurs.

It has been proposed that the use of relative judgments would be adequate in perpetrator present lineups, but represents a danger in perpetrator absent lineups where the innocent suspect may be likely to resemble the perpetrator. A witness who relies on relative judgment processes under these conditions may be at increased risk of making a false identification. It is argued that these risks may be mitigated by the sequential lineup procedure, where witnesses view lineup members one at a time and are asked to make a judgment about whether each person was the person they saw commit the offence, in isolation of the other lineup members. The contention is that the sequential lineup procedure discourages reliance on relative judgments and necessitates absolute judgment processes, where a witness compares each individual presented with the image of the perpetrator in their mind (Wells, 1984). This proposition is argued by proponents of sequential lineups to be supported by research demonstrating that the simultaneous lineup procedure produces more false identifications relative to sequential lineup procedures (Lindsay & Wells, 1985; Steblay et al., 2001; Steblay, Dysart & Wells, 2011). Further, Grondlund’s (2004) research using height has investigated the difference between relative encoding (i.e. taller than…) and absolute encoding (i.e. 1.83m/6ft tall) on performance in simultaneous and sequential lineups. Their research found that the type of encoding (absolute versus relative) interacts with lineup procedure such that performance under simultaneous lineup procedures is best when information was encoded relatively, whereas performance under sequential lineup procedures is best when information was encoded absolutely (Gronlund, 2004). These findings were interpreted as suggesting that simultaneous lineups encouraged relative judgments whereas sequential lineups encouraged absolute judgments, which would lend support to the relative versus absolute judgment explanation. However, other
researchers have argued that the measurement of this theoretical presumption is particularly difficult, and so it has not conclusively been demonstrated that this is the way identification decisions are made in simultaneous and sequential lineups (McQuiston-Surrett et al., 2006).

Some researchers have focussed on false identifications of the target replacement, rather than all false positive identifications in their analyses (Amendola & Wixted, 2015; Mickes et al., 2012). While this is relevant in terms of real life eyewitness lineups (only false identification of the suspect is relevant in a real life situation), theoretical explanations should take into account false identification of any non-target member of the lineup. This is because differentiating between a false identification of a foil and a miss (both incorrect responses where the target replacement was not chosen) gives important information about the nature of the decision making process - one indicates that people are mistaking a foil for the target, and one indicates the person simply did not see anyone they were certain was a match with the target. This is relevant to the theoretical question of how lineup decisions are made under different conditions, and particularly to the question of whether people tend to engage in relative judgment in simultaneous lineups in comparison to sequential ones.

**Signal detection theory- higher decision criterion in sequential lineups**

An alternative interpretation of data demonstrating the sequential superiority effect has been that sequential lineups may make eyewitnesses more conservative in general, rather than the difference being due to absolute judgments (Clark, 2005, 2012; Clark & Davey, 2005; Clark, Howell, & Davey, 2008; Flowe & Ebbesen, 2007; Malpass & Devine, 1981a; Mickes et al., 2012; Palmer & Brewer, 2012). For example, it has been argued that the aforementioned study in which participants were instructed that the perpetrator may not be in the lineup (Malpass & Devine, 1981a) may not have encouraged participants to rely on absolute judgments but rather made them more conservative causing them to choose less often. Indeed, participants in the study who were told that the perpetrator may or may not be present also chose significantly less often than those who were given biased instructions. This finding suggests that the reduced false identification rates observed in that research may not have been due to participants’ recognition of relative judgment strategies, but rather
may have been due to participants being more reluctant to choose, and thus choosing less often in general.

An alternative explanation is that people may use relative judgments when making decisions in response to both types of lineup. In simultaneous lineups witnesses may select the member of the lineup who looks most similar to their memory of the offender, that is, they make a relative judgment. In contrast, in sequential lineups it has been assumed that each member of the lineup is compared independently by the witnesses to the memory they have of the offender, an absolute judgment. Original propositions that relative judgments accounted for weaknesses in the simultaneous lineup procedure came from findings that when a target that was originally identified was removed from a lineup, the tendency of witnesses in both high and low visual similarity lineups was not to reject the lineup, but rather to drift towards identifying a foil sharing superficial similarities with the suspect (Lindsay & Wells, 1980; Wells, 1984). However, more recent data replicate these findings, and also report a similar target to filler identification drift using a sequential lineup procedure (Clark & Davey, 2005). Flowe and Ebbesen (2007) also reported similar findings whereby visual similarity between the target and foils had an effect on accuracy in both sequential and simultaneous lineups. This finding is contrary to the proposition that the sequential superiority effect is due to increased reliance on absolute judgments compared with simultaneous lineups. If participants were more likely to use absolute judgments as described by Wells (1984) in sequential lineups, target similarity with foils should not have impacted accuracy in the same way as in simultaneous lineups, as participants would not have been comparing between lineup members, but only between their memory of the target and each lineup member individually. Rather, these findings suggest that participants may be using relative judgments in both types of lineup, selecting the ‘best fit’ relative to the other lineup members they have viewed. Further, measurement of whether individuals actually use relative or absolute judgments necessitates reliance on self-report data (McQuiston-Surrett et al., 2006). This method carries the limitation that individuals who are asked to what extent they compared images or considered each on its own after doing a simultaneous or sequential procedure, may respond based on their recollection of the task, rather than a recollection of their cognitive strategy (McQuiston-Surrett et al., 2006). Indeed, an assumption underlying self-report methods is that a person will be aware of, recollect, and report on the item in
question. However, it is unclear whether individuals would be aware of their cognitive strategy when completing self-report items retrospectively.

Some writers suggest that the eyewitness findings may be better explained by signal detection theory (Banks, 1970; Clark, 2012; Ebbesen & Flowe, 2002; Flowe & Ebbesen, 2007; Malpass & Devine, 1981a; McQuiston-Surrett et al., 2006; Mickes et al., 2012; Palmer & Brewer, 2012). In the context of recognition, signal detection theory suggests that the probability of recognition can be understood in terms of an abstract dimension known as the ‘likelihood axis’, or in terms of memory, the ‘familiarity axis’ (Banks, 1970- See Figure 2.1).

Figure 2.1. Normal distributions of new (peak represented by $\mu_n$) and old items (peak represented by $\mu_o$) along the familiarity axis (with CA- CF representing varying levels of familiarity thresholds) assumed with the $d'$ model of signal detection theory (Banks, 1970).

New (not seen before) and old (previously seen) items are both assumed to be normally distributed along the familiarity axis, with the distribution of old items generally being higher on the familiarity axis than the distribution of new items. The measure of observer sensitivity or recognition acuity, $d'$, is the difference between the means of the two distributions. Points along the familiarity axis represent different levels of criteria, and can be conceptualised as the individual’s leniency or strictness when making a judgment (Banks, 1970). For example, an individual
applying very strict criteria (such as $C_A$ in Figure 2.1 above) may require a very high level of familiarity with the item before judging that they recognise it. On the other hand, a very lenient criterion threshold (such as $C_F$ in Figure 2.1 above), would result in a judgment that the item is “old” for a great many items, both old and new (or all items lying above $C_F$ on the familiarity axis). Further, signal detection theory assumes the decision criterion along the familiarity axis and difference between the means of the familiarity distributions of new and old items. A basic assumption underlying signal detection is that the observer's recognition acuity sensitivity and response criterion are independent from each other and may be influenced by different factors.

One explanation for the differences in hits and false positives in sequential compared with simultaneous lineup procedures is that a stricter criterion level is set in sequential lineups (Ebbesen & Flowe, 2002; McQuiston-Surrett et al., 2006). Individuals may be more conservative when making a judgment in a sequential lineup. This explanation is consistent with findings that sequential lineups are superior for minimising false identifications and increasing correct lineup rejections when the perpetrator is absent, whereas simultaneous lineups are superior for securing correct identifications and reducing false lineup rejections when the target is present (Brewer & Palmer, 2010; Clark, 2012; McQuiston-Surrett et al., 2006). Meissner and colleagues (2005) sought to investigate this possibility using a memory paradigm whereby they presented participants with a number of target faces followed by a number of lineup tasks, and collected signal detection measures. Their results were consistent with predictions of signal detection theory, indicating a conservative criterion shift with the sequential presentation of lineups. Their findings also suggested that the conservative shift in criterion may have been due to a reduction in familiarity-based responding in sequential lineups which they recognised may also be consistent with the idea that simultaneous judgments may encourage relative judgments in comparison to sequential lineups.

Further, Flowe (2005) used a memory paradigm to investigate this issue and hypothesised that if relative versus absolute judgment processes explained differences in accuracy between sequential and simultaneous lineups, then visual similarity between lineup foils and the target lineup member would only affect accuracy in simultaneous lineups. In other words, different levels of target to foil similarity should not affect accuracy in sequential lineups because the participants
should not be comparing between lineup members, but rather only between their memory of the perpetrator and each lineup member. On the contrary, they found that visual similarity between foils and the target in the lineup affected accuracy in both sequential and simultaneous lineups (Flowe, 2005; Flowe & Ebbesen, 2007). One explanation for this finding may be that a witness must compare a given member of the lineup with their memory of the perpetrator, regardless of the type of lineup administered. Therefore, the more similar the lineup member, the more difficult the task, regardless of whether the lineup completed is presented in a simultaneous or sequential manner. Flowe’s (2005) studies also found that accuracy was higher in simultaneous than sequential lineups when fewer features were available for discriminating between faces, but under forced choice conditions differences in accuracy between lineup procedures were reduced. These findings indicated that when told the target was in the lineup and forced to choose, participants were equally likely to select the target in both lineup procedures, but that differences in accuracy were due to participants in sequential lineups rejecting the lineup more often. This suggested that not being able to compare faces in sequential lineups may lead witnesses to adopt a higher decision criterion. On the basis of these findings, rather than sequential lineups being superior, they may simply make witnesses more conservative, decreasing errors of false identification, but possibly increasing errors of false rejection and decreasing the likelihood of correct identifications. The latter outcome would be contrary to the aims of the police, one of which is to maximise the identification of the perpetrator of a crime when the perpetrator is present in the lineup (Clark, 2012).

Research inconsistent with prominent theoretical explanations for differences between simultaneous and sequential lineups

Neither a criterion shift in signal detection theory nor relative versus absolute judgments explain findings that sequential lineups improve accuracy in terms of false positive identifications without a cost to hits (Lindsay & Wells, 1985; Steblay et al., 2001; Steblay, Dysart & Wells, 2011; Wells, 1984). For example Lindsay and Wells (1985) found that the sequential procedure resulted in fewer false positive identifications without reducing the number of hits when the target was present. The adoption of a more conservative criterion threshold should result in fewer choices in
general, and therefore not only decrease false positive identifications but also decrease hits.

Some researchers have argued that the sequential advantage may only result under certain circumstances (Carlson et al., 2008; Clark, Erickson, & Breneman, 2011). Gronlund (2004, 2005) has proposed that this advantage will only occur if distinctive information is encoded and recollection is used to access that information. Additionally, one study which sought to replicate the original findings of a sequential advantage shown by Lindsay and Wells (1985), found no such advantage, reporting more false identifications of the innocent suspect under sequential lineup conditions (Carlson et al., 2008). Carlson et al. (2008) found that the advantage only occurred for unfair lineups where the suspect stood out and was enhanced when the target was placed late in the lineup (Carlson et al., 2008). These findings suggested that the sequential advantage was dependent on target position and lineup fairness. Another large study of 2529 undergraduate participants found that there was no evidence for an advantage of either sequential or simultaneous lineup procedures under real world conditions where other factors, such as photograph quality vary randomly (Gronlund et al., 2009). Further, Mickes (2012) found that simultaneous lineups were superior at discriminating the presence versus absence of a guilty suspect in a lineup as measured by receiver operating characteristics. These findings cast some doubt as to whether the sequential lineup procedure is superior as has been claimed.

Questions have also been raised as to whether the differences found between sequential and simultaneous lineup procedures are due to the lineup procedures per se and not to other variables that accompany the respective procedures (Zimmerman et al., 2006). For example, sequential lineup procedures often use back-loading, where administrators lead witnesses to believe they will be presented with more photos than they will actually get to see, and they also often require that witnesses make multiple identification decisions, as they are asked an identification question for each lineup member. Zimmerman et al. (2006) manipulated these variables independently of lineup procedure. They replicated previous findings within the literature that when sequential and simultaneous lineups were used along with the variables typically associated with that lineup (sequential lineups are back-loaded and multiple decisions are made, while simultaneous lineups are not back-loaded and one decision is made), correct rejections of perpetrator absent lineups were more frequent in sequential than simultaneous lineups. However, when the variables
accompanying the lineups were held constant or reversed, (sequential lineups were not back-loaded and required only a single decision while simultaneous lineups were back-loaded, in that participants were led to believe they would see more sets of photographs than the set they were currently viewing, and required multiple decisions) the sequential superiority effect disappeared. These data suggest it may not be the lineup procedure itself, but rather the typical differences in procedural administration of each lineup that may be responsible for the sequential superiority effect so often reported in the literature. These findings, collectively, call into question the robustness of the sequential superiority effect, suggesting that many other factors may be relevant to determining whether this effect is observed. The current research aims to clarify patterns of accuracy in sequential compared to simultaneous lineups so that the inconsistency between research findings in this area may be further resolved. Through more in depth exploration of the impact of lineup type on accuracy, a greater understanding of the theoretical mechanisms underlying eyewitness identification can be developed.

2.2.5 Position of the suspect in the lineup

The position in which the suspect is presented in the lineup is a system variable that is thought to be particularly pertinent to sequential lineups due to the sequential presentation of lineup members, although the relative position of lineup members in a simultaneous lineup may also have some effect (Gonzalez, Davis, & Ellsworth, 1995). Position effects have generally not been reported within the eyewitness literature with many studies finding no position effects for sequential lineups (Flowe & Ebbesen, 2007; Lindsay & Wells, 1985; Sporer, 1993). Perhaps as a result of the early studies reporting no position effects, studies within the eyewitness literature have tended not to examine or report the effects of target position on accuracy (Carlson et al., 2008; Clark & Davey, 2005). McQuiston-Surrett and colleagues (2006) examined whether target position could moderate the relationship between lineup procedure and accuracy. Their findings indicated that counterbalanced target position (i.e. counterbalancing the target position as either early or late in the lineup) failed to find an advantage of sequential lineups, whereas studies that did not counterbalance found a strong advantage of sequential over simultaneous lineups. Similarly, Gronlund et al. (2008) and Carlson (2008) found that when the target, or suspect, was placed early in the lineup, a simultaneous
advantage in terms of accuracy tended to result, however when the suspect was presented late in the lineup, a sequential advantage tended to result. Clark and Davey (2005) also reported position effects for sequential lineups finding that when the next best (or most visually similar) lineup foil preceded the target in the lineup, accuracy was decreased and participants more often falsely identified the next best alternative as the target. This finding was interpreted as due to participants identifying the next best alternative early in the lineup and thus never being able to view the target person. However, most participants identified the target person when they preceded the next best alternative. Further, when the target was absent from the lineup, participants more often chose the target replacements (lineup member most similar to the target) when it was presented late in the lineup. This was thought to be evidence of a downward shifting in participants’ decision criterion such that some witnesses desired to make an identification such that they lowered their decision criterion threshold over the sequential lineup in order to do so.

Given the conflicting results in relation to target or suspect position in the lineup within the eyewitness literature, it is pertinent that this variable receive further study. As a system variable that may impact upon eyewitness accuracy, particularly in sequential lineups, more information should be gathered about the effects of target position. This would be in the interests of minimising wrongful convictions as a result of system variables. Particularly, in line with the assertions of McQuiston-Surrett and colleagues (2006), even if target position is not the focus of the study, target position should be counterbalanced in studies examining eyewitness accuracy in sequential and simultaneous lineups due to its ability to bias results, particularly in regard to sequential lineups. The current investigation will counterbalance target position and also investigate the impact of target position on patterns of accuracy in sequential and simultaneous lineups.

2.3 Estimator variables

Several estimator variables, and whether or not they are related to accuracy have also been investigated in the empirical psychological literature. Knowledge of the impact of estimator variables on accuracy informs the capacity of a particular witness to provide reliable eyewitness identification evidence. The estimator variables that have been commonly investigated include duration of observation; obstacles to viewing, distance and lighting conditions; familiarity with, and visual
similarity of, the target person; length of time between viewing the target and identification in the lineup; consistency between the description of the perpetrator and the appearance of the accused; and eyewitness attention. The findings from this research are reviewed in the sections below.

2.3.1 Duration of observation

Studies which have varied the length of time observers have been exposed to photos and target faces have generally found that the quality of recall and recognition of targets have increased with increased exposure in terms of recognition accuracy and the number of details recalled (DiNardo & Rainey, 1991; Memon, Hope, & Bull, 2003; Reynolds & Pezdek, 1992; Shapiro & Penrod, 1986; Thomson, 2003; Wells & Murray, 1983). A meta-analysis of 128 eyewitness identification and facial recognition research studies has revealed that exposure duration is significantly related to facial identification accuracy (Shapiro & Penrod, 1986). This finding is found with both generic presentation methods, where faces are shown and participants are instructed to remember the faces (Reynolds & Pezdek, 1992), and under simulated lineup conditions, where participants are witnesses to a simulated crime and then asked to identify the perpetrator (Memon et al., 2003). These empirical findings from psychological research suggest that duration of exposure to a given target significantly contributes to accuracy of subsequent recognition of that target. Therefore, the current investigation standardised the amount of time the target photograph was viewed to five seconds in order to control for the effects of exposure duration.

2.3.2 Obstacles to viewing, distance from the offender, and lighting conditions

Generally, with increasing distance, more obstacles, and lower light, accuracy in identification is reduced (Thomson, 2003). Indeed, Yarmey (1986) conducted an experiment where 128 participants viewed a series of slides accompanied by an audiotape implying a rape scene and were told to imagine they were witnesses to a crime in a park. Under low illumination conditions or where participants viewed the scene at the end of twilight or night time, identifications by male participants were drastically reduced. Similarly, in their experiment involving 100 participants, Malpass and Devine (1981a) found that witnesses’ distance from a staged vandalism was positively correlated with errors in identification in target absent and target
present lineups. As such, the psychological literature generally supports the notion that obstacles occluding the vision of the witness viewing the offender, a greater the distance between the witness and the offender, and poor lighting conditions all impair identification accuracy. Photographs in the current investigation included no obstacles to viewing the target and were standardised in terms of distance as well as lighting conditions in order to control for the aforementioned effects.

### 2.3.3 Familiarity and visual similarity

The idea that a familiar person is more easily recognised than an unfamiliar person receives some support within the literature (Thomson, 2003; Thomson, Robertson, & Vogt, 1982). Differential processing of familiar and unfamiliar faces has been demonstrated by a number of researchers (Bruce, 1982; Ellis, Shepherd, & Davies, 1979; Thomson, 2003). For example, Bruce (1982) examined the effect of changing faces on recognition for familiar and unfamiliar faces across two different experiments. The first experiment tested 36 participants on their ability to recognise photographs of people unfamiliar to them when the photograph either remained the same or was changed, in terms of angle, expression, or both upon testing. The second experiment tested a different 24 participants who were presented with photographs of faces, half of which were highly familiar to them and half of which were unfamiliar to them, and upon recognition testing the photographs either remained the same or were changed in both angle and expression. Results indicated that overall, familiar faces were recognised more accurately than unfamiliar faces, although false positive rates were similar for familiar and unfamiliar faces. Further, unfamiliar faces were recognised more slowly and less accurately if changed in angle and expression on test, while familiar faces were recognised more slowly but just as accurately if changed on test. These findings suggested that familiar faces are generally recognised more accurately than unfamiliar faces and changes to unfamiliar faces reduce accuracy whereas recognition of familiar faces is preserved when changed in angle and expression.

Similarly, Thomson et al. (1982) conducted a comprehensive study on the role of context and familiarity in recognition of people. Results indicated that changing the pose, clothing and context of the person viewed drastically reduced the likelihood that unfamiliar people would be identified as having been seen in the earlier series of slides, but these changes had little or no impact on identification of
familiar people. Further, unfamiliar people were much more likely to be falsely identified if their pose and clothing were similar to someone in the previous series of slides compared with familiar people. Additionally, researchers suggest that recognition is facilitated when the target person is of the same race or gender as the observer, which has been suggested to be due to increased familiarity with people of the same race and gender as oneself (Brigham & Malpass, 1985; Sporer & Horry, 2011; Wright & Sladden, 2003). These findings together suggest that recognition of unfamiliar people is a much more fragile process than that of familiar people, and is affected by changes in context, hairstyle, clothing, ethnicity, gender and pose of the individual originally observed (Brigham & Malpass, 1985; Sporer & Horry, 2011; Thomson et al., 1982; Wright & Sladden, 2003). More recent research has also found that the length of time a face is examined in simultaneous lineups is related to how familiar a face is, or how similar to a previously seen face the examined face is, suggesting that similarity or familiarity of faces plays an important part in decision making in eyewitness identification tasks (Flowe & Cottrell, 2010).

While familiarity and visual similarity of the perpetrator, in general, represent an estimator variable, visual similarity in terms of context and external features can represent a system variable because it can be manipulated in an identification lineup. For example, if the perpetrator wore a black jacket and a red cap, the lineup administrators could conduct the lineup with all members wearing a black jacket and a red cap. As such, research into the effect of visual similarity on eyewitness accuracy is pertinent to recommendations relating to lineup procedures within the legal system. Thus, the current investigation will investigate the impact of visual similarity between the original observation of the target and the observation of that target in the lineup, as well as its interaction with other relevant system variables, on identification accuracy.

### 2.3.4 Length of time between the offence and identification

The length of time between the offence and the identification may be a system variable, an estimator variable, or both. If the reasons for delay between the offence occurring and the identification taking place are due to time taken for the system to operate, then it may be considered a system variable. However, often these reasons may be due to the witness not coming forward or delaying, for some other
reason, participation in the lineup. In these cases, this variable may be considered an estimator variable and will be discussed here accordingly.

In the memory literature it has been proposed that the process of forgetting happens as a function of a decay of the memory traces over time (Altman & Gray, 2002; Pavlik & Anderson, 2005). If memory decays over time, then logically the more time that elapses, the worse the memory for the event, or the person will be. The rate of memory decay has been the subject of much debate. In the short term visual memory empirical literature, memory decay is generally not observed within the first 600 milliseconds, but is observed over the subsequent 10 to 15 seconds (Alvarez & Cavanagh, 2004; Curby & Gauthier, 2007; Hollingworth, 2004; Phillips, 1974; Phillips & Baddeley, 1971). Some literature also suggests that such visual memory has a maximum capacity of a certain number of objects or faces (often four or five), but that this is also impacted upon by complexity, or information load of each item in short term memory (Alvarez & Cavanagh, 2004; Curby & Gauthier, 2007).

In the eyewitness literature, very short delays are rarely examined. Nevertheless, although the most rapid decline in memory has been found to be within minutes after encoding, the rate of decline in accuracy of descriptions and recognition is relatively modest over short delays of minutes to days, but accelerates over weeks to months (Deffenbacher, Bornstein, McGorty, & Penrod, 2008; Malpass & Devine, 1981a; Shapiro & Penrod, 1986; Thomson, 2003). Several factors seem to influence the rate of decay in terms of accuracy of witness descriptions. For example, Ebbesen and Rienick (1998) investigated eyewitness memory for an encounter with a stranger and found that obtaining a description of the target person immediately after the event, made recall more robust to the effects of elapsing time. Further, some events occurring in the intervening period, such as repeated questioning, exposure to new information or conferring with other witnesses have been found to decrease accuracy in recognition (Hastie, Landsman, & Loftus, 1978; Loftus, 2005; Henkel, 2017). These intervening events have been shown to be susceptible to being incorporated into a witness’s original memory of an event after the fact (Loftus & Greene, 1980). Thus, the psychological literature overall suggests that the longer the amount of time between observation of a target and subsequent identification of that target, the poorer accuracy is likely to be. Thus the current investigation allowed participants to view lineups directly after engaging in a short
distractor task of less than thirty seconds so as to remain practically relevant and avoid a long delay between viewing of the original target and that of the lineups.

2.3.5 Consistency between the description of the perpetrator and the appearance of the accused

Psychological research has found that the relationship between recall (i.e. the description of the offender) and recognition (identification of the offender) is at best weak (Thomson, 2003; Wells & Murray, 1983). Early theories proposed that the relationship between recall and recognition could be explained by the memory strength theory, recognition of an item required a lower threshold of strength than recall. [See Anderson & Bower (1972) for a review.] Kintsch (1968; 1965) argued that recognition significantly reduced retrieval difficulties. As such, studies have found that individuals who were accurate at recalling words were generally accurate at recognising them, but those who were accurate at recognising were not necessarily accurate at recall (Anderson & Bower, 1972). This finding was interpreted as indicating that some of the same processes underlay recognition and recall, but recognition was an easier task requiring a lower threshold.

Also relevant is the phenomenon called ‘recognition failure’. Recognition failure describes the finding that previously recalled words are not recognised (Flexser & Tulving, 1978; Tulving & Thomson, 1973). Similar studies have found that those who provide accurate face descriptions cannot necessarily accurately recognise faces (Wells & Murray, 1983). Recognition failure is inconsistent with the contention that recognition is the same process as recall but merely requires a lower threshold (Anderson & Bower, 1972). Flexser and Tulving (1978) argue that an explanation for the recognition failure phenomenon is that retrieval cues in recognition are quite independent of those in recall. More specifically, in recall the individual must retrieve the target information from some context, however in recognition, the individual must retrieve contextual aspects of the original image when the target information is provided. Thus, from the psychological research literature, the consistency between the description of the offender (recalled by a witness) and the appearance of the accused (recognised by a witness) is not likely to reflect increased accuracy.

Further, consistency between a witness’s description of the offender and the appearance of the accused may be for reasons other than reliable memory. First, if a
witness encoded the perpetrator as having a particular feature, then that witness may incorrectly pick an innocent suspect with that particular feature from the lineup. A second reason for consistency between description and appearance of the accused is the exposure of the witness to other witnesses’ reports. Loftus and Greene (1980) investigated the effect of witnesses being exposed to a description from another witness. Their results indicated that witnesses exposed to other witness descriptions which included an incorrect detail were misled into incorporating this detail into their own reconstructions of the original face, whereas those who did not hear this detail did not incorporate it. This memory reconstruction led to the witnesses picking individuals from the lineup who were consistent with that reconstruction. This finding suggests that consistency between the description and appearance of the alleged suspect is not likely to be a reliable indicator of accuracy. Lineups in the current experiments were therefore not constructed according to a participant’s description of the perpetrator and instead similarity was manipulated using similarity ratings from several other participants not involved in the witness identification task.

2.3.6 Eyewitness attention

In order to acquire memories, the visual system must be attending to the target object (Cowan, 1998). Indeed, the relationship between attention and memory dates back over a century (James, 1890). It has been consistently demonstrated in psychological research that attention is critical for effectively acquiring memories [See Underwood (2013) and Cowan (1998) for reviews.] Many things can impair attention, such as alcohol consumption (Givens & McMahon, 1997; Goodwin, 1995), distraction or attending to multiple items (Craik et al., 1996), and stress (Vedhara, Hyde, Gilchrist, Tytherleigh, & Plummer, 2000). All of these variables have also been found to decrease memory performance, supporting the association between attention and memory (Craik et al., 1996; Deffenbacher, Bornstein, Penrod, & McGarty, 2004; Goodwin, 1995; White, 2003). Thus, from the psychological research it would appear that the degree of eyewitness attention at the time of observing the perpetrator would impact upon the accuracy of the eyewitness’s identification. In the present research, eyewitness attention was not measured, but the nature of the online study meant that participants were able to engage in the study at a time that suited them which likely increased the level of attention paid to the
task. If participants lost attention or interest in the task and did not complete the survey, their answers were not recorded and they were excluded from the studies.

2.4 Eyewitness confidence

The reported findings investigating the relationship between confidence and accuracy are conflicting. Some research finds little to no relationship between confidence and accuracy whereas other research finds a moderate effect or a relationship in some respondents and not others.

Studies that have found a significant relationship between confidence and accuracy have tended to use between-subjects designs and very simple tasks, or to manipulate variables that impair or facilitate accuracy (Bradfield, Wells, & Olson, 2002; Brewer & Wells, 2006; Lindsay et al., 1998; Thomson, 2003). For example, Lindsay et al. (1998) manipulated exposure time to a video tape depicting individuals that participants were later asked to identify. The participants who viewed the video for longer were both more accurate in their identifications and more confident in their identification responses. Here, the confidence-accuracy relationship may have been accounted for by the manipulation of exposure time, which would be likely to yield higher levels of both accuracy and confidence independent of each other. Bradfield et al. (2002) examined the effect of post-identification confirmatory feedback on the confidence-accuracy relationship in 245 eyewitnesses. They found that there was a significant relationship between confidence and accuracy, with the correlation being .58. However, the videos participants watched in this study were three minutes and showed the target in a variety of contexts and positions, making the identification task very easy. Therefore, the positive correlation found in Bradfield and colleagues’ (2002) study logically related to a high relationship between actual accuracy and belief in accurate identifications. As such, these findings suggest that confidence may be related to accuracy, particularly for simple tasks. However the study also indicated that post-confirmatory feedback inflated confidence ratings more for inaccurate eyewitnesses than for accurate eyewitnesses. This finding suggests that while related, confidence and accuracy may be independent of one another, as memory accuracy cannot be changed post-event. The larger inflation of confidence for inaccurate compared to accurate eyewitnesses further indicates that confidence and accuracy may be dissociated. In other words, confidence may be higher in less accurate witnesses.
Similarly, Weber and Brewer (2004) examined the confidence-accuracy calibration for absolute and relative judgments across three different exposure durations to create three levels of difficulty. They found that the confidence-accuracy calibration was significantly better for positive decisions than negative ones. However, exposure duration was varied between blocks, rather than between participants and it is possible that participants would have been able to adjust their confidence criteria to produce positive calibrations for positive decisions based on exposure time. This effect would not have occurred for negative decisions, as the manipulation would not influence the familiarity of non-studied faces. Nevertheless there are some studies that report that the calibration between confidence and accuracy is better for choosers, or positive decisions, than for non-choosers in the absence of possible confounders in between-subjects designs (Sauer, Brewer, Zweck, & Weber, 2010; Weber & Brewer, 2006). Indeed, Sporer et al. (1995) found that the average correlation for choosers was moderate, being consistently and reliably higher than that for non-choosers. Taken together these findings suggest that although the overall relationship between confidence and accuracy may be minimal for non-choosers, it may be higher for choosers.

A vast number of studies, both in the general memory literature and the eyewitness identification literature have failed to find evidence for a relationship between confidence and accuracy in between-subjects designs (Brewer, Keast, & Rishworth, 2002; Cutler & Penrod, 1989; Cutler, Penrod, & Martens, 1987; Krug, 2007; Malpass & Devine, 1981a; Penrod & Cutler, 1995; Sporer et al., 1995; Thomson, 2003). For example, Cutler and his associates conducted several experiments where participants watched videotapes of a robbery and then rated their confidence in identifying the perpetrator before and after viewing the lineup and attempting to identify the perpetrator (Cutler & Penrod, 1989; Cutler et al., 1987). They also conducted a meta-analysis of nine studies ranging in sample size from 100 to 320 participants, including five of their own, that all assessed the relationship between pre-lineup confidence and accuracy (Cutler & Penrod, 1989). For pre-lineup confidence, correlations were between 0 and .20, indicating that pre-lineup confidence is not a useful indicator of accuracy. Their studies, however, suggested that post-lineup confidence may be a more useful predictor of accuracy with correlations being slightly higher, ranging between .10 and .45. Many other studies have also assessed the relationship between post-lineup confidence and accuracy.
(Brewer et al., 2002; Brewer, Sampaio, & Barlow, 2005; Brewer, Weber, Wootton, & Lindsay, 2012; Brewer & Wells, 2006; Fleet, Brigham, & Bothwell, 1987; Krug, 2007; Lindsay et al., 1998; Memon et al., 2003; Sauer et al., 2010; Smith, Kassin, & Ellsworth, 1989; Sporer, 1993; Sporer et al., 1995; Weber & Brewer, 2003; Weber & Brewer, 2004, 2006). Comprehensive reviews of these studies have found that confidence is at best a weak indicator of accuracy (Krug, 2007; Wells & Murray, 1984). Indeed, a review of 31 studies investigating the relationship between post-lineup confidence and accuracy revealed a small average correlation of .07 between the two (Wells & Murray, 1984). Similarly, a meta-analysis of 30 studies that used staged event methods in between-subjects designs with sample sizes ranging from 30 to 320 and a collective sample size of 4060 conducted by Sporer et al. (1995), found that the confidence-accuracy correlation was only .29.

Some researchers have argued that one explanation for the low confidence-accuracy correlations may be related to the use of point-biserial correlation, and the confidence-accuracy relationship should be studied in terms of diagnosticity and calibration (Juslin, Olsson & Winman, 1996). However, the weakness of looking at diagnosticity, commonly using the diagnosticity ratio, is that it combines the rate of correct and false identifications into a single number (the correct identification rate divided by the false identification rate), which does not allow researchers to understand the numbers of each occurring, or the confidence in decisions in each separate response type. Further, the researchers using the calibration approach often disregard false identifications of lineup foils from the analysis, which while practically relevant, does not allow a comprehensive analysis of data (Brewer & Wells, 2006). Considering the rate of false identification of foils and confidence related to these decisions could provide important information about the confidence-accuracy calibration. Calibration approaches of assessing confidence and accuracy, however, have generally reported findings consistent with the wider literature, with the confidence and accuracy being calibrated for choosers but not non-choosers (Brewer & Wells, 2006), although some have found that accuracy and confidence may not always be calibrated for choosers (Brewer et al., 2002).

Some studies have found a relationship between confidence and accuracy, under certain conditions. Some report a positive relationship only for choosers, or in other words, participants who make a positive identification (Brewer & Wells, 2006; Fleet et al., 1987; Sauer et al., 2010; Weber & Brewer, 2004, 2006). For example,
Chapter 2 | Factors Influencing Eyewitness Identification: The Psychological Literature

Fleet et al. (1987) conducted a between-subjects study whereby 142 participants who witnessed a staged crime and were asked to identify the perpetrator from either target present or target absent photographic lineups. They were also randomly assigned into three groups, which either received positively biased instructions, negatively biased instructions, or unbiased instructions. The authors reported that confidence-accuracy correlations for choosers were significantly higher, at .50 compared to .14 for non-choosers. No other manipulations affected accuracy, although participants were significantly more confident and there were significantly more choosers in the target present condition. Although accuracy was not significantly higher in the target present condition, the percentage of participants who correctly identified the perpetrator was numerically higher (61% compared to 45%). As such, it is possible that the higher overall correlation between confidence and accuracy in choosers may have been contributed to by the higher level of accuracy and confidence in choosers in the target present condition in Fleet and colleagues’ (1987) study. More recently, Wixted and Wells (2017) suggest that despite many historical findings to the contrary, where confidence is taken immediately after the identification, there are no suggestible influences, and the lineup is completely fair, then confidence may be a useful indicator of accuracy. However, this proposition is difficult to evaluate as it is unclear what suggestible influences, or unfair elements, may have been present in the past research reporting findings to the contrary.

The literature on the relationship between eyewitness confidence and accuracy reveals that eyewitness confidence is highly malleable and a range of non-memory factors can also substantially reduce the relationship between confidence and accuracy (Krug, 2007; Wells & Murray, 1983). These factors have typically been investigated using between-subjects designs. Biased lineup instructions have been found to lead to higher confidence and lower accuracy, in particular more false positive identifications (Malpass & Devine, 1981a). On the other hand, disconfirmation (where witnesses reflect upon encoding conditions and characteristics of their identification, thinking about reasons why their identification might be incorrect) can serve to enhance the relationship (Brewer et al., 2002). Similarly, consistent with data indicating that simple identification tasks enhance the confidence-accuracy correlation, a delay of several weeks before making an identification reduces the diagnosticity of the confidence-accuracy relationship (Sauer et al., 2010). Further, believing that the identification is real and will lead to
the identification of a real criminal has been found to decrease the relationship between confidence and accuracy, but the mechanism by which this decrease occurs is unclear (Wells & Murray, 1983). Possible explanations include that belief that the identification procedure is real may lead to some anxiety which interferes with the confidence-accuracy relationship, or that pressure to identify a suspect may be higher when witnesses believe the identification is in service of a real crime (Wells & Murray, 1983). Studies have also found that confidence can be influenced by factors independent of the quality of memory, and therefore accuracy. For example, repeated questioning of witnesses after the event can increase eyewitness confidence, despite no changes in the amount of information available and thus no possible change in actual accuracy (Hastie et al., 1978; Shaw, 1996). Additionally, many factors, such as suggestiveness, misinformation, supportive negative feedback, biased instructions and post-identification confirmation all artificially inflate confidence ratings independent of accuracy (Brewer & Wells, 2006; Dodson and Dololyi, 2016; Dodson & Dobolyi, 2017; Hastie et al., 1978; Henkel, 2017; Malpass & Devine, 1981a; Shaw, 1996; Thorley, 2015; Thorley & Kumar, 2017; Wells & Murray, 1983). Further, research has suggested that accuracy may be influenced independent of confidence ratings. For example, post-event information provided to witnesses can be integrated into a witness’s memory of the original event leading to decreased accuracy independent of confidence (Köhnen & Brockmann, 1987; Loftus, 2005; Loftus, 1979; Loftus & Greene, 1980).

These findings together reveal two problems for reliance on eyewitness confidence as an indicator of accuracy. First, events or stimuli that make the identification task more difficult, such as recognition delays reduce the relationship between confidence and accuracy significantly. Indeed, it has been proposed that confidence is based on witnesses’ metacognitive beliefs that a recall or recognition is accurate (Brewer et al., 2005). A metacognitive belief is understood to be a belief about one’s own thoughts or cognitions (Brewer et al., 2005). As such, increasing the difficulty of the task would decrease the relationship between accuracy and the metacognitive belief that a recognition is accurate because the witness is less likely to be aware of having made any errors. Second, eyewitness confidence is highly malleable independent of accuracy. Indeed, as mentioned above, there is a great range of factors that change confidence ratings independent of memory factors, and thus accuracy (Hastie et al., 1978; Shaw, 1996; Wells & Murray, 1983). Many of
these factors are likely to be present in the criminal justice system. Standard police procedure involves repeated questioning of witnesses and implicit confirmation of choice (as the suspect goes to trial if the witness’ choice matches the police suspect). These factors are likely to inflate confidence (Bradfield et al., 2002; Shaw, 1996). Similarly, it is often a delay of several weeks or months before the police can identify a suspect and conduct an identification lineup. Given that recognition delay is one factor which diminishes the relationship between confidence and accuracy, based on the psychological literature eyewitness confidence is unlikely to be a good predictor of accuracy in the criminal justice system. The current research aims to further investigate the relationship between confidence and accuracy. Particularly, it will include within-subjects designs to examine this relationship as most of the previous psychological literature has tended to rely on findings from between-subjects experiments only.

2.5 Legal safeguards: Are they supported by research?

Although many of the legal safeguards instigated by the courts are logical, most arise as a matter of legal utility and decisions are made based on the citation of previous court findings, rather than being based on empirical research findings. The USA Supreme Court ruling that eyewitness certainty is an important indicator of eyewitness reliability gains limited support from empirical research - a finding of particular relevance to the rationale of the current research. As such, the following section will outline whether several distinctions made by legal practice, which tends to be based on the interpretation of the law in previous cases within the relevant jurisdiction, are supported by empirical psychological research, which focuses on human behaviour. Issues of the legal distinction between identification and similarity evidence, as well as court rulings relating to system variables, estimator variables, and eyewitness confidence will be addressed.

2.5.1 The distinction between identification and similarity evidence

The distinction between positive identification evidence and similarity evidence, sometimes referred to as resemblance evidence, is made in the law on the assumption that a witness is capable of making both an identity and a similarity judgment, and that these judgments are different processes. Psychological literature on visual processing reveals that judgment of visual similarity is a complex issue
with many problems that computational models have not yet solved (Basri, Costa, Geiger, & Jacobs, 1998). The way identity is judged, however, is even more complex and may depend on several factors such as on the consistency of the present image with previous images, complexity, context and prior knowledge (Basri & Jacobs, 1997).

Judging that an object, or person, that has been seen before is the same as the one that is currently being seen comes with several challenges (Basri & Jacobs, 1997; Bedford, 2001). First, the movement of eyes, head and body, can distort or create a confusing retinal image (Bedford, 2001). In other words, the image changes based on the viewer’s position and movement in relation to the object viewed. For example if the viewer turns sideways while looking at an object, the image changes, so how is the person or object judged to be the same? The literature would suggest that this decision may be made based on the application of knowledge of constraints of the world (Bedford, 2001). Objects do not just change because people move, so they attribute the changed image to the movement. However, it is possible that if the individual was to remain unaware of their movement, the decision may be made in a different way. Second, the objects that are perceived can move, changing the way they look and the way they are perceived (Bedford, 2001). When objects change, in order to make an identity decision, people need to be able to both perceive the changes to the object and know that the object is the same one. Finally, identifying objects is complicated if there is a lack of continuity of sensory information (Bedford, 2001). For example, if an object were to transform before someone’s eyes, they would accept it as the same object, no matter how bizarre the transformation, however if the person were to see the object, look away, and then view the bizarrely transformed object, they would not have the sensory information to determine that the transformed object is the same one - no matter how ordinary the transformation.

The importance of object identity in our understanding of the world around us is apparent when paradoxes of visual perception, such as apparent motion, an optical illusion in which a series of static images on a screen can create the illusion of movement, are considered (Bedford, 2001). Indeed, the decision of object identity is important for following objects through time, separating different objects from each other and experiencing a coherent world. Importantly, however, the object identity decision is often not solely based on the perception of incoming sensory information, but often the gaps in our sensory perception are filled based on our knowledge about
the world around us. The implications of this visual processing literature for eyewitness identification is that often a witness does not have the necessary information to determine identity. Usually, there has been a period of intervening time between the witness seeing the alleged perpetrator commit the crime and identifying the suspect in a completely different context at a police lineup. In the intervening time the visual characteristics of the suspect may have been changed in a number of ways, for example, their clothes, hair colour or even weight may have changed. Additionally, changes in the context in which the person is viewed may change how the person is perceived. If human identity decisions are based on our prior knowledge about the world and on the context in which the target is seen, then in the context of a police lineup, several weeks or months after the original viewing of a previously unfamiliar suspect, a witness does not have the information required to make an identity decision (Basri et al., 1998; Basri & Jacobs, 1997; Bedford, 2001). The witness has not been able to track the suspect over time viewing the ways in which their appearance may have changed, and without prior knowledge of the suspect’s appearance to inform their decision, an identity decision becomes unfeasible. In other words, the witness has no information on which to base a decision about whether the suspect they are viewing and the person they viewed previously are the same person. This is particularly true in the common situation that a witness is unfamiliar with the perpetrator and has no prior knowledge other than the original viewing to apply to the situation. This puts into the question the value of the identification decision - after all if a witness is not able to make such a decision, requiring them to make that decision may be of limited utility. Further, the outcome of requiring this decision may be that processes unrelated to identity are contributing to the ultimate decision. The value of identification decisions has also been questioned by recent data finding that allowing witnesses to rate their confidence that each member in the lineup was the culprit under short deadline conditions provides a better indication of identity than simple yes/no judgments (Brewer et al., 2012).

The above research demonstrates that a simple identification provided by a witness is not complex enough to capture all the desired information. Instead, a more nuanced response system that takes into account the extent to which each lineup member is recognised by the witness (a confidence judgment, which is reminiscent of a visual similarity judgment), may be more diagnostic of accuracy, and therefore more informative. It is therefore likely that the best a witness can do is
to make a judgment about how much a suspect in a lineup resembles, or is similar to, the person seen at the scene of the crime. As such, the distinction the criminal justice system makes between visual similarity and identity, may be an artificial one. Indeed, it has been suggested that witnesses only be allowed to make a judgment about the degree of visual similarity of lineup members, whereas the issue of identity should be left up to the courts (Thomson, 2003). In other words, a witness would be allowed to select any person, or multiple people, from the lineup who they judge to look similar to the offender and this information would then be provided to the courts. The idea of selecting multiple persons from the lineup is novel and could be useful. However, thus far the criminal justice system has failed to acknowledge the possibility that the legal distinction between visual similarity and identity evidence may be flawed.

2.5.2 System variables

The common law rulings to exclude identifications made when the suspect is viewed by the witness alone, or in the dock at their trial without any foils (Davies and Cody v R, 1937) are consistent with evidence in the empirical psychological literature that increasing the number of foils increases protection for the innocent suspect (Wells et al., 1998). Further, the USA common law ruling to dismiss identifications made where the procedures are unnecessarily suggestive, such as when a suspect is depicted in prison clothing or other circumstances suggesting the guilt of the suspect (Alexander v R, 1981; R v Wainwright, 1925; R v Williams, 1983), is supported by psychological research experiments showing the susceptibility of witnesses to suggestion and social normative pressure (Loftus, 2005).

There are no specific legal guidelines concerning instructions given to the witness when the witness views the lineup in the USA, UK or Australia, and, as such, the words used in such a context differ across jurisdictions within these countries. Given that there are no specific guidelines, instructions may also differ between lineup administrators within the same jurisdiction. Often, the exact wording of the instructions remains unknown to the jury as it is not mandatory to video tape eyewitness lineups. Videotaping the lineup process so that instructions given to witnesses and the fairness of lineup procedures can be assessed by jurors has been recognised by eyewitness identification researchers as a desirable procedure (Wells
et al., 1998). However, no legal rulings or safeguards governing lineup instructions to eyewitnesses have been put in place.

Similarly, while the clear preference of the courts in Australia for live identification parades is apparent (Alexander v R, 1981), there is no legal guidance on the type of lineup that is preferred (Evidence Act, 2008). This is understandable given that this subject is still contentious within the empirical literature, although the influence of these different lineup procedures on the likelihood of correct identifications, and in particular, false identifications could ultimately have some utility to the legal system. Further there is generally no guidance on how foils should be selected for lineups, whether they should be selected in terms of visual similarity to the suspect, other foils, or witness descriptions of the perpetrator, or in the selection of foils, who determines visual similarity of the foils. As such, these factors may be considered either irrelevant to the judgment of eyewitness reliability, or factors upon which the jury can make judgment without any guidance from the court.

2.5.3 Estimator variables

Several estimator variables have been cited in the decisions of judges in cases tried in both the USA and UK courts as important indicators of eyewitness reliability (Neil v Biggers, 1972; R v Turnbull, 1976). Thus legal safeguards have been dictated by these decisions in that it was determined that judges and jurors must turn their mind to several estimator variables when assessing the reliability of eyewitness identification evidence. In both legal jurisdictions the opportunity of the witness to view the offender, the consistency between the description given by the witness and the appearance of the accused, and the length of time between the offence and the eyewitness identification have been named as indicators of eyewitness reliability. However, there are some differences between the factors identified by judges in courts in the USA and the UK. Namely, the factors mentioned by judges in courts in the UK revolve mostly around the circumstances, or prevailing conditions, at the time the witness observed the offender. As such, they specifically mention quality of the observation in terms of obstructions and lighting conditions that prevailed at the time of viewing (R v Turnbull, 1976). These judge’s directions also mention familiarity of the suspect to the witness as a factor to consider which the USA courts do not. On the other hand, the USA courts identify the witness’ degree of attention
and certainty as factors relevant to eyewitness reliability which are two factors the courts of England and Wales do not mention. In particular, the Devlin Inquiry (1976) in the UK came to the opposite conclusion about eyewitness certainty, and instead reported that juries should be warned about putting too much emphasis on this factor. As such, the extent to which psychological research findings support the legal safeguards put in place by the decisions of judges to consider these estimator variables as indicators of eyewitness identification reliability is considered.

2.5.3.1 Duration of observation

The length of time that the witness viewed the offender at the time of the crime is the first factor that the decisions of both R v Turnbull (1976) and Neil v Biggers (1972) identified as important for the jury to consider. Indeed, it has previously been a concern of the courts that in cases where the witness’s encounter with the offender is fleeting, there is a substantial risk of misidentification (R v Oakwell, 1978). The presumption that fleeting encounters are more likely to lead to misidentification than lengthy observations of suspects has been tested in empirical psychological research and these findings generally support the view expressed in legal court cases (DiNardo & Rainey, 1991; Memon et al., 2003; Reynolds & Pezdek, 1992; Shapiro & Penrod, 1986; Thomson, 2003; Wells & Murray, 1983). Psychological research generally shows that exposure duration is significantly related to facial identification accuracy (Memon et al., 2003; Reynolds & Pezdek, 1992; Shapiro & Penrod, 1986). However, it must be noted that this consideration would only be relevant under conditions where exposure time can be objectively measured: Asking a witness to retrospectively estimate the length of time is subjective and potentially unreliable as people are not always accurate estimators of time (Thomson, 2003).

2.5.3.2 Obstacles to viewing, distance from the offender, and lighting conditions

Prevailing conditions when the witness originally viewed the perpetrator are important when assessing eyewitness identification evidence. If witnesses are so obstructed that they cannot see the offender clearly, critical details may be unable to be observed resulting in the witness being unable to make an identification or being unable to discriminate accurately between similar looking persons. Psychological research finds that obstacles to viewing, increased distance from the offender, and
poor lighting conditions all decrease identification accuracy (Malpass & Devine, 1981a; Thomson, 2003; Yarmey, 1986). Wagenaar and Van Der Schrier (1996) tested seven distances (3 to 40 metre) and nine illumination levels (0.3 to 3000 lux), and found a systematic increase in recognition performance with decreasing distance and increasing illumination. These collective findings suggest that prevailing conditions at the time a witness observes a perpetrator of a crime can impair eyewitness identification accuracy. Thus psychological research is consistent with views expressed by courts that such prevailing conditions should be taken into account when assessing the reliability of eyewitness identification evidence (R v Turnbull, 1976).

2.5.3.3 Length of time between the offence and identification

Logically, the more time that has elapsed between the original observation of the perpetrator and the recognition at the lineup, the less accurate the recognition of the witness (Thomson, 2003). Both the psychological literature relating to memory (Altmann & Gray, 2002; Pavlik & Anderson, 2005) and eyewitness identification (Deffenbacher et al., 2008; Malpass & Devine, 1981a; Shapiro & Penrod, 1986; Thomson, 2003) suggest that increasing time between observation of a target and subsequent identification of the target decreases identification accuracy. Generally memory decay is observed over the subsequent 11 to 16 seconds following observation of a target (Alvarez & Cavanagh, 2004; Curby & Gauthier, 2007; Hollingworth, 2004; Phillips, 1974; Phillips & Baddeley, 1971). Further, the decline in accuracy of descriptions and recognition is most marked within minutes, and then is relatively modest over short delays of minutes to days, but accelerates over weeks to months (Deffenbacher et al., 2008; Malpass & Devine, 1981a; Shapiro & Penrod, 1986; Thomson, 2003). Further, with increasing time between observation and identification, the chance that intervening events susceptible to being incorporated into a witness’s original memory of an event will impair accuracy is increased (Loftus & Greene, 1980). Thus the court’s ruling that the length of time between the offence and the identification should be taken into account when assessing the reliability of eyewitness identification evidence is solidly supported by psychological literature.
2.5.3.4 Consistency between the description of the perpetrator and the appearance of the accused

Both courts in the UK and USA identified the consistency between the description of the suspect provided by the witness and the appearance of the accused as important factors to consider when assessing eyewitness reliability (Neil v Biggers, 1972; R v Turnbull, 1976). However, there is little empirical evidence to suggest that this is related to identification accuracy.

The psychological literature indicates that the relationship between recall and recognition is weak due to different retrieval cues underlying these memory processes (Flexser & Tulving, 1978). Thus, recall and recognition appear to be independent from one another and cannot be assumed to be related. Psychological experiments have found that people who have recalled words and faces do not necessarily recognise them (Flexser & Tulving, 1978; Tulving & Thomson, 1973; Wells & Murray, 1983). Furthermore, psychological research has indicated that consistency between a witness’s description of an offender and the appearance of the accused may be for reasons separate from memory retention, such as incorrect encoding, and exposure to other witnesses’ descriptions (Loftus & Greene, 1980). These latter findings suggest that, if anything, courts should warn juries about the dangers of assuming that consistency between a description of the offender and the appearance of the accused has the potential to be misleading. Thus, the UK and USA court rulings that consistency between the description of the offender and the appearance of the accused reflects increased accuracy has little support from relevant empirical findings in the psychological research literature.

2.5.3.5 Eyewitness attention

The USA Supreme Court specifically mentions a witness’s degree of attention when they first observe an offender, whereas the courts of England and Wales were silent on this matter (Neil v Biggers, 1972; R v Turnbull, 1976). Psychological research demonstrating the relationship between attention and memory has consistently demonstrated that the two are interrelated in that decreased attention can impair memory (Cowan, 1998). Thus, considering a witness’s degree of attention when assessing the reliability of eyewitness identification evidence is appropriate based on psychological research. It has also been suggested that the type of attention also matters when making a judgment about accuracy (Wells & Quinlivan, 2009).
Paying attention to specific facial features may take a great deal of time and effort, and be useful for reconstructing faces, but not for recognising faces, whereas global attention to the entire face may be useful for recognition but not reconstruction of faces. However, while it would be logical to consider the extent to which a witness was paying attention at the time of observation, how accurately a witness’s degree of attention at the time of an observed event can be measured or ascertained is problematic. As others have argued, attention is a purely psychological variable that cannot be checked against any objective facts in a criminal case (Wells & Quinlivan, 2009). Thus, while the instruction of the USA courts that attention is relevant to assessing a witness’s testimony receives good support in psychological research, unless that degree of attention can be objectively established, the instruction has little value.

2.5.4 Eyewitness confidence

The USA courts and courts of England and Wales have very different views on eyewitness certainty (Neil v Biggers, 1972; R v Turnbull, 1976). The assumption implicit in the USA court ruling is that eyewitness certainty is positively related to the accuracy of the eyewitness identification. Certainly, a number of judges and courtroom officials believe that subjective confidence of a witness is a good indication of their testimonial accuracy (Wise & Safer, 2004). However, the view of the judges in R v Turnbull (1976) was that eyewitness certainty should not be considered when assessing reliability of eyewitness identification evidence, as it was not mentioned as an important factor. This was due to their reference to the Devlin Report (Devlin, 1976) which warned against putting too much emphasis on eyewitness certainty.

There is little evidence to support a consistent relationship between confidence and accuracy in the empirical psychological literature. While some research indicates that there is a moderate or strong positive relationship between accuracy and confidence particularly for choosers (Sauer et al., 2010; Sporer et al., 1995; Weber & Brewer, 2006; Wixted & Wells, 2017), a number of studies in the psychological literature relating to memory and the eyewitness identification have failed to find evidence for such a relationship (Brewer et al., 2002; Cutler & Penrod, 1989; Cutler et al., 1987; Krug, 2007; Malpass & Devine, 1981a; Penrod & Cutler, 1995; Sporer et al., 1995; Thomson, 2003). Furthermore, the psychological literature
suggests that eyewitness confidence is highly malleable and susceptible to influence by a range of factors independent of memory, and this influence can substantially reduce the relationship between confidence and accuracy (Hastie et al., 1978; Krug, 2007; Malpass & Devine, 1981a; Shaw, 1996; Wells & Murray, 1983). Several of these factors such as repeated questioning of witnesses, implicit confirmation of choice, and a delay between observation and recognition are likely to be present in the criminal justice system. As such, the empirical psychological literature suggests that in a real lineup situation, confidence is not only likely to be a questionable predictor of accuracy, but several factors, such as repeated questioning and confirming feedback inherent to the legal system can lead to inflation of eyewitness confidence (Bradfield et al., 2002; Hastie et al., 1978; Shaw, 1996).

Despite the considerable research that has been conducted on eyewitness confidence and how it relates to accuracy, very little consideration has been given to the differences in confidence within individuals as opposed to that between individuals. Nearly all studies within the eyewitness literature examine the between-participant effects as they question whether confident witnesses are more accurate than hesitant witnesses (Sporer et al., 1995). As already mentioned, the relationship between confidence and accuracy overall appears to be negligible. However, there is substantial variation between eyewitnesses in terms of confidence (Smith et al., 1989). It might be expected that the relationship between confidence and accuracy may be higher when considering variations in confidence within a certain individual. In other words, a person’s confidence may be higher when they give accurate responses relative to their confidence when they are inaccurate. There is some support for this proposition in the general memory literature. These studies show that feelings of knowing are significantly related to recognition accuracy of non-recalled items on general knowledge recognition tests (Hart, 1965; Nelson, Gerler, & Narens, 1984). Further, research on text recognition shows similar results, with subjective confidence after text recall being significantly related to accuracy (Stephenson, Wolfgang, & Brandstatter, 1983). Additionally, while studies in this domain show that there is substantial individual variation, the within-subjects correlation between confidence and accuracy judgments is fairly high (Stephenson, 1984). However, eyewitness researchers have rarely assessed confidence within participants. The little research that has assessed confidence within participants has suggested that, like between-subjects confidence, within-participants confidence
only very weakly correlated with accuracy, even when taking into account people of high and low confidence separately and in the absence of floor and ceiling effects (Smith et al., 1989). Nonetheless, given the wealth of research demonstrating that the within-subjects confidence-accuracy correlation is high in the general memory literature, it may be that within-participant confidence is a better predictor of accuracy than between-participant confidence, particularly in certain witnesses. Witnesses whose confidence levels have a range (rather than those whose confidence is systematically low or high) may demonstrate a greater confidence-accuracy relationship in their responses.

The predictive value of within-participant confidence, however, has limited utility within the practical context of the criminal justice system. A jury only has one opportunity to observe the confidence of a witness. The jury does not have an opportunity to observe a witness’s confidence under other circumstances. For example, a jury’s opinions about the relationship between confidence and accuracy may be changed if the jury were to observe that the witness was not only highly confident about their identification, but also about all other responses or decisions that they made. Without this information, the confidence exuded by witnesses in making identifications may be quite misleading.

Very little research has explored how various system variables may affect the confidence exuded by witnesses when making their identifications, and the relationship between confidence and accuracy. In an electronic search of five popular databases (PsycInfo; Medline Complete; Web of Knowledge; Academic Search Complete; and SCOPUS) using the search terms “eyewitness confidence and accuracy”; “eyewitness confidence and lineup type” and “eyewitness confidence”, 687 articles were revealed. Of these articles, 158 pertained to the examination of the confidence-accuracy relationship, and out of the 158 articles only two included lineup type as a variable (Dobolyi & Dodson, 2013; Sporer, 1993). Of these, only one study specifically examined the impact of lineup procedure on the confidence-accuracy relationship (Dobolyi & Dodson, 2013). This study, however, did not examine the impact of other system variables such as target position in lineups. Further, the same search revealed that little research has been dedicated to examining the influence of system variables, such as lineup procedure and lineup composition, on confidence, although the little research that has been conducted conflicts. One study suggests that lineup procedure has little impact on confidence, at least for
choosers, or people who identify a lineup member (Sporer, 1993), whereas the other relevant study suggests that sequential lineups may lead to overconfidence in false identifications when compared to simultaneous lineups. If the courts propose to rely on confidence as an indicator of accuracy, examining any variables which may systematically influence confidence would be relevant to understanding the relationship between confidence and accuracy. As such, more research examining the nature of this relationship and the variables that may affect it is needed to assess the USA court’s proposition that eyewitness confidence should be taken into account by juries when assessing the reliability of eyewitness identification evidence - a proposition that currently remains insufficiently supported by the empirical psychological literature. A major aim of the current research was to examine the influence of several system variables on patterns of accuracy and confidence in experiments using within-subjects designs: lineup type, photograph similarity, and target position.

2.5.4.1 Eyewitness confidence and jury decision making

Extensive experimental psychological research suggests that eyewitness confidence may not be a particularly good predictor of accuracy and thus reliability of testimony (Krug, 2007; Wells & Murray, 1984). However, studies also show that judges and juries find eyewitness confidence a compelling factor when making decisions relating to guilt or innocence (Fox & Walters, 1986; Lindsay, Wells, & O'Connor, 1989; Wells et al., 1979; Wise & Safer, 2004). Wise and Safer (2004) conducted a survey of 160 judges in the USA and found that despite their extensive experience, judges averaged only 55% on a 14-item knowledge scale about eyewitness factors, and only 32% of judges disagreed with the statement that confidence is a good predictor of accuracy. These findings provide evidence that USA judges do not necessarily have a good knowledge of psychological factors relating to eyewitness identification decisions, particularly in relation to beliefs regarding the confidence-accuracy relationship.

Similarly, empirical psychological studies suggest that members of a jury are likely to rely heavily on eyewitness confidence when deciding upon a verdict (Fox & Walters, 1986; Lindsay et al., 1989; Wells et al., 1979). Wells et al. (1979) asked witnesses to a staged crime to identify a thief from a six-person photograph array. While there was no relationship between witness confidence and accuracy, witness
confidence accounted for 50% of the variance in jurors’ decisions to believe witnesses. Further, Lindsay et al. (1989) showed 178 participants videotapes of 16 eyewitnesses to a staged crime being questioned by either experienced or inexperienced lawyers about their identification. They found that participants were more likely to believe witnesses who were perceived as confident in their decisions, regardless of the expertise of the lawyer, and that confidence was not related to witness accuracy. Indeed, eyewitness credibility has consistently been found to be highly dependent on eyewitness confidence (Whitley & Greenberg, 1986), even when expert testimony of witness unreliability is given (Fox & Walters, 1986). These findings suggest that in cases where the guilt or innocence of a suspect relies solely on eyewitness identification evidence, the jury may find a confident witness detrimentally persuasive. As such, the investigation of the relationship between confidence and accuracy, and variables which may influence confidence, would be pertinent to educating jurors of the potential strengths and dangers of relying on eyewitness confidence when engaging in legal decision making. All three experiments in the current investigation aimed to examine the relationship between confidence and accuracy in identification decisions.

2.6 Summary

Previous findings within the empirical psychological literature have revealed several variables that impact upon eyewitness accuracy and confidence, such as lineup type, target position, and similarity between observations of the suspect. These variables have commonly been investigated through experiments approximating real life situations in which a crime is simulated and several witnesses to the staged event make an identification. Thus, the vast majority of the psychological literature on eyewitness identification has used between-subjects designs and investigated few variables at one time due to experimental constraints of recruiting vast numbers of participants and staging more than one crime. For this reason, interactions between variables that may impact on accuracy and confidence, such as target position, lineup type, and similarity, have not commonly been explored. Further, although lineup type has been clearly implicated as affecting identification accuracy, a cohesive theoretical explanation of the way in which this occurs has not been agreed upon. The current investigation aims to provide a thorough examination of the ways in which lineup type, similarity, and target
position may impact on accuracy and confidence using both within- and between-subjects designs. Such investigation will aid in developing a more cohesive theoretical understanding of the way in which lineup judgments are made. It will also inform practical legal policy relating to lineup administration by adding to the current understanding of the extent to which variables commonly at play in the administration of lineups, such as lineup type, similarity, and target position, interact to impact upon eyewitness accuracy and confidence.
CHAPTER 3
Experiment 1: Target Present Lineups

The literature outlined in previous chapters demonstrates links between system and estimator variables that contribute to a lineup, eyewitness misidentification and wrongful convictions. More specifically, it is well known that wrongful convictions have disastrous consequences for the individuals they affect, and mistaken eyewitness identification plays a major role in this phenomenon (Colvin, 2009; Huff, 2004; Leo & Gould, 2009; Weathered & Blewer, 2009). Many aspects of the police lineup, or system variables, are shown in the literature to affect eyewitness accuracy and some of these have been recognised by the courts, whereas others, such as eyewitness confidence, remain disputed (Alexander v R, 1981; Manson v Brathwaite, 1977; R v Burchielli, 1981; R v Clarke, 1997; Wells & Murray, 1983; Wells et al., 1998). In particular, there is some argument in the literature as to whether sequential or simultaneous lineups are superior in terms of eyewitness accuracy (Lindsay & Wells, 1985; McQuiston-Surrett et al., 2006; Steblay et al., 2001; Steblay, Dysart & Wells, 2011; Zimmerman et al., 2006). Further in their review of the eyewitness literature, McQuiston Surr$et$ et al. (2006) reported that some studies do not investigate the effects of target position or counterbalance target position in the lineup which may contribute to observed differences between simultaneous and sequential lineups. These collective findings have raised some disagreement as to the mechanism behind the sequential superiority effect observed in the literature, particularly whether the observed effects are due to relative versus absolute judgments, or a higher criterion threshold in sequential lineups as posited by signal detection theory (Clark, 2005, 2012; Clark, et al., 2011; Flowe & Ebbesen, 2007; Palmer & Brewer, 2012; Wells, 1984). Clarifying the pattern of accuracy, in terms of correct identifications, false identifications, incorrect rejections and correct rejections of the lineup in sequential and simultaneous lineups would be useful in elucidating the process of decision making underlying the sequential superiority effect.

Decision makers in the courts have also recognised a number of variables as important to consider when assessing the reliability of eyewitness identification evidence (Neil v Biggers, 1972; R v Turnbull, 1976), however, the role of many of these variables has not been supported in the empirical psychological literature (Thomson, 2003; Wells & Murray, 1983; Wells & Murray, 1984). More specifically,
the USA Supreme court’s presumption that eyewitness confidence is a reliable indicator of eyewitness accuracy is heavily contested within the empirical psychological literature, with this effect only being shown under certain circumstances (Krug, 2007; Sporer et al., 1995). Additionally, confidence is highly malleable and many factors inherent to the criminal justice system influence witness confidence independent of accuracy: repeated questioning, biased lineup instructions, suggestibility, exposure to other witness reports, exposure to mug shots, or exposure to incorrect details subsequent to eyewitness identification taking place (Brewer & Wells, 2006; Dobolyi & Dodson, 2013; Hastie et al., 1978; Shaw, 1996; Wells & Murray, 1983). Very little literature has comprehensively addressed the way in which different lineup procedures may affect patterns of confidence and accuracy, while also investigating potential impacts and interactions with other relevant system variables. If the courts, particularly the USA Supreme court, wish to use eyewitness confidence as an indicator of accuracy, then any variable that may impact upon confidence, including lineup procedure and target position, would be important to investigate. Thorough evaluation of the effect of lineup procedure on the relationship between confidence and accuracy is essential to objectively examining the USA Supreme Court’s ruling on eyewitness confidence.

With the preceding background in mind, the primary aim of the current investigation was to examine the relative efficacy of simultaneous and sequential lineups using a memory paradigm as a function of the position of the target or target replacement, and as a function of the similarity of the photograph of the target or target replacement in the lineup to the target photograph originally seen. Second, the present experiments also aimed to investigate how differences in accuracy between lineup procedures aligned with differences in confidence under the aforementioned conditions in order to examine the USA Supreme Court’s evaluation that eyewitness confidence is a reliable indicator of eyewitness accuracy.

Thus, Experiment 1 investigated patterns of accuracy and confidence in sequential and simultaneous lineups where the target was present and where either the same or a different photograph of the target person was presented either early or late in the lineup. It was hypothesised that:
Chapter 3 | Experiment 1: Target Present Lineups

1. Evidence of a sequential superiority effect would be observed, at least in terms of a reduced occurrence of false identifications in sequential compared to simultaneous lineups.

2. Visual similarity between the originally observed photograph and the photograph observed in the lineup in target present lineups would impact upon accuracy and confidence. In other words, viewing the same photograph of the target person in the lineup as that originally observed would lead to more correct identifications and less incorrect rejections of the lineup as well as higher confidence in correct identifications compared to viewing a different photograph of the target person.

3. There would be higher confidence in correct than false identifications, and there would be higher confidence in correct than incorrect rejections of the lineup but to a lesser extent than the difference seen between correct and false identifications. Further if the proposition of the USA Supreme courts is accurate, then: 1) confidence in making correct identifications would be higher than confidence in making false identifications; 2) patterns of accuracy would align with patterns of confidence such that under conditions where increased accuracy was observed, higher confidence would also be observed. If the proposition of the courts in the UK is more accurate, then either the level of confidence in correct identifications would not differ from that in false identifications, or conditions where increased accuracy was observed would not align with conditions where higher confidence was observed, or both would occur.

3.1 Method

Experiment 1 was part of a larger study and as such data collection was undertaken jointly with several other students, after which the author analysed a subset of the collected data. Further, due to practical constraints related to the number of lineups that were able to be administered to participants, only lineups where the target was present were included in Experiment 1. A discussion of each set of results for all experiments in the current research will be included with that set of results and a general discussion of all the findings will be presented in the final chapter. Additionally, being part of a larger study Experiment 1 included several
variables such as the ethnicity and gender of members of the lineup, which were not the focus of the research to be reported in the present investigation. These variables were not included in the design or analyses and will not be reported on or discussed any further.

3.1.1 Participants

Participants were 546 (348 women) volunteers aged 18-70 years, with a mean age of 28.07 years (SD= 10.39 years). Participants were recruited through an advertisement to complete an anonymous online survey involving photographic lineup judgments. Participants were excluded from the study if they did not complete the survey. On the basis of this criterion, 387 participants were excluded out of 933 that volunteered for the study, leaving an effective sample of 546. The protocol conformed to the Declaration of Helsinki and had prior approval of the local Human Subjects Ethics Committee. All participants gave informed consent online prior to their participation in the survey. A copy of the advertisement and the online plain language statement with consent checkbox can be seen in Appendix A. Participants were given no incentive to participate in the study. Procedures were approved by Deakin University’s Human Ethics Committee (#DUHREC-H92_2013; Appendix B).

3.1.2 Design

Patterns of accuracy and confidence were studied using a 2 (sequential vs simultaneous lineups) x 2 (same vs different photographs) x 2 (early vs late position of the target photograph). Lineup procedure was the between-subjects factor, and photograph condition as well as target position were within-subjects factors. The dependent variables were participant accuracy (in terms of hits or correct identifications of the target, misses or incorrect rejections of the target, and false positives or incorrect identification of a lineup foil) and confidence ratings in the decisions (confidence in hit, miss and false positive responses), measured on a 5-point Likert scale. The target person was present in all lineups.

3.1.3 Procedures

Participants carried out the anonymous online survey on their own computers in their own time. The online survey utilised the Qualtrics™ (Provo, USA) survey
building computer software. Participants clicked through the survey, answering each question when prompted until the survey was complete, or they exited from the survey prematurely. (See Appendix C for an example of the survey. This example is an internet copy of the survey downloaded directly from the website and thus retains that format.) Participants were not allowed to continue the survey if they had missed an item or not answered one of the questions.

The online survey consisted of eight lineups and participants were randomly allocated to either sequential or simultaneous lineup procedure conditions. Before completing the survey, participants completed two practice lineups. For each lineup, participants were asked to look at a target photograph displayed on the computer screen for five seconds. The target photograph was followed by a photograph of another person for five seconds. Participants were asked to estimate the age of the person in the second photograph and then to estimate how confident they were on a 5-point Likert scale (one being “not confident at all” and five being “extremely confident”) of their age estimate. The ethnicity and gender of the person in the second photograph differed from that of the target person. Participants were then presented a lineup of eight photographs, one of which was the target person and asked to identify which member of the lineup was the target person or indicate if they thought the target person was not present. Specifically they were given the following instructions upon viewing the lineup: “If you are able to identify the target person, please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs”. In accordance with the literature, this wording was utilised in order to clarify that the target person may not be present in the lineup. Participants were allowed as much time as they wanted to respond. In the simultaneous lineup procedure condition, all photographs in the lineup were displayed on the screen at the same time along with a “not present” option. In the sequential lineup procedure condition, the photographs were displayed on the screen one at a time. If the response was “no”, the next photograph was presented until all eight photographs had been shown. If the response was “yes”, the lineup was discontinued. In the sequential condition, participants were not told how many possible photographs were in the lineup. After each lineup decision, participants were asked to rate on a 5-point Likert scale how confident they were as to the accuracy of their answer. At the conclusion of the survey, participants completed demographic information, including age, gender, and ethnicity before
being prompted to quit the survey. The length of time taken to complete the survey varied depending on how quickly the participant made the relevant choices. On average, participants took 11.15 (SD= 6.87) minutes to complete the survey when viewing simultaneous lineups and 16.02 (SD= 11.73) minutes when viewing sequential lineups.

3.1.4 Materials and Apparatus

3.1.4.1 Photographs

Photographs were all 100 x 149 pixels (180 dpi horizontal and vertical resolution) close up, eye level, well lit, camera shots of a person’s face. The photographs were drawn from CMU Multi-PIE Face Database (Gross, Matthews, Cohn, Kanade, & Baker, 2010) and included male and female Caucasian and Asian photographs. (See lineup construction, section 3.1.4.2 below for an example of a male Caucasian lineup.) All lineup photographs had plain blue backgrounds with the face in the foreground (Figure 3.1). Target photographs either had plain blue backgrounds, or a high backed chair in the background with the face in the foreground. For the simultaneous lineup procedure condition, lineups consisted of eight photographs of faces displayed in one line directly next to each other in the centre of the computer screen alongside a “not present” option (Figure 3.1). For the sequential lineup procedure condition, eight photographs were displayed one at a time at approximately eye level in the centre of the computer screen immediately to the left of a “not present” option (Figure 3.1).

3.1.4.2 Lineup composition

Each lineup comprised eight lineup photographs. The number of eight lineup members was chosen due to suggestions in the literature that there should be more than five foils, and also due to practical limitations relating to the photographic database of including more than eight lineup members.

A target was selected for each lineup and then eight other photographs were selected on the basis of being visually similar to the target. For instance, if the target was a blond Caucasian male with short hair, the foil photographs were selected to match these features. Photographs containing distinctive traits such as glasses or facial hair were excluded from the lineups. Two different photographs of the target were used that were taken at different time points. The target photograph initially
Presented to participants differed in terms of background and clothing to that presented in the lineup. A total of 10 lineups were constructed which included two practice lineups.

Half of the target photographs were photographs of males and half were females. Half of the male and female target photographs were Caucasian and half were Asian. The gender and ethnicity of lineup photographs matched the gender and ethnicity of the target person. A panel of four university students (three Caucasian and one mixed race) classified all photographs into “Caucasian” or “Asian” categories. To be included in lineups, all four students had to agree on the ethnicity of the person in the photographs. Half of the photographs of the target person displayed in the lineups were the same as the target photograph previously observed (Figure 3.1), and the other half were a different photograph of the target person to the photograph of the target person previously observed.

Target photographs were counterbalanced in terms of target position, such that half of the targets were displayed early in the lineup (at position 2 or 3) and half were displayed late in the lineup (at position 6 or 7). Presentation order of the eight lineups was randomised.

Figure 3.1. Example of photograph displays for simultaneous (top pane) and sequential (bottom pane) identification decisions. Photographs taken from Gross and colleagues (2010).
3.1.5 Data reduction & statistical analyses relevant to the entire thesis

The alpha level for all analyses throughout the thesis was set at .05. All effects for accuracy were assessed for normality and tested for equality of variances using Box’s M equality of variances test or Levene’s homogeneity of variances test. When parametric assumptions were violated, main and interaction effects were assessed using appropriate nonparametric equivalents (Mann-Whitney U or Wilcoxon signed-rank tests) where possible. Further, where t-tests were performed and the equal variances assumption was not met as tested by Levene’s equality of variances test, adjusted tests with adjusted degrees of freedom are reported. Where post hoc testing was carried out, Bonferroni adjustment to the significance threshold value was utilised to test significance.

Violations of the parametric assumptions were noted in several places throughout the thesis. It must be noted, however, that no appropriate nonparametric test was capable of investigating all of the within and between-subjects factors included in the current experiments and their relative interaction effects. Thus, for conservativeness, parametric tests were conducted and reported alongside the nonparametric tests for the effects that it was possible to do so. This was to confirm consistency between the ANOVA and nonparametric equivalents, and it should be noted that these analyses were consistent with the few exceptions noted below. However, given that ANOVA has been found to be robust to violations of equality of error variances, results from the parametric analyses will be interpreted (Tabachnick & Fidell, 2007).

3.1.6 Data reduction & statistical analyses for Experiment 1
3.1.6.1 Accuracy

Participant answers for all eight lineups were collected via Qualtrics™ (Provo, USA) and participant means for accuracy (number of hits, misses and false positives) were collated. In order to examine patterns of accuracy, three 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph condition) x 2 (target position early vs late) mixed model ANOVAs were conducted to examine differences in the number of hits, the number of false positives and the number of misses. Paired samples t-tests were carried out post hoc to assist interpretation of any interaction effects.
3.1.6.2 Confidence

Participant answers were collected via Qualtrics™ (Provo, USA), participant ratings for confidence in hit, miss and false positive responses were collated and mean ratings were calculated. Similar analyses to those conducted for accuracy measures were conducted for confidence ratings on hits, false positives and misses, namely, three 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph condition) x 2 (target position early vs late) mixed model ANOVAs. In addition, differences in confidence between hit and false positive responses, were assessed using a 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph) x 2 (hit vs false positive responses) x 2 (target position early vs late) mixed model ANOVA. These analyses were conducted to allow investigation of the premise that a person would be more confident in their correct identification than their incorrect identifications, as posited by the USA Supreme Court’s proposition that confidence is a reliable indicator of accuracy. Paired samples t-tests were carried out post hoc to assist interpretation of any interaction effects.

For the first three analyses, only participants who had responses appropriate to the analysis (hit, miss or false positive) in both the same and different photograph condition where the target was presented early and where it was presented late, were included in the analyses. Where this was not possible due to small numbers of responses in all four cells, two analyses were run to investigate the relevant effects - one with photograph condition as the within-subjects factor and one with target position as the within-subjects factor. For the fourth analysis, the same method was used, however, participants without both hit and false positive responses in all of the relevant conditions were excluded from this analysis. Given the substantial reduction in sample size due to only the inclusion of participants who had both hit and false positive responses in all cells of the relevant conditions in the latter analysis, a 2 (hit vs false positive responses) x 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph) univariate ANOVA treating lineup judgments as participants was conducted post hoc to confirm observed differences across all lineup judgments.
3.2 Results and discussion

3.2.1 Accuracy

Table 3.1 displays sample sizes and participants’ mean number of hit, miss and false positive responses. Given that the analyses were run on participants’ mean number of hits, misses and false positives across the eight lineups they completed, all 546 participants were included in all analyses.

For all three analyses, Box’s M equality of variances test was violated, as was Levene’s homogeneity of variances test for the same photograph condition. Levene’s homogeneity of variances test was also violated for the different photograph condition for the analyses on the mean number of false positive responses.

3.2.1.1 Lineup procedure: Hits

There was a main effect of lineup procedure on the number of hits. There were more hits in simultaneous than sequential lineups \([F(1, 544)= 21.40, p< .001; U= 29262.00, z= -3.94, p< .001, r= .169; M(SD)= 5.17 (1.18) vs 4.65 (1.43)]\). The main effect of lineup procedure was driven by an interaction between photograph and lineup procedure \([F(1, 544)= 17.95, p< .001]\). There were consistently more hits when participants completed simultaneous lineups, relative to sequential lineups, for only the same photograph \([t(316.647)= 7.37, p< .001; M(SD)= 3.80 (0.47) vs 3.32 (0.89)]\). These parametric analyses were confirmed by two independent samples nonparametric Mann-Whitney U tests on the mean number of hits in simultaneous versus sequential lineups for same \([U(544)= 25002.50, z= -7.72, p< .001]\) and different photograph \([U(544)= 35714.00, z= -.28, p= .778]\). The two-way interaction of photograph and lineup procedure was not driven by the observed three-way interaction between target position, photograph and lineup procedure \([F(1,544)=16.91 , p< .001]\) as there were more hits in simultaneous compared with sequential lineups for the same photograph condition in both early \([t(375.627)= 2.49, p= .007; U= 33801.00, z= -2.43, p= .008, r= .104; M(SD)= 1.91 (0.309) vs 1.82 (0.446)]\) and late \([t(298.715)= 8.06, p< .001; U= 25347.50, z= -8.19, p< .001, r= .351; M(SD)= 1.89 (0.32) vs 1.50 (0.68)]\) target presentations.
Table 3.1
*Mean (SD) number of responses and sample sizes for hit, miss and false positive responses for sequential and simultaneous lineups in the same or different photograph condition for early and late target presentations.*

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Target Position</th>
<th>Lineup Procedure</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Photograph</td>
<td>Early (n= 546)</td>
<td>Simultaneous Lineup (n= 319)</td>
<td>1.91 (0.31)</td>
<td>0.05 (0.22)</td>
<td>0.04 (0.20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential Lineup (n= 227)</td>
<td>1.82 (0.45)</td>
<td>0.06 (0.24)</td>
<td>0.11 (0.38)</td>
</tr>
<tr>
<td></td>
<td>Late (n= 546)</td>
<td>Simultaneous Lineup (n= 319)</td>
<td>1.89 (0.32)</td>
<td>0.08 (0.28)</td>
<td>0.03 (0.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential Lineup (n= 227)</td>
<td>1.50 (0.68)</td>
<td>0.05 (0.22)</td>
<td>0.45 (0.67)</td>
</tr>
<tr>
<td>Different Photograph</td>
<td>Early (n= 546)</td>
<td>Simultaneous Lineup (n= 319)</td>
<td>0.69 (0.67)</td>
<td>1.08 (0.76)</td>
<td>0.23 (0.45)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential Lineup (n= 227)</td>
<td>0.64 (0.65)</td>
<td>1.08 (0.73)</td>
<td>0.28 (0.56)</td>
</tr>
<tr>
<td></td>
<td>Late (n= 546)</td>
<td>Simultaneous Lineup (n= 319)</td>
<td>0.68 (0.70)</td>
<td>1.05 (0.75)</td>
<td>0.27 (0.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential Lineup (n= 227)</td>
<td>0.69 (0.72)</td>
<td>0.88 (0.75)</td>
<td>0.43 (0.61)</td>
</tr>
</tbody>
</table>

There was a main effect of target position \( [F(1,544)=10.77, p=.001; T=23723.00, z=-2.89, p=.002, r=.124] \). When the target was presented early in the lineup there were more hits than when it was presented late in the lineup overall \( [M(SD)=2.54(0.79)\ vs\ 2.41(0.91)] \). This main effect was driven by two interactions between target position and lineup procedure \( [F(1,544)=6.48, p=.011] \), and photograph condition and target position \( [F(1,544)=17.48, p<.001] \). However, these two-way interactions will not be discussed further as they were driven by the three-way interaction between target position, photograph condition and lineup.
procedure. There was a greater number of hits when the target was presented early compared to when it was presented late in sequential lineups for the same photograph \([t(227)= 6.74, p< .001; T= 3879.00, z= 6.12, p< .001, r= .287; M(SD)= 1.82 (0.45) vs 1.50 (0.68)]\) condition but not the different photograph condition \([t(227)= 0.87, p= .193; T= 3524.00, z= -0.95, p= .173; M(SD) early= 0.64 (0.65) vs late= 0.69 (0.72)]\). Further, there were more hits in simultaneous compared to sequential lineups when the target was presented late in the lineup only for the same photograph condition \([late: t(298.715)= 8.06, p< .001; U= 25347.50, z= -8.19, p< .001, r= .351; M(SD)= 1.89 (0.32) vs 1.50 (0.68); early: t(375.627)= 2.49, p= .007; U= 33801.00, z= -2.42, p= .008, r= .104; M(SD)= 1.91 (0.31) vs 1.82 (0.45), but not the different photograph condition \([late: t(544)= 0.23, p= .407; U= 36521.00, z= 0.19, p= .425; M(SD)= 0.68 (0.70) vs 0.69 (0.72); early: t(544)= 0.94, p= .175; U= 34730.50, z= -0.90, p= .185; M(SD)= 0.69 (0.67) vs 0.64 (0.65)]\). In other words, target position exaggerated differences in accuracy between lineup procedures such that the higher number of correct identifications in simultaneous lineups was accentuated when the target was presented late in the lineup and the same photograph of the target person was viewed.

Findings for hits were consistent with those for misses. There was a main effect of target position for misses \([F(1, 544)= 6.11, p= .014; T= 23938.50, z= -2.08, p= .019, r= .063]\). There were more misses when the target was presented early \([M(SD)= 1.14 (0.79)]\) than when they were presented late \([M(SD)= 1.05 (0.80)]\) in the lineup. This main effect was driven by an interaction between target position and lineup procedure \([F(1, 544)= 6.11, p= .014]\). There were more misses when the target was presented early than when it was presented late by participants who viewed sequential lineups \([t(226)= 3.44, p= .001; T= 3318.00, z= -3.40, p= .001, r= .160; M(SD)= 1.14 (0.79) vs 0.93 (0.77), but not those who viewed simultaneous lineups \([t(318)< 0.01, p= .500; T= 9284.50, z= 0.03, p= .489; M(SD)= 1.13 (0.80) vs 1.13 (0.81)]\). There were also more misses in simultaneous than sequential lineups when the target was presented late in the lineup \([t(544)= 2.92, p= .002; U= 31426.00, z= -2.80, p= .003, r= 120; M(SD)= 1.13 (0.814) vs 0.93 (0.773)], but not when it was presented early \([t(544)= 0.14, p= .447; U= 36239.00, z= 0.02, p= .493; M(SD)= 1.13 (0.798) vs 1.14 (0.786)]\).

Findings with respect to correct identifications are only partially consistent with previous literature. On the one hand, the finding that there are few differences
between lineup procedures, in terms of correct identifications, when a different photograph of the target person was viewed is consistent with previous research finding little difference in rates of correct identifications between lineups types (Cutler et al., 1987; Lindsay & Wells, 1980; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Wells, 1984, 1993; Yarmey, 1986). Given that these previous studies have tended to use staging of live crimes as the original observation followed by photographic or live lineup, the differences between the target’s appearance at the time of the staged crime and the target’s appearance in the lineup, would make this procedure analogous to the different photograph condition in the current experiment. Although the similarity of the appearance of the perpetrator of the staged crime to the appearance of this person in the lineup or photographic identification is often not reported (i.e. clothing type and colour, lighting, orientation, hairstyle), some differences in terms of at least orientation and lighting between a staged crime and a photographic or live lineup would be anticipated.

On the other hand, current findings suggest that the closer the match between the two photographs of an unfamiliar target person, the greater the superiority of simultaneous lineups compared to sequential lineups, in terms of correct identifications. This appeared to be the case when the targets were presented late in the lineup. This result is inconsistent with past research and has not been observed in previous psychological research. Some previous research has indicated that sequential lineups consistently produced a similar number of correct identifications to simultaneous lineups when the level of similarity of perpetrator to suspect (perpetrator replacement) was moderate or low but produced fewer correct identifications in comparison to simultaneous lineups when the level of similarity was high (McQuiston-Surrett et al., 2006). Very little research has investigated the impact on differences, or similarity, between two presentations of the same target person, rather than similarity between a perpetrator and an innocent suspect. As such, the current results add to the findings of previous research, suggesting that increasing the similarity between two photographs of an unfamiliar target person increases the superiority, in terms of correct identifications, of simultaneous compared to sequential lineups.

Increases in correct identifications with increasing similarity between original observation and observation of the target in the lineup can be understood in terms of the encoding specificity principle (Tulving & Thomson, 1973). In other words, the
properties of an effective retrieval cue is dependent on properties of the input stimuli upon encoding. Thus, the greater the match between what was originally encoded about the target person and what is encoded of the target person at retrieval, the more likely successful recognition will occur.

There are some differences between the simultaneous and sequential lineup procedure in general that may be relevant to understanding why greater similarity between observations accentuates the superiority of simultaneous over sequential lineups. When completing simultaneous lineups, participants view the target, then after a short delay, they are presented with all of the members of the lineup at once. It is likely that participants inspect each lineup member, then after inspecting all of the lineup members, they make a decision about which lineup member they believe best matches their memory of the target person. When completing sequential lineups, participants view the target, then after a short delay they are presented with one lineup member. They inspect the first lineup member and make a decision about whether that lineup member matches their memory of the target person without having viewed any of the other lineup members. If they do not choose that lineup member, then they are presented with the next lineup member and repeat the procedure until they either choose a lineup member as the target or they run out of lineup members and reject the lineup.

There is one plausible explanation, on the basis of current data, for why similarity between photographs of the target person may have increased the superiority, in terms of correct identifications, of simultaneous compared to sequential lineups, particularly when the targets were presented late in the lineup. Simultaneous lineups may have provided greater opportunity in comparison to sequential lineups for the match between the properties of the stimulus upon encoding (original observation of the target person) and the properties of the retrieval cue (observation of the target person in the lineup) to have enhanced accuracy in terms of correct identifications. This is because in sequential lineups participants can choose a lineup foil before viewing the target.

The viewing of foils prior to viewing the target may have interfered with accuracy. Presenting multiple lineup members all of which represented a good match with the context (background, orientation) in which the target person was encoded, may have increased the likelihood that the relevant sets of memory traces required for recognition were activated by a lineup foil prior to viewing the target.
person. In other words, in a sequential lineup, the witness may never have had the opportunity of viewing the target because prior to the appearance of the target the witness has incorrectly identified a foil. For example, if a target person is originally observed in a red T-shirt and sitting in a black chair facing the camera, then a lineup member is observed in a red T-shirt and sitting in the same black chair facing the camera, it may be more likely for a participant to confuse that lineup member for the target due to the match in clothing and context. Thus, more false identifications may occur and the increase in false identifications would logically decrease the number of correct identifications under these conditions.

In general, results demonstrate that when the target is presented late in the lineup, there is a decrease in accuracy in sequential lineups. Further, different photographs are observed to result in lower accuracy than same photographs in general. Thus, it would be expected that when different photographs were viewed, even though there would be less correct identifications in general, there would still be more correct identifications early than late in sequential lineups. However, results indicate that while there is a similar increase in false identifications late compared to early in sequential lineups in the different photograph condition, there is no difference in the number of correct identifications. The reasons behind this finding are unclear and no immediate explanation for this change in the general pattern of results can be provided. Experiments 2 and 3 will aim to further investigate and test the robustness of the current findings.

3.2.1.2 Lineup procedure: False positives

There was a main effect of lineup procedure on the number of false positive responses. More false positive responses were made when participants completed the sequential compared to the simultaneous lineups \(F(1, 544)= 47.26, p< .001; U= 45523.00, z= 5.60, p< .001, r=.246; M(SD)= 1.28 (1.58) vs 0.57 (0.81)\). There was also an interaction between photograph and lineup procedure \(F(1, 544)= 18.82, p< .001\). The magnitude of the difference in false positives between simultaneous and sequential lineups was greater in the same photograph condition. However, this two-way interaction was driven by a three-way interaction between photograph condition, target position and lineup procedure (see below).

Further, there was an interaction between photograph condition, target position and lineup procedure \(F(1, 544)= 12.05, p=.001\). There were more false
identifications in the different than same photograph condition for both target positions for participants who viewed simultaneous lineups [early: t(318) = 7.25, p < .001, M(SD) = 0.23 (0.45) vs 0.04 (0.12); late: t(318) = 8.19, p < .001, M(SD) = 0.27 (0.50) vs 0.03 (0.18)], but for those who viewed sequential lineups there were only more false positive responses in the different photograph condition when the target was presented early in the lineup [t(226) = 4.42, p < .001, M(SD) = 0.28 (0.56) vs 0.11 (0.38)] but not when it was presented late in the lineup [t(226) = 0.57, p = .284, M(SD) same = 0.45 (0.67) vs late = 0.43 (0.61)]. This was confirmed by nonparametric related samples Wilcoxon signed-rank tests [simultaneous same vs different: early, T= 1990.00, z= 6.71, p< .001, r= .266; late, T= 3216.50, z= 7.43, p< .001, r= .294; sequential same vs different: early, T= 1153.00, z= 4.18, p< .001, r= .196; late, T= 2008.50, z= 0.56, p= .285]. This interaction did not drive the two-way interaction between lineup procedure and target position as the difference between lineup procedures was larger when the target was presented late in the lineup in both the same [early: t(312.466) = 2.67, p= .008, M(SD) = 0.11 (0.38) vs 0.04 (0.20); late: t(248.202) = 9.32, p< .001, M(SD) = 0.45 (0.67) vs 0.03 (0.18)] and different [early: t(419.054) = 1.26, p= .104, M(SD) = 0.28 (0.56) vs 0.23 (0.45); late: t(424.420) = 3.21, p= .001, M(SD) = 0.43 (0.61) vs 0.27 (0.50)] photograph conditions, however the difference between lineup procedures in the different photograph condition when the target was presented early was not significant. The effect of late target position in both photograph conditions was confirmed by nonparametric independent samples Mann-Whitney U tests for same [early: U= 381113.00, z= 2.51, p= .006, r= .109; late: U= 48101.00, z= 0.11, p< .001, r= .433] and different [early: U= 37086.50, z= 0.67, p= .251; late: U= 40806.00, z= 3.18, p< .001, r= .163] photograph conditions. There were also more false identifications in sequential lineups when the target was presented late compared with when it was presented early in both the same [t(226) = 7.43, p< .001; T= 435.00, z= -6.55, p< .001, r= .308; M(SD) = 0.45 (0.67) vs 0.11 (0.38)] and different photograph conditions [t= 3.11, p= .001; T= 1369.00, z= -3.01, p= .002, r= .141; M(SD) = 0.43 (0.61) vs 0.28 (0.56)], whereas differences between late and early target presentations in simultaneous lineups were not significant for both photograph conditions [same: t(318) = 0.65, p= .257; T= 132.00, z= 0.67, p= .513; M(SD)= 0.03 (0.18) vs 0.04 (0.20); different: t(318) = 1.25, p= .107; T= 2019.00, z= -1.25, p= .213; M(SD)= 0.27 (0.50) vs 0.23 (0.50)].
Chapter 3 | Experiment 1: Target Present Lineups

There was a main effect of target position \[F(1, 544)= 53.21, p< .001; T= 17085.50, z= 5.86, p< .001, r= .177\]. There were more false positive responses when the target was late in the lineup \[M(SD)= 0.54 (0.85)\] compared to when the target was early in the lineup \[M(SD)= 0.32 (0.64)\]. There was also an interaction between target position and lineup procedure \[F(1, 544)= 40.01, p< .001\] that drove the main effect of target position. There were more false identifications when the target was presented late compared with when it was presented early only in sequential lineups \[t(226)= 7.46, p< .001; T= 5539.50, z= 6.67, p< .001, r= .313\], but not in simultaneous lineups \[t(318)= 0.90, p= .184; T= 2988.00, z= 0.91, p= .182\]. The interaction also indicated that the magnitude of difference in terms of false identifications between lineup procedures was greater when the target was presented late \[t(307.039)= 7.54, p< .001; U= 46993.50, z= 6.94, p< .001, r= .297, M(SD)= 0.88 (1.07) vs 0.30 (0.54)\] than when it was presented early \[t(371.070)= 2.22, p= .014; U= 38198.00, z= 1.46, p= .072, r= .406, M(SD)= 0.40 (0.77) vs 0.27 (0.52)\].

The current results suggest that there are changes occurring across sequential lineups that elevate false positive responding. Unlike simultaneous lineups, sequential lineups involve serial presentation of lineup members and this may impact upon a participant’s decision making as the lineup progresses. This suggestion would be clearly consistent with the present finding that false identifications were more likely to occur when the target was presented late in sequential lineups. Some previous research has found that sequential lineups may, in fact, be inferior to simultaneous lineups for discriminating between the presence or absence of a guilty suspect in a lineup (Amendola & Wixted, 2015a, 2015b; Mickes et al., 2012). However due to the analysis utilised by these past studies, it is unclear whether these results were due to an elevation in false identifications or decrease in correct identifications. Further, studies that have reported effects of target position have typically shown that if a lineup member similar to the target is presented before the target face in sequential lineups, participants tend to identify the similar lineup member before they have viewed the target person, thus resulting in more false identifications (Clark & Davey, 2005; Flowe & Ebbesen, 2007). The current results are consistent with these findings, suggesting that for sequential lineups, the increase in false identifications, and corresponding decrease in correct identifications, when the target was presented late in the lineup may have been due to a lineup foil being chosen before participants had a chance to view the target person in the lineup.
However, the observation that this occurred when similarity of the lineup foils was not manipulated suggests that some change in decision making or in memory processes over sequential lineups may have occurred. Given no systematic differences in the relative similarity of lineup foils to the target, if there was no shift in decision making or memory, there would be no reason for an increase in false identifications in sequential lineups, particularly when the target was presented late in the lineup. The present findings are consistent with Clark and Davey’s (2005) data which indicated that some witnesses did display a change in their decision making, increasing their propensity to identify a lineup member in lineups where the target person was absent. The current results extend Clarke and Davey’s (2005) findings, suggesting that participants may change their decision making across sequential lineups, being more likely to identify a lineup member as the target as the lineup progresses, resulting in an increase in false identifications, not only when the target is absent, but also when the target person is present in the lineup. Possible theoretical explanations for this change in decision making will be discussed later in the general discussion (Chapter 6).

Findings also demonstrated that the different photograph condition interacted with target position in sequential lineups to augment the impact of photograph condition on accuracy. When the target was presented late in sequential lineups, there was no difference in the number of false identifications between the same and different photograph condition. This result appeared to be due to elevated numbers of false identifications occurring when the same photograph of the target person was viewed late in sequential lineup. As previously mentioned, in the same photograph condition in sequential lineups, the lineup members presented before the target could have interfered with accuracy. The reasons for why the different photograph condition did not further elevate the number of false identifications when the target was presented late in sequential lineups are not clear. While the difficulty of the task when there was a different photograph presented late in sequential lineups may have caused participants to simply guess, the hit rate of 0.69 for this cell of the design is well above chance levels. Thus, no immediate explanation for this result can be offered. Experiments 2 and 3 in the current investigation may aid to shed light on the reliability of the current findings.

Nevertheless, current results demonstrates that late presentation of the target in a sequential lineup that terminates once a decision is made increases rates of false
identifications. This is consistent with previous studies indicating that lineup participants may be more likely to falsely identify a similar looking foil if that foil is presented before the target (Clark, 2005; Clark et al., 2008). Further, studies on the VIPER (video identification parade electronic recording) system, a system developed to produce video identification parades, have found that allowing witnesses to see the entire lineup more than twice when using sequential lineups results in more correct identifications than only presenting faces once and terminating the lineup once a decision has been made (Valentine, Darling, & Memon, 2007). This previous research, in conjunction with the current study, suggest that some degree of ‘relative’ decision making, or being able to make some comparison across lineup members may be beneficial to accuracy. In sequential lineups, participants may engage in a comparison between the lineup members they have viewed in order to find the ‘best match’ with their mental representation of the target person, just the same as they may compare between lineup members in simultaneous lineups. However, the difference in sequential lineups is that participants do not know which faces, or how many faces, are to come. This possibility would be consistent with previous studies suggesting that participants may compare relative similarity between the lineup members they have viewed in sequential as well as simultaneous lineups (Clark & Davey, 2005; Flowe & Ebbesen, 2007). If this kind of process were occurring, it may also suggest that the threshold for choosing would shift with changing desire to see more or less faces. In other words, utilizing a strictly ‘absolute judgment’ strategy and not allowing eyewitnesses to view all lineup members before making a decision may result in poorer decision making. This poor decision making may result from witnesses feeling pressure to choose before running out of lineup members, increasing the likelihood of a lineup member being chosen before the target is viewed. However, this possibility is speculative and cannot be confirmed based on the present study. Additionally, the current research only used the same or a different photograph to that originally presented as the target and thus it is not known what impact varying levels of similarity would have on accuracy in terms of false identifications. As such, Experiment 3 will investigate the impact of varying levels of similarity on accuracy.
3.2.1.3 Photograph condition (same versus different)

There was a main effect of photograph condition on the mean number of hits [F(1,544)=1825.99, p< .001; T= 394.00, z= -19.53, p< .001, r=.591], miss [F(1, 544)= 1346.50, p< .001] and false positive [F(1, 544)= 75.20, p< .001; T= 19978.50, z= 8.66, p< .001, r=.370] responses. More hits occurred in the same photograph condition than in the different photograph condition [M(SD)= 3.60 (0.71) vs 1.35 (1.04)]. Although there were several interactions involving photograph condition reported, none of these drove the main effect. There were more hits in the same than the different photograph condition for early and late target presentations in both simultaneous [early: t(318)= 29.88, p< .001; T= 70.50, z= -14.52, p< .001, r=.575; M(SD)= 1.91 (0.31) vs 0.69 (0.67); late: t(318)= 28.37, p< .001; T= 350.00, z= -14.37, p< .001, r=.569; M(SD)= 1.89 (0.32) vs 0.68 (0.70)] and sequential lineups [early: t(226)= 23.70, p< .001; T= 348.00, z= -12.19, p< .001, r=.572; M(SD)= 1.82 (0.45) vs 0.64 (0.65); late: t(226)= 12.74, p< .001; T= 1060.00, z= -9.56, p< .001, r=.448; M(SD)= 1.50 (0.68) vs 0.69 (0.72)]. Less misses were made to the same photograph [M(SD)= 0.12 (0.36)] compared to the different photograph condition [M(SD)= 2.06 (1.18)]. This was confirmed by a nonparametric related samples Wilcoxon signed-rank test comparing same to different photograph conditions [T= 112850.50, z= 19.00, p< .001, r=.575]. Again, although there was an interaction with target position, this did not drive the main effect of photograph condition. Less misses were made in the same than different photograph condition for both early [t(545)= 31.10, p< .001; T= 78772.00, z= 17.77, p< .001, r=.538; M(SD)= 0.05 (0.23) vs 1.09 (0.75)] and late [t(545)= 26.93, p< .001; T(546)= 71816.50, z= 16.97, p< .001, r=.514; M(SD) 0.07 (0.26) vs 0.98 (0.76)] target presentations. The interaction between target position and photograph condition [F(1, 544)= 7.89, p= .005] indicated that early target position increased misses in the different photograph condition [t(545)= 2.47, p=.007; T= 36239.00, z= 0.02, p=.493; M(SD) early= 1.08 (0.75) vs late= 0.98 (0.76)] but not in the same photograph condition [t(545)= 0.79, p=.215; T= 36239.00, z= 0.02, p=.493; M(SD) early= 0.06 (0.23) vs late= 0.07 (0.26)]. There were more false positive responses made to different than same photographs [F(1, 544)= 75.20, p< .001; T= 19978.50, z= 8.66, p< .001, r=.370; M(SD)= 0.58 (0.81) vs 0.28 (0.63)]. This was impacted by the three-way interaction between lineup procedure, photograph condition and target position reported above, whereby for those who viewed sequential lineups no difference between photograph
conditions was found when the target was presented late in the lineup [t(226)= 0.57, 
p=.284, M(SD) same= 0.45 (0.67) vs late= 0.43 (0.61)]. The pattern of current results indicate that presenting the same photograph of the target person as was viewed earlier increased accuracy dramatically, compared to when a different photograph of the same person was presented. When a different photograph was presented the previously discussed impact of early target position led to decreased accuracy, specifically more incorrect rejections of the lineup. The impact of target position when a different photograph was shown may be due to participants being more likely to decide that the target person was not in the lineup when there were differences between original observation and observation of the target early in the lineup. As discussed above, when targets were presented late in the lineup, participants appeared more likely to choose a lineup member. Findings are consistent with previous research which indicates that consistency in the context and external features of an unfamiliar target person between the original observation of the target and subsequent observation of that target in the lineup, leads to increased accuracy in recognition of the target (Thomson, 2003; Thomson et al., 1982). Even small changes in context, clothing, distance, lighting or hairstyle have been found to lead to an increase in errors for facial recognition and identification tasks (Malpass & Devine, 1981a; Thomson, 2003; Thomson et al., 1982; Yarmey, 1986). Indeed, Thomson (1982) conducted a series of experiments testing the ability of people to recognise unfamiliar individuals under several different conditions. These experiments revealed that person recognition was far superior when the background, clothing and orientation at the time of testing was the same as the original observation compared to when these variables were changed. They also found that reinstating a previously seen context with a different target increased the number of false identifications for that target significantly. The current findings lend support to, and further strengthen, this research emphasising the importance of context, clothing and external features in the accuracy of unfamiliar person recognition. Findings suggest that attempting to maintain the constancy of such features in real life identification situations, at least in simultaneous lineups, may increase the likelihood of correct recognition of the suspect from a lineup, should that suspect be the perpetrator of the crime in question. The robustness of these findings will be further examined in Experiment 2.
3.2.2 Confidence

Three separate analyses were planned to investigate differences in confidence between lineup procedure and photograph condition for hit, miss and false positive responses. A fourth analysis was planned to investigate differences between hit and false positive responses in simultaneous and sequential lineups under the same and different photograph conditions. Participants were included in the analyses on the basis of the incidence of hit, miss and false positive responses. As such, it must be noted that some participants may have contributed to only one analysis whereas others to multiple analyses. Box’s test of equality of covariance matrices and Levene’s test of equality of error variances were violated for some of the analyses conducted on confidence ratings and these violations are noted below.

Samples which included less than 10 participants per cell of the design were considered too small, and not powerful enough, to conduct the planned ANOVAs on. Therefore, some additional analyses collapsing within-subjects factors were conducted post hoc due to small numbers of miss and false positive responses being observed in some cells of the study design. The samples analysed were considered statistically powerful enough to conduct ANOVAs and thus the analyses reported were used to examine the impact of lineup type, photograph condition and target position on confidence in miss and false positive responses. It must be noted, however, that several of the samples used in the analyses reported below represented a very small proportion of the total sample. As such, interpretation of findings related to these samples should be undertaken cautiously.

3.2.2.1 Lineup procedure

In order to allow for within-subjects comparisons, only participants who had hits in both same and different photograph conditions with both early and late target presentations were included in the analysis for hits. On the basis of this criterion, 368 participants were excluded leaving 178 (106 female) in the sample. This sample was considered large enough to conduct the planned analyses. Table 3.2 displays means and standard deviations for both simultaneous and sequential lineups participants who had hit responses in the same and different photograph conditions when the target was presented early and late in the lineup.
Table 3.2
*Mean (SD) confidence levels for hit responses for sequential and simultaneous lineups in the same or different photograph condition for early and late target.*

<table>
<thead>
<tr>
<th>Lineup Procedure (n=178)</th>
<th>Same Photograph (n=178)</th>
<th>Different Photograph (n=178)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target Early (n=178)</td>
<td>Target Late (n=178)</td>
</tr>
<tr>
<td>Simultaneous Lineups (n=110)</td>
<td>4.59 (0.69)</td>
<td>4.57 (0.69)</td>
</tr>
<tr>
<td></td>
<td>Target Early (n=178)</td>
<td>Target Late (n=178)</td>
</tr>
<tr>
<td>Sequential Lineups (n=68)</td>
<td>4.83 (0.47)</td>
<td>4.72 (0.59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.43 (1.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.28 (1.12)</td>
</tr>
</tbody>
</table>

Levene’s homogeneity of variances test was violated for the same photograph condition when the target was presented early, but Box’s M equality of variances test was not violated for this analysis. There was no main effect of lineup procedure on the mean number of hits [F(1, 176)= 1.80 p= .182; M(SD) simultaneous= 4.06 (0.66) vs sequential= 4.23 (0.52)].

Only four participants (two female) had misses in both same and different photograph conditions for both early and late target presentations. When the analysis using photograph condition as the within-subjects factor was prepared only participants who had misses in both the same and different photograph conditions were included. On the basis of this criterion, 488 (307 female) participants were excluded, leaving 58 (41 female) participants in the sample. Table 3.3 displays means and standard deviations for both simultaneous and sequential lineups for participants who had miss responses in the same and different photograph conditions.
Table 3.3
Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had miss responses in the same and different photograph condition.

<table>
<thead>
<tr>
<th>Lineup Procedure</th>
<th>Same Photograph (n= 58)</th>
<th>Different Photograph (n= 58)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous Lineup</td>
<td>2.43 (1.30)</td>
<td>2.97 (0.97)</td>
</tr>
<tr>
<td>(n=37 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential Lineup</td>
<td>2.43 (1.33)</td>
<td>3.27 (1.23)</td>
</tr>
<tr>
<td>(n= 21)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the analysis using target position as the within-subjects factor was conducted, 223 (134 female) participants were excluded from the analysis on the basis of this criterion, leaving 323 (214 female) in the sample. Table 3.4 displays means and standard deviations for both simultaneous and sequential lineups for participants who had miss responses for both early and late target presentations. There was no significant main effects of, or interactions with, lineup procedure observed for any analysis on confidence of miss responses.

Table 3.4
Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had miss responses for early and late target presentations.

<table>
<thead>
<tr>
<th>Lineup Procedure</th>
<th>Target Early (n= 323)</th>
<th>Target Late (n= 323)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous Lineup</td>
<td>3.17 (1.21)</td>
<td>3.28 (1.17)</td>
</tr>
<tr>
<td>(n=194 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential Lineup</td>
<td>3.02 (1.24)</td>
<td>3.10 (1.31)</td>
</tr>
<tr>
<td>(n= 129)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only eight participants (five female) had false positive responses in both
same and different photograph conditions for both early and late target presentations. When the analysis using photograph condition as the within-subjects factor was prepared, only participants who had false positive responses in both same and different photograph conditions were included. On the basis of this criterion, 468 (297 female) participants were excluded, leaving 78 (51 female) participants in the sample. Table 3.5 displays means and standard deviations for both simultaneous and sequential lineups for participants who had false positive responses in the same and different photograph conditions. No violation of the parametric assumptions was observed for this analysis.

When the analysis using target position as the within-subjects factor was conducted for false positive responses, 472 (303 female) participants were excluded from the analysis on the basis of this criterion, leaving 74 (45 female) in the sample. Table 3.6 displays means and standard deviations for both simultaneous and sequential lineups for participants who had false positive responses for both early and late target presentations.

Table 3.5
Mean (SD) confidence levels for sequential and simultaneous lineups in participants who falsely identified a lineup foil in the same and different photograph condition.

<table>
<thead>
<tr>
<th>Lineup Procedure (n= 78)</th>
<th>Same Photograph (n= 78)</th>
<th>Different Photograph (n= 78)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous Lineup (n=14)</td>
<td>2.39 (1.11)</td>
<td>2.50 (1.21)</td>
</tr>
<tr>
<td>Sequential Lineup (n= 64)</td>
<td>2.67 (0.90)</td>
<td>2.71 (0.92)</td>
</tr>
</tbody>
</table>
Table 3.6

<table>
<thead>
<tr>
<th>Lineup Procedure</th>
<th>Target Early (n=74)</th>
<th>Target Late (n=74)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous Lineup</td>
<td>3.00 (1.25)</td>
<td>2.49 (1.17)</td>
</tr>
<tr>
<td>(n=30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential Lineup</td>
<td>2.61 (0.94)</td>
<td>2.59 (0.77)</td>
</tr>
<tr>
<td>(n=44)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Box’s test of equality of covariance matrices was violated for this analysis as was Levene’s test of equality of variances for late target position. No main effects or interactions involving lineup procedure were observed for either of the above analyses on confidence in false positive responses.

Given that the samples when including within-subjects factors were a very small proportion (less than one sixth) of the total sample, an analysis collapsing all within-subjects factors was conducted in order to assess the impact of lineup type on confidence in false positive responses. When the variables of photograph and target position were collapsed, all participants who gave a false positive response in any of the conditions were included in the sample. Therefore, when a one-way univariate ANOVA with lineup procedure as the only variable was performed, only 283 (181 female) participants were excluded, increasing the sample to 263 (167 female). Using the univariate ANOVA, there was a main effect of lineup procedure, such that reported confidence was higher for sequential than for simultaneous lineups \([F(1, 261)= 9.33, p=.002; M(SD)= 2.54 (0.92) \text{ vs } 2.20 (0.96)]\). As such, analyses indicate overall that there was higher confidence in false identifications in sequential than simultaneous lineups. No other main effects or interactions were significant.

The current findings suggest that differences in confidence between simultaneous and sequential lineups do not mirror differences in accuracy. One noteworthy finding was observed: participants who gave false identifications demonstrated a higher level of overall confidence in the sequential compared to
simultaneous lineups. These results are inconsistent with previous research in the area. Although few studies have examined differences in confidence when people complete simultaneous compared to sequential lineups, previous studies that have, have found no differences in confidence whether participants complete simultaneous or sequential lineups (Kneller et al., 2001; Lindsay & Wells, 1985). One study did find that sequential lineups produced higher confidence, but only among nonchoosers, or people who did not identify a person from the lineup as the target (Sporer, 1993). However, this difference in confidence between sequential and simultaneous lineups was not discussed. The current results are inconsistent with this latter study, instead indicating that generally sequential lineups produced higher confidence, but only among those who falsely identified a person from the lineup as the target, but not among nonchoosers (people who incorrectly rejected the lineup). Some previous eyewitness research has indicated that asking a person to rate their confidence in having seen a lineup member before provides a better index of whether a face had been previously seen or unseen than asking a person for a binary response as to whether or not the face had been seen (Sauer, Weber, & Brewer, 2012). As such, the utility of confidence ratings may be enhanced if confidence in having seen a lineup member previously were asked for, rather than confidence in the participants’ binary response. Further study would be needed to explore this possibility.

Nevertheless, there is no clear explanation for the finding that sequential lineups led to higher confidence for false identifications. If a decision criterion threshold or memory explanation could be applied, it should have been seen consistently in all cells of the design and also correspond to differences in accuracy. Therefore, it is unclear why confidence in sequential lineups in the current experiment was higher than simultaneous lineups, when participants falsely identified a lineup member. Given that this is the first observation within the eyewitness literature of elevated confidence in sequential lineups despite decreased accuracy, replication of the current findings would be necessary to confirm presently observed effects. Further, it must be noted that current analyses were unable to determine whether there were any three-way interactions between lineup procedure, target position and photograph condition contributing to differences in confidence in false identifications between sequential and simultaneous lineups. Thus, these data should be interpreted cautiously. Thus, Experiments 2 and 3 will further examine the
impact of lineup type on accuracy in order to ascertain the robustness of presently observed effects.

Regardless of the mechanisms behind such results, current data do suggest that in many instances, sequential lineups may be associated with not only more false identifications, but also elevated confidence in those false identifications compared with simultaneous lineups. These results are noteworthy, given that they demonstrate that witness confidence may be higher for false identifications given in sequential lineups compared to simultaneous lineups. Thus, inferring accuracy from confidence in responses in an eyewitness context within the criminal justice system may lead to more, rather than fewer, wrongful convictions.

3.2.2.2 Photograph

There was a main effect of photograph condition on confidence in hit responses [F(1, 176)= 437.09 p < .001]. Confidence when viewing the same photograph in the lineup as the originally presented target photograph was higher than when viewing a different photograph [M(SD)= 4.66 (0.58) vs 3.31 (0.96)]. A nonparametric related samples Wilcoxon signed-rank test confirmed that same photograph condition yielded higher confidence levels than the different photograph condition [T= 23.500, z= -11.11, p< .001, r= .589]. Similarly, there was a main effect of photograph condition on confidence in miss responses [F(1, 56)= 12.22, p= .001]. Confidence in misses when different photographs were viewed was higher than when the same photographs were viewed [M(SD)= 3.08 (1.07) vs 2.43 (1.30)]. No main effect of photograph [F(1, 76)= 0.17, p= .682] on confidence in false positive responses and no interaction effects were observed.

The current results are consistent with the proposition that seeing the same photograph of the target person as was viewed originally impacts upon both accuracy and confidence, increasing correct identifications and confidence in those correct identifications. Participants were more confident in correctly identifying a target from the lineup and less confident in incorrectly rejecting a lineup when the same photograph of the target person as was viewed earlier was presented compared to when a different photograph of the target was presented. These findings are consistent with previous studies indicating that confidence and accuracy are closely related, particularly when the identification task is simple (Bradfield et al., 2002; Weber & Brewer, 2004). For example, Bradfield and colleagues (2002) found that
post-identification confidence was related to accuracy, with a correlation of .58, however participants were allowed to watch targets for three minutes in a variety of different contexts and positions, making the identification task highly conducive to recognition. Other studies, however, that required participants to engage in an identification task after one brief encounter with a perpetrator have found less support for a relationship between confidence and accuracy (Cutler & Penrod, 1989; Cutler et al., 1987). It is well established that people are more accurate at recognition when observation of the target occurs in the same context and with the same clothing and lighting as originally observed (Thomson, 2003; Thomson et al., 1982; Yarmey, 1986). This finding has been explained in terms of encoding specificity (Tulving & Thomson, 1973). Upon encoding a face, information about the item and the context in which the item is viewed is encoded and stored in memory. Thus, the more the context matches what was originally seen, the more the information resonates with stored memories (Atkinson & Juola, 1974; Tulving & Thomson, 1973). The consistency in clothing, context and hairstyle when the same photograph of the target person as viewed earlier was presented in the current study would likely have facilitated correct recognition of the target, decreasing the likelihood of incorrect rejection of the lineup, when compared to when a different photograph of the target person was presented, as is consistent with findings relating to accuracy.

Findings relating to confidence in the same versus different photograph conditions are in line with accuracy, suggesting that participants may have been aware that the same photograph condition facilitated their recognition of the target. When recognition is fast and effortless, participants may feel more confident in their answer. On the other hand, on the basis of current results, in the different photograph condition participants were likely not aware that the target was in the lineup due to the differences in target presentation across observations, as confidence in incorrect rejections of the lineup was higher in the different photograph condition. Similar to the process in correct recognition, when there are differences between observations of the target causing participants to reject the lineup, internal feedback of the mismatch between the memory trace and observation of the target may be received, such that the participant is confident that they correctly rejected the lineup.

As such, when the same photograph of the target person as was viewed originally was presented in the current study, participants’ awareness of successful
recognition may have strengthened the relationship between confidence and accuracy. On the other hand, in the different photograph condition, their lack of awareness of an incorrect rejection may have decreased the relationship between confidence and accuracy. While, the same photograph may have strengthened participants’ awareness of possibly missing the target due to a sense of familiarity of the target in the lineup, the different photograph condition may have decreased awareness of missing the target due to lack of familiarity caused by a mismatch between observations of the target. This would be logically related to a high relationship between actual accuracy and the belief in accurate (and conversely disbelief in inaccurate) identifications in the same photograph condition, whereas the converse would be true for the different photograph condition. These findings therefore lend further support to the notion that when consistency between the originally observed target and target observed in the lineup is high, people who are accurate are also likely to be confident. However, when consistency between observations is low, the confidence-accuracy relationship is also likely to be lower.

3.2.2.3 Target position

There was a main effect of target position for confidence in hit responses \[F(1, 176)= 6.69 \ p= .011; \ T= 7335.50, \ z= 2.97, \ p= .002, \ r= .158\]. Overall, confidence was higher when the target was presented early than when it was presented late in the lineup \[M(SD)= 4.20 (0.69) \ vs 4.05 (0.70)\]. No other interactions or main effects were significant. These results indicate that early target position increased both the number of, and confidence in, correct identifications. This may be related to the short span of time and viewing of less faces between observation of the target and viewing the target in the lineup. However given that interactions with photograph and lineup procedure were observed for accuracy, it is not clear why no interactions were observed for confidence. The effect of target position on confidence has not often been discussed in the empirical psychological literature on eyewitness identification as most studies that examine confidence do not include target position as a variable in analyses (Brewer et al., 2002; Brewer et al., 2012; Brewer & Wells, 2006; Carlson et al., 2016; Deffenbacher, 1980). Issues relating to target position will be considered and discussed further in the general discussion (Chapter 6).
3.2.2.4 Confidence in correct versus false identifications

No participants had hit and false positive responses for both the same and different photograph condition for both early and late target presentations. As such, two different analyses comparing confidence in hits to false positive responses were prepared, one using photograph condition as the within-subjects factor and one using target position as the within-subjects factor. This was to elucidate whether each of these within-subjects factors had an effect on confidence ratings or any interactions with differences in the confidence of hit versus false positive responses.

When the analysis including only photograph as the within-subjects factor was prepared, only participants who had both hits and false positives in the same and different photograph condition were included. On the basis of this criterion, 487 (309 female) participants were excluded leaving 59 (39 female) participants in the sample. This sample was considered statistically powerful enough to conduct the ANOVA, however due to the small proportion of the total sample included in this analysis (less than one sixth), findings should be interpreted cautiously. Table 3.7 displays means and standard deviations for both simultaneous and sequential lineups for participants who had both hit and false positive response in same and different photograph conditions. No violations of the parametric assumptions were observed for this analysis.

Table 3.7
Mean (SD) confidence levels for sequential and simultaneous lineups in participants who had both hit and false positive responses for both same and different photograph conditions.

<table>
<thead>
<tr>
<th>Photograph Condition (n= 59)</th>
<th>Lineup Procedure (n= 59)</th>
<th>Hits (n= 59)</th>
<th>False Positives (n= 59)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Photograph (n= 59)</td>
<td>Simultaneous Lineups (n=11)</td>
<td>4.30 (0.59)</td>
<td>2.41 (1.07)</td>
</tr>
<tr>
<td></td>
<td>Sequential Lineups (n=48)</td>
<td>4.76 (0.41)</td>
<td>2.72 (0.94)</td>
</tr>
<tr>
<td>Different Photograph (n= 59)</td>
<td>Simultaneous Lineups (n=11)</td>
<td>3.23 (1.13)</td>
<td>2.63 (1.10)</td>
</tr>
<tr>
<td></td>
<td>Sequential Lineups (n=48)</td>
<td>3.57 (0.99)</td>
<td>2.76 (0.97)</td>
</tr>
</tbody>
</table>

There was a significant main effect of response type \[F(1, 57)= 125.75, p<\]
Reported confidence for hits was higher than for false positive responses overall \([M(\text{SD}= 4.21 (0.56) vs 2.68 (0.71))\), however, the range of ratings for both hits and false positives was the same (1-5). There was also an interaction between response type and photograph \([F(1, 57)= 27.11, p < .001]\). The magnitude of the difference in confidence between hit and false positive responses was greater when the same photograph was presented \([t(58)= 15.63, p < .001; M(\text{SD}= 4.67 (0.48) vs 2.66 (0.96))\) than when a different photograph of the target person as was originally observed was presented \([t(58)= 5.71, p < .001; M(\text{SD}= 3.50 (1.01) vs 2.73 (0.99))\) in the lineup. In other words, it appeared that photograph condition made a difference in the number of hits but not the number of false identifications observed. The interaction did not drive the main effect of response type.

There was also a main effect of photograph \([F(1, 57)= 11.36, p = .001]\) which reflected higher confidence for same than different photographs for hit responses. No other interactions were observed.

When the analysis comparing hits and false positive responses including only target position as the within-subjects factor was prepared, only participants who had both hits and false positives when the target was presented early and when it was presented late were included in the analysis. On the basis of this criterion, 483 (310 female) participants were excluded, leaving 63 (38 female) participants in the sample. No violations of parametric assumptions were observed with the exception that Levene’s test of equality of error variances was violated for false positive responses when the target was presented late in the lineup. Means and standard deviations for confidence in hit versus false positive responses in sequential and simultaneous lineups when the target was presented early or late can be seen in Table 3.8.
There was a significant main effect of response type [F(1, 61)= 203.80, p< .001; T< 0.001, z= -6.85, p< .001, r= .610]. Reported confidence for hits (range= 1-5) was higher than for false positive responses (range= 1-5) overall [M(SD)= 4.21 (0.63) vs 2.63 (0.86)]. There was also a three-way interaction between lineup procedure, response type and target position [F(1, 61)= 5.72, p= .020]. Confidence in hits for sequential lineups was higher when the target was presented early than when it was presented late in the lineup [t(33)= 2.75, p= .010], but this was not the case for simultaneous lineups [t(28)= 0.44, p= .663], nor was it the case for confidence ratings of false positives in sequential lineups [t(33)= 0.27, p= .792; T= 202.50, z= 0.32, p= .751] and simultaneous lineups [t(28)= 1.93, p= .064; T= 169.50, z= 1.90, p= .058]. However, this interaction did not drive the main effect of response type, as confidence in hits was higher than that of false positive responses for both simultaneous [t(28)= 6.70, p< .001; and t(28)= 7.20, p< .001, T= 11.00, z= -4.29, p< .001, r= .563] and sequential lineups [t(33)= 10.52, p< .001; and t(33)= 6.92, p< .001, T= 34.50, z= -4.51, p= .001, r= .547] for both early and late target presentations, respectively. (See Table 3.8 for means and standard deviations.)

There was also a main effect of target position [F(1, 61)= 203.80, p< .001; T< 0.001, z= -6.85, p< .001, r= .610]. Confidence was higher when the target was presented early than when it was presented late overall.

In order to exclude the possibility that the higher confidence for hit when compared to false positive responses may be due to the particular sample used, namely only participants who had hit and false positive responses in the relevant
conditions, the analysis was run across all lineup judgments, rather than across participants. When this analysis was conducted, 3176 (2705 hit and 471 false positive responses) lineup judgments were included in the analysis. However, for transparency nonparametric as well as parametric tests are reported. Table 3.9 displays means and standard deviations for both simultaneous and sequential lineups for participants who had both hit and false positive response in same and different photograph conditions when the target was presented early or when it was presented late in the lineup.

There was still a main effect of response type, such that confidence for hit responses was higher than that for false positive responses \([F(1, 3160) = 563.22, p < .001]\). This was confirmed by a nonparametric Mann-Whitney U test \([U(3174) = 180784.00, z = -26.730, p < .001, r = .474]\). The ANOVA further indicated that there was a main effect of lineup procedure such that participants who viewed sequential lineups reported higher levels of confidence than those who viewed simultaneous lineups \([F(1, 3160) = 29.63, p < .001]\). However, the nonparametric Mann-Whitney U test did not confirm this analysis, indicating no significant difference in confidence between sequential and simultaneous lineup procedures \([U(3174) = 1216225.00, z = -0.65, p = .517]\). Given the violation of parametric assumptions and the non-significant finding using the nonparametric test, it is not likely that the main effect of lineup procedure using the parametric test is reliable. Analyses conducted in the individual response types are likely more informative when considering the impact of lineup procedure on confidence in hit and false positive responses.
In line with effects for hits reported above, there was also a main effect of photograph condition \([F(1, 3160)= 100.19, p< .001]\), such that participants responded with higher levels of confidence in the same than different photograph condition. The nonparametric test also concurred with this result \([U(3174)= 180784.00, z = -.26.73, p< .001, r= .474]\). In line with findings in the individual response types, the main effect of photograph condition was driven by an interaction between photograph condition and response type \([F(1, 3160)= 110.79, p< .001]\), such that for hit responses, participant confidence was higher for same than different photograph conditions \([t(984.434)= 28.07, p< .001; U(2703)= 274536.00, z= -.28.01, p< .001]\), but this was not the case for false positive responses \([t(469)= 1.32, p= .188]\). This

### Table 3.9

Mean (SD) confidence levels and sample sizes \((n)\) for hit and false positive responses \((FPs)\) in sequential and simultaneous lineups when the target was presented early and when it was presented late in same and different photograph conditions.

<table>
<thead>
<tr>
<th>Photograph (n=3176)</th>
<th>Target Position (n=3176)</th>
<th>Lineup Procedure (n=3176)</th>
<th>n</th>
<th>Hits (n=2705)</th>
<th>n</th>
<th>False Positives (n=471)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same (n=2118)</td>
<td>Early (n=1062)</td>
<td>Simultaneous Lineups (n=622)</td>
<td>609</td>
<td>4.54 (0.83)</td>
<td>13</td>
<td>2.00 (1.16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential Lineups (n=440)</td>
<td>414</td>
<td>4.69 (0.70)</td>
<td>26</td>
<td>2.88 (1.18)</td>
</tr>
<tr>
<td>Late (n=1056)</td>
<td>Simultaneous Lineups (n=613)</td>
<td>603</td>
<td>4.52 (0.84)</td>
<td>10</td>
<td>2.10 (1.10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential Lineups (n=443)</td>
<td>340</td>
<td>4.68 (0.65)</td>
<td>103</td>
<td>2.62 (0.91)</td>
<td></td>
</tr>
<tr>
<td>Different (n=1058)</td>
<td>Early (n=502)</td>
<td>Simultaneous Lineups (n=293)</td>
<td>221</td>
<td>3.27 (1.22)</td>
<td>72</td>
<td>2.42 (1.15)</td>
</tr>
<tr>
<td></td>
<td>Sequential Lineups (n=209)</td>
<td>145</td>
<td>3.52 (1.06)</td>
<td>64</td>
<td>2.45 (0.93)</td>
<td></td>
</tr>
<tr>
<td>Late (n=556)</td>
<td>Simultaneous Lineups (n=302)</td>
<td>216</td>
<td>3.00 (1.25)</td>
<td>86</td>
<td>2.21 (1.03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential Lineups (n=254)</td>
<td>157</td>
<td>3.35 (1.15)</td>
<td>97</td>
<td>2.66 (1.11)</td>
<td></td>
</tr>
</tbody>
</table>
interaction, however, did not drive the main effect of response type. This was confirmed with Mann-Whitney U tests indicating that there was a significantly higher level of confidence for hits compared to false positives in both the same [U(3174)= 22635.50, z= -21.10, p< .001] and different photograph conditions [U(3174)= 74014.50, z= -9.88, p< .001]. There was also an interaction between lineup procedure, response type and photograph [F(1, 3160)= 5.44, p= .020]. This indicated the differing patterns of confidence between sequential and simultaneous lineups for same and different photograph conditions for hit and false positive responses (see individual response types analyses). No other interactions or main effects were significant.

The current findings are consistent with literature showing that there is a correlation between confidence and accuracy (Bradfield et al., 2002; Brewer et al., 2002; Brewer & Wells, 2006; Lindsay et al., 1998; Lindsay & Wells, 1985) and literature showing that there is a positive relationship between confidence and accuracy among choosers, or people who make a positive identification choice (Brewer & Wells, 2006; Fleet et al., 1987; Sauer et al., 2010; Weber & Brewer, 2004, 2006). Conversely, present findings are inconsistent with previous studies that have found little evidence for a relationship between confidence and accuracy (Brewer et al., 2002; Cutler & Penrod, 1989; Cutler et al., 1987; Krug, 2007; Malpass & Devine, 1981a; Sporer et al., 1995; Thomson, 2003). The current results, collectively, therefore appear consistent with the proposition of the USA Supreme Court that eyewitness confidence is a good predictor of accuracy (Neil v Biggers, 1972).

Further consideration of confidence ratings, however, support that the predictive value of confidence is likely limited, and that there may be significant problems with relying on confidence as a measure of accuracy. Although on average confidence for correct identifications was higher than that for false identifications, the spread of scores for both response types spanned the 5-point Likert scale. As such, while there may have been a significant difference between the group means, the spread of scores indicates that some participants may have been highly confident in their false identifications compared with other participants who may have been far less confident in their correct identifications. This observation indicates that while confidence may be related to accuracy in general, people may have different baselines. Some people may be highly confident all the time, whereas others may
lack confidence or have a normal spread in their confidence ratings. Therefore, a witness’s confidence rating alone has little meaning without some knowledge of that witness’s base level of confidence. Given that only one observation of confidence is afforded in criminal investigations and the courts are not afforded a baseline for a witness’s confidence, the current experiment suggests that confidence does not represent a reliable indicator of a witness’s accuracy, as the base level of confidence of one witness may be much higher than the base level of confidence of another witness. As such, using confidence ratings as indicators of accuracy in criminal investigations is may be risky. The observation that people who completed sequential lineups exuded more confidence in their false identifications, and higher confidence on average, than those who completed simultaneous lineups despite being less likely to be accurate exemplifies the inconsistency of the relationship between confidence and accuracy. As such, current findings suggest that it is likely to be dangerous to rely on eyewitness confidence as an indicator of accuracy and, therefore, that jurors should be warned of the inconsistencies in the relationship between confidence and accuracy.

It must also be noted that many variables have been shown to augment and exaggerate confidence ratings, such as repeated questioning, suggestibility, exposure to other witness reports, exposure to mug shots, or exposure to incorrect details subsequent to eyewitness identification taking place (Deffenbacher, Bornstein, & Penrod, 2006; Dysart, Lindsay, Hammond, & Dupuis, 2001; Hastie et al., 1978; Loftus, 1979; Shaw, 1996; Wells & Murray, 1983). As such, although the current study indicates that confidence for some witnesses may be a useful indicator of accuracy immediately following an identification, it may also be misleading in the absence of an examination of the presence or absence of other factors which may artificially influence witness confidence prior to participation in the lineup. The current investigation, however, did not assess the relationship between confidence and accuracy for non-choosers, or people who did not identify a target from the lineup due to the inclusion of only target present lineups. This issue will be examined in Experiment 2.

3.3 Summary

The major findings central to the hypotheses and aims of Experiment 1 are as follows:
- Simultaneous lineups resulted in more hits than sequential lineups when the same photograph of the target person as was originally presented was in the lineup, but there was no consistent advantage of simultaneous lineups over sequential lineups when a different photograph of the target was in the lineup.

- Simultaneous lineups produced less false positives than sequential lineups, more so when the same photograph of the target as had previously been viewed was in the lineup.

- Early presentation resulted in more misses than late presentation in sequential lineups when the photograph of the target person in the lineup was different from that originally presented.

- With late presentations, there was an increase in false positives for sequential lineups compared to early presentations. There was also more false positives in sequential than simultaneous lineups there was little difference in the number of false positives between lineup procedures when the target was presented early.

- Participants who viewed sequential lineups reported higher confidence in hits, and false identifications when within-subjects factors were collapsed, than those who viewed simultaneous lineups.

- Where the same photograph of the target person as viewed earlier was presented in the lineup, participants were more confident in their hits and were less confident in their misses than when a different photograph was presented.

- Participants reported being more confident in their hits than their false positive identifications.

Many of the findings from Experiment 1 are novel within the psychological literature. More specifically, findings of no evidence of a sequential superiority effect was observed and results suggested that a close match between two photographs of an unfamiliar target person increases the superiority of simultaneous over sequential lineups in terms of correct identifications. Furthermore, the current findings suggested that confidence did not necessarily align with accuracy and particularly that confidence in false identifications may be elevated for sequential lineups. Given that these findings have not been observed elsewhere in the literature,
further study testing the replicability of these effects would be need to confirm the reliability of findings. Experiments 2 and 3 will aim to test the reliability of results observed in Experiment 1. Additionally, the current experiment only included lineups where the target person was present. Experiment 2 aims to build on Experiment 1 by examining effects of lineup procedure, photograph, and target position in lineups where the target was absent.
CHAPTER 4
Experiment 2: Target Present and Absent

The results of Experiment 1 generally indicate that while identification accuracy in simultaneous and sequential lineups are similar when the target is presented early in the lineup, there is a dramatic decrease in accuracy in sequential lineups when the target is presented late in the lineup. Most importantly, every measure of the current data shows markedly different results to those found by Lindsay and Wells (1985). Further, although Experiment 1 provided some evidence that confidence of correct identifications may be higher than that of incorrect identifications, some inconsistencies in this relationship were apparent. Specifically, there was some evidence that confidence in incorrect identifications may have been higher in sequential lineups. Thus, replication of these findings would be important to ascertain their strength and robustness.

In criminal court cases it is not known whether or not the true perpetrator is present in the lineup or not. However, in the empirical psychological literature it is known that the risk to a suspect is enhanced in target absent lineups (Clark et al., 2008; Malpass & Devine, 1981a; Wells, 1993). Logically this may be the case in real life lineups - if the true perpetrator is not present in the lineup, then a suspect may be at heightened risk of mistaken identification by eyewitnesses. In the empirical psychological literature, this risk has been found to be particularly great if identifying witnesses believe, or receive information implying, that the perpetrator is present in the lineup (Clark, 2005; Malpass & Devine, 1981a; Steblay, 1997). As previously mentioned, some empirical psychological literature has shown that when a target person is removed from the lineup, witnesses shift their responses to pick foils rather than rejecting lineups (Wells, 1993; Wells et al., 1998). Historically, the target to foils shift was assumed to be more problematic in simultaneous lineups, where lineup members are presented together (Wells, 1993; Wells et al., 1998), however more recent research demonstrated a similar target to foil shift for sequential lineups as well as simultaneous lineups (Clark & Davey, 2005). On the basis of these collective findings, the absence of the target in lineups may be more likely to lead to identification of an innocent suspect and thus represents an important factor for consideration.
Similarly, from a practical point of view considering outcomes in target absent lineups has been considered to be extremely important (Clark et al., 2008; Clark & Tunnicliff, 2001; Malpass & Devine, 1981a; Wells, 1993). This is because procedural changes that enhance accuracy in target present lineups (i.e. increase correct identifications of the target) may come at the cost of decreasing accuracy in target absent lineups (i.e. increasing false identifications of foils or the target replacement). Including the target as present or absent in the lineup in psychological experiments has been seen as a way to simulate the real life situation where the true perpetrator may be present in the lineup, or absent and replaced with an innocent suspect. Thus, the outcomes of procedural changes in target absent lineups represents a measure of risk to innocent suspects in real life situations (Clark et al., 2008; Clark & Tunnicliff, 2001; Malpass & Devine, 1981a; Wells, 1993). The consideration of the impact of system variables on accuracy in target absent lineups as well as target present lineups is pertinent.

Both the empirical psychological literature, and practical considerations of real life lineup situations, indicate that examining outcomes when the target is absent as well as when they are present is important. This import is highlighted when considering it is when the perpetrator is absent from a lineup in real life criminal cases that an innocent suspect may be at risk of conviction. Identification accuracy in Experiment 1 was limited to lineups where a target was always present. Experiment 2 investigated the impact of lineup procedure, photograph condition and target position on accuracy and confidence when the target was present in the lineup and when the target was absent from the lineup. In light of previous literature and the findings of Experiment 1, it was hypothesised that:

1. Evidence of a sequential superiority effect would be observed, at least in terms of a reduced occurrence of false identifications in sequential compared to simultaneous lineups.

2. Visual similarity between the originally observed photograph and the photograph observed in target present lineups would impact upon accuracy and confidence. In other words, viewing the same photograph of the target person in the lineup as was originally observed would lead to more correct identifications and less incorrect rejections of the lineup as well as higher
confidence in correct identifications compared to viewing a different photograph; whereas placing the same background as seen behind the originally observed target behind the target replacement in target absent lineups would interfere with accuracy leading to more false identifications compared to a different background.

3. There would be higher confidence in correct than false identifications, and that there would be higher confidence in correct than incorrect rejections of the lineup but to a lesser extent than the difference seen between correct and false identifications. Further, if the proposition of the USA Supreme courts is accurate, then: 1) confidence in making correct identifications would be higher than confidence in making false identifications; 2) patterns of accuracy would align with patterns of confidence such that under conditions where increased accuracy was observed, higher confidence would also be observed. If the proposition of the courts in the UK is more accurate, then either the level of confidence in correct identifications would not differ from that in false identifications, or conditions where increased accuracy was observed would not align with conditions where higher confidence was observed, or both would occur.

4.1 Method

Experiment 2 was an independent study from Experiment 1 and utilised different participants. Thus, in contrast to Experiment 1, Experiment 2 was not part of a larger study and was not designed to investigate the variables of ethnicity and gender.

4.1.1 Participants

Participants were 297 (207 women) volunteers aged 18-70 years, with a mean age of 33.23 years (SD= 13.20 years). Of this sample, 242 (179 women) participants were Caucasian (81.5% of the sample), whereas 17 participants were Asian and 6 were of African origin. The remaining 32 participants were comprised of a mixture of Pacific Islanders, Middle Eastern persons and persons of mixed or other ethnicities. Participants were recruited through advertisements to complete an anonymous online survey involving photographic lineup judgments. Recruitment
numbers prior to exclusions were 465. Participants were excluded from the study if they did not complete the survey. Participants under 18 years of age were excluded for ethical restrictions involving participants under the age of 18. No respondents reported being under the age of 18 years. On the basis of not completing the survey, 168 participants were excluded, leaving an effective sample of 297.

The protocol conformed to the Declaration of Helsinki and had prior approval of the local Human Subjects Ethics Committee. All participants gave informed consent online prior to their participation in the survey. A copy of the advertisement and the online plain language statement with consent checkbox can be seen in Appendix A. Participants were given no incentive to participate in the study. Procedures were approved by Deakin University’s Human Ethics Committee (#DUHREC-H92_2013; Appendix B).

4.1.2 Design

Patterns of accuracy and confidence of participants were studied using a between-subjects design with three factors. The design of the study was a 2 (sequential vs simultaneous lineups) x 2 (target present vs absent) x 2 (same or different photograph of the target/same or different background as the target behind the target replacement) x 2 (early vs late position of the target or target replacement photograph) between-subjects design. A between-subjects design was considered appropriate for comparing confidence levels given that in a courtroom scenario, only one confidence rating is provided by a witness so no comparison with other kinds of responses from that witness can be made.

For target present lineups, dependent variables were the same as for Experiment 1 (the mean number of hits, misses or false positives of foils, as well as confidence ratings in the decision, measured on a 5-point Likert scale). For target absent lineups, dependent variables were the mean number of correct rejections or not identifying any lineup member as the target, false identification of the target replacement, and false identification of a lineup foil, as well as confidence ratings in these decisions measured on a 5-point Likert scale. Whether or not the target replacement was selected was also added as an additional independent variable in analyses relating to confidence of false positive responses.
4.1.3 Procedures

All lineup procedures were identical to the first experiment with the exception of the number of lineups included in the survey. The format of the online survey was the same as in Experiment 1. The online survey consisted of four lineups and participants were randomly allocated to either sequential or simultaneous lineup procedure conditions. On average, participants took 7.45 (SD= 5.80) minutes to complete the survey when viewing simultaneous lineups and 12.78 (SD= 12.32) minutes when viewing sequential lineups.

4.1.4 Materials and apparatus

4.1.4.1 Photographs

Faces were drawn from CMU Multi-PIE Face Database (Gross et al., 2010), and thus were identical in orientation, composition and style to that reported in Experiment 1.

4.1.4.2 Lineup composition

All lineups included eight Caucasian male lineup members. Lineup foils were selected based on similarity to the target person, as is the standard procedure in the wider literature. Similarity was judged utilising the same procedure as that in Experiment 1. The lineup foil ranked as most similar to the target person by four experimenters was used as the target replacement for target absent photographs.

In half of the four lineups, the target person was present and in the other half the target person was absent from the lineup. In the target absent lineups target replacements were photographs that were similar to that of the photograph of the target. For three out of four lineups, the target replacement was the most frequently identified foil from the previous experiment. For the remaining lineup, photographs not already included in the other three lineups were judged by a panel of five research staff (four female) to determine the most visually similar photograph for use as the target replacement.

Of the four lineups, either participants viewed target present lineups with same photographs and target absent lineups with different photographs, or they viewed target absent lineups with same photographs and target present lineups with different photographs. For target present lineups, the photograph conditions referred to presenting either the same photograph as was originally viewed or a different
photograph than was originally viewed of the target person which included a different background. In the different photograph condition, targets in photographs did not differ in distance from or orientation to the camera compared to photographs in the same photograph condition. For target absent lineups, the same and different photograph conditions referred to presenting the target replacement with either the same background as that viewed behind the originally presented target person, or a different background from that viewed behind the originally presented target person. As such, the same and different photograph condition for target absent lineups referred to the background of the photograph, rather than the whole photograph because the target person was not presented in the lineup. This was to investigate the effect of placing the target replacement within the same, or a different, context compared with the the target person. See Figure 4.1 for examples of the same and different photograph condition for target absent and target present lineups.

![Figure 4.1.](image-url) Examples of photographic lineups including target photographs (left-most column), for the same (top two panes) and different (bottom two panes) photograph conditions when the target was present (first and third panes, target person is lineup member 2) and absent (second and fourth panes, target replacement is lineup member 2). Different photographs for both target present and target absent lineups contain a different background of the target person.

Similarly, early and late presentation for target present lineups referred to the target being presented early or late in the lineup, whereas for target absent lineups
this referred to the target replacement being presented early or late. Similar to photograph condition, participants either viewed targets in target present lineups early (position 2 or 3) and target replacements in target absent lineups late (position 6 or 7), or visa versa. No participant contributed to both levels of the target position and the photograph condition variable for target present or target absent lineups. However, participants all participated in different conditions, such that a different combination of these three variables was viewed in lineups. (See Table 4.1 for possible conditions.) Participants were randomly, but evenly, allocated into one of these conditions. Presentation order of the four lineups was counterbalanced.

4.1.5 Data reduction & statistical analyses

Data reduction and statistical analyses were similar to that of Experiment 1. Participant answers were collected via Qualtrics™ (Provo, USA) and participant data for accuracy (hits, misses, false positives of foils, false positives of target replacements, and correct rejections) and confidence ratings were collated and numbers for each response type for accuracy, and the confidence level for each judgment was calculated. The four lineups for each participant were separated and entered into the analyses as separate subjects so that both accuracy and confidence could be assessed as between-subjects factors, or across lineups using lineups as participants. This was done because target absent and target present lineups were analysed separately and as such, none of the variables of interest could be used as within-subjects variables. In order to thoroughly investigate patterns of accuracy and confidence in each of the response types, 14 separate analyses were conducted and are outlined below.

4.1.5.1 Accuracy measures

Patterns of accuracy were investigated using six separate analyses: Three 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph) x 2 (target position early vs late) univariate ANOVAs were conducted on the dependent measures of hits, misses and false positives in target present lineups. Similarly, in target absent lineups, three 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph condition) x 2 (target replacements position early vs late) univariate ANOVAs were conducted on dependent measures of correct rejections, false positives of foils and false positives of target replacements.
Table 4.1
Participants were randomly allocated to conditions. They either participated in simultaneous or sequential lineups. Black font corresponds to target present lineups, whereas grey font corresponds to target absent lineups. Note, that the variables of photograph (same vs different photographs of the target or same vs different backgrounds behind the target replacement) and target position (early presentation at position 2 or 3 or late position at position 6 or 7), are perfectly balanced for each participant, such that they represent between-subjects factors when considering target present and target absent lineups in isolation. NB: TP = target present, TA = target absent.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Lineup 1</th>
<th>Lineup 2</th>
<th>Lineup 3</th>
<th>Lineup 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Same/Early/TP</td>
<td>Same/Early/TP</td>
<td>Different/Late/TA</td>
<td>Different/Late/TA</td>
</tr>
<tr>
<td>2</td>
<td>Different/Early/TP</td>
<td>Different/Early/TP</td>
<td>Same/Late/TA</td>
<td>Same/Late/TA</td>
</tr>
<tr>
<td>3</td>
<td>Same/Late/TP</td>
<td>Same/Late/TP</td>
<td>Different/Early/TA</td>
<td>Different/Early/TA</td>
</tr>
<tr>
<td>4</td>
<td>Different/Late/TP</td>
<td>Different/Late/TP</td>
<td>Same/Late/TA</td>
<td>Same/Late/TA</td>
</tr>
<tr>
<td>5</td>
<td>Same/Early/TA</td>
<td>Same/Early/TA</td>
<td>Different/Late/TP</td>
<td>Different/Late/TP</td>
</tr>
<tr>
<td>6</td>
<td>Different/Early/TA</td>
<td>Different/Early/TA</td>
<td>Same/Late/TP</td>
<td>Same/Late/TP</td>
</tr>
<tr>
<td>7</td>
<td>Same/Late/TA</td>
<td>Same/Late/TA</td>
<td>Different/Early/TP</td>
<td>Different/Early/TP</td>
</tr>
<tr>
<td>8</td>
<td>Different/Late/TA</td>
<td>Different/Late/TA</td>
<td>Same/Early/TP</td>
<td>Same/Early/TP</td>
</tr>
<tr>
<td>Sequential</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Same/Early/TP</td>
<td>Same/Early/TP</td>
<td>Different/Late/TA</td>
<td>Different/Late/TA</td>
</tr>
<tr>
<td>10</td>
<td>Different/Early/TP</td>
<td>Different/Early/TP</td>
<td>Same/Late/TA</td>
<td>Same/Late/TA</td>
</tr>
<tr>
<td>11</td>
<td>Same/Late/TP</td>
<td>Same/Late/TP</td>
<td>Different/Early/TA</td>
<td>Different/Early/TA</td>
</tr>
<tr>
<td>12</td>
<td>Different/Late/TP</td>
<td>Different/Late/TP</td>
<td>Same/Early/TA</td>
<td>Same/Early/TA</td>
</tr>
<tr>
<td>13</td>
<td>Same/Early/TA</td>
<td>Same/Early/TA</td>
<td>Different/Late/TP</td>
<td>Different/Late/TP</td>
</tr>
<tr>
<td>14</td>
<td>Different/Early/TA</td>
<td>Different/Early/TA</td>
<td>Same/Late/TP</td>
<td>Same/Late/TP</td>
</tr>
<tr>
<td>15</td>
<td>Same/Late/TA</td>
<td>Same/Late/TA</td>
<td>Different/Early/TP</td>
<td>Different/Early/TP</td>
</tr>
<tr>
<td>16</td>
<td>Different/Late/TA</td>
<td>Different/Late/TA</td>
<td>Same/Early/TP</td>
<td>Same/Early/TP</td>
</tr>
</tbody>
</table>

4.1.5.2 Confidence ratings
Patterns of confidence were investigated using eight separate analyses: Three 2 (sequential vs simultaneous lineups) x 2 (same vs different) x 2 (target position early or late) univariate ANOVAs were conducted to assess differences in confidence for hits, misses and false positives to foils in lineups where the target was present. Similarly where the target was absent, three 2 (sequential vs simultaneous lineups) x 2 (same vs different) x 2 (target replacement position early vs late) univariate
ANOVA were conducted to assess differences in confidence for correct rejections, false positives to foils and false positives to target replacements.

In addition, confidence ratings to hits and false positives were compared using a 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph) x 2 (target absent vs present) x 2 (hit in target present lineups vs all false positive responses in both target present and absent lineups) x 2 (target/target replacement position early vs late) univariate ANOVA. Differences in confidence between miss and correct rejection responses in sequential and simultaneous lineups, were assessed using a 2 (sequential vs simultaneous lineups) x 2 (same vs different photograph) x 2 (miss in target present lineups vs correct rejections in target absent lineups) x 2 (target/target replacement position early vs late) univariate ANOVA.

4.2 Results and discussion

All analyses for Experiment 2 assessed variables across lineups, rather than within individuals. There were 1188 lineups in total (594 target present and 594 target absent lineups). Table 4.2 displays sample sizes for hit, miss, false positive and correct rejection responses.

4.2.1 Accuracy

There were 401 hit, 134 miss and 59 false positive responses. Table 4.3 displays total lineup sample sizes and participant mean number of responses for hit, miss and false positive responses when the target was present.

For false positive responses in target present lineups, there were two conditions that had a mean and standard deviation of zero due to no false positive responses in these conditions. As such, it was not appropriate to conduct an ANOVA on the mean number of false positive responses when the target was present including lineup procedure as a variable. In order to investigate the effects of target position and photograph condition on the number of false positive responses where the target was present in sequential lineups, a 2 (target position) x 2 (photograph condition) univariate ANOVA was conducted on the mean number of false positive responses in only the 358 sequential lineups (Table 4.3).
Table 4.2
Sample sizes for hit, miss, correct rejection and false positive confidence ratings for target absent and present sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: N/A refers to conditions where the relevant response type was not possible.

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Lineup Procedure</th>
<th>Target Presence</th>
<th>Target Position</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives (Replacement Selected)</th>
<th>Correct Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same (n= 594)</td>
<td>Simultaneous (n= 236)</td>
<td>Present (n=114)</td>
<td>Early (n=62)</td>
<td>60</td>
<td>2</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=52)</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Absent (n=122)</td>
<td></td>
<td>Early (n=60)</td>
<td>N/A</td>
<td>N/A</td>
<td>15 (4)</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=62)</td>
<td>N/A</td>
<td>N/A</td>
<td>17 (3)</td>
<td>45</td>
</tr>
<tr>
<td>Sequential (n= 358)</td>
<td>Present (n=178)</td>
<td></td>
<td>Early (n=86)</td>
<td>75</td>
<td>1</td>
<td>10</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=92)</td>
<td>70</td>
<td>5</td>
<td>17</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Absent (n=180)</td>
<td></td>
<td>Early (n=92)</td>
<td>N/A</td>
<td>N/A</td>
<td>33 (4)</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=88)</td>
<td>N/A</td>
<td>N/A</td>
<td>39 (11)</td>
<td>49</td>
</tr>
<tr>
<td>Different (n= 594)</td>
<td>Simultaneous (n= 236)</td>
<td>Present (n=122)</td>
<td>Early (n=62)</td>
<td>33</td>
<td>24</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=60)</td>
<td>34</td>
<td>24</td>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Absent (n=114)</td>
<td></td>
<td>Early (n=52)</td>
<td>N/A</td>
<td>N/A</td>
<td>12 (0)</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=62)</td>
<td>N/A</td>
<td>N/A</td>
<td>9 (2)</td>
<td>53</td>
</tr>
<tr>
<td>Sequential (n= 358)</td>
<td>Present (n=180)</td>
<td></td>
<td>Early (n=88)</td>
<td>40</td>
<td>40</td>
<td>8</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=92)</td>
<td>38</td>
<td>37</td>
<td>17</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Absent (n=178)</td>
<td></td>
<td>Early (n=92)</td>
<td>N/A</td>
<td>N/A</td>
<td>19 (3)</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Late (n=86)</td>
<td>N/A</td>
<td>N/A</td>
<td>27 (2)</td>
<td>59</td>
</tr>
</tbody>
</table>
Table 4.3
*Mean (SD) numbers of responses and sample sizes for hits, misses, and false positive responses in lineups where the target was present for sequential and simultaneous lineups in the same or different photograph condition for early and late target presentations.*

<table>
<thead>
<tr>
<th>Photograph Target Position</th>
<th>Lineup procedure</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Photograph (n= 292)</td>
<td>Simultaneous (n=62)</td>
<td>0.97 (0.18)</td>
<td>0.03 (0.18)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td>Sequential (n=86)</td>
<td>0.87 (0.33)</td>
<td>0.01 (0.11)</td>
<td>0.12 (0.32)</td>
</tr>
<tr>
<td>Early (n= 148)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late (n= 144)</td>
<td>Simultaneous (n=52)</td>
<td>0.98 (0.14)</td>
<td>0.02 (0.14)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td>Sequential (n=92)</td>
<td>0.76 (0.43)</td>
<td>0.05 (0.23)</td>
<td>0.19 (0.39)</td>
</tr>
<tr>
<td>Different Photograph (n= 302)</td>
<td>Simultaneous (n=62)</td>
<td>0.53 (0.50)</td>
<td>0.39 (0.49)</td>
<td>0.08 (0.28)</td>
</tr>
<tr>
<td>Early (n= 150)</td>
<td>Sequential (n=88)</td>
<td>0.46 (0.50)</td>
<td>0.46 (0.50)</td>
<td>0.09 (0.29)</td>
</tr>
<tr>
<td>Late (n= 152)</td>
<td>Simultaneous (n=60)</td>
<td>0.57 (0.50)</td>
<td>0.40 (0.49)</td>
<td>0.03 (0.18)</td>
</tr>
<tr>
<td></td>
<td>Sequential (n=92)</td>
<td>0.41 (0.50)</td>
<td>0.40 (0.49)</td>
<td>0.19 (0.39)</td>
</tr>
</tbody>
</table>

There were 423 correct rejection, 142 false positive responses where a foil was selected and 29 false positive responses where the target replacement was selected. Table 4.4 displays total lineup sample sizes and participants’ mean number of responses for correct rejection and false positive responses when the target was absent and either the target replacement or a lineup foil was selected. Levene’s test of equality of error variances was violated for all reported analyses relating to accuracy.

For the analysis on the number of the target replacements being selected in target absent lineups there was one condition with a mean and standard deviation of zero due to no target replacements being selected in that condition (Table 4.4). As such, it was not appropriate to conduct an ANOVA on the number of target replacements selected including lineup procedure as a comparison, or including both
lineup procedures in the analysis. Results are presented here for descriptive purposes only and no firm conclusions can be drawn.

Table 4.4

<table>
<thead>
<tr>
<th>Photograph</th>
<th>Target Position</th>
<th>Lineup Procedure</th>
<th>Correct Rejections</th>
<th>False Positives (foil)</th>
<th>False Positives (replacement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>Early (n= 152)</td>
<td>Simultaneous (n=60)</td>
<td>0.75 (0.44)</td>
<td>0.18 (0.39)</td>
<td>0.07 (0.25)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=92)</td>
<td>0.64 (0.48)</td>
<td>0.32 (0.47)</td>
<td>0.04 (0.21)</td>
</tr>
<tr>
<td></td>
<td>Late  (n= 150)</td>
<td>Simultaneous (n=62)</td>
<td>0.73 (0.45)</td>
<td>0.23 (0.42)</td>
<td>0.05 (0.22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=88)</td>
<td>0.56 (0.50)</td>
<td>0.34 (0.48)</td>
<td>0.13 (0.33)</td>
</tr>
<tr>
<td>Different</td>
<td>Early (n= 144)</td>
<td>Simultaneous (n=52)</td>
<td>0.77 (0.43)</td>
<td>0.23 (0.43)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=92)</td>
<td>0.79 (0.41)</td>
<td>0.17 (0.38)</td>
<td>0.03 (0.18)</td>
</tr>
<tr>
<td></td>
<td>Late  (n= 148)</td>
<td>Simultaneous (n=62)</td>
<td>0.86 (0.36)</td>
<td>0.11 (0.32)</td>
<td>0.03 (0.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=86)</td>
<td>0.69 (0.47)</td>
<td>0.29 (0.46)</td>
<td>0.02 (0.15)</td>
</tr>
</tbody>
</table>

4.2.1.1 Lineup procedure

There was a main effect of lineup procedure on the number of participants’ hit responses [F(1, 586)= 15.00, p< .001]. There were more hits in the simultaneous than the sequential lineup procedure [M(SD)= 0.75 (0.43) vs 0.62 (0.49)]. This result was confirmed by a nonparametric independent samples Mann-Whitney U Test [U= 36696.00, z= -3.34, p= .001, r= .137]. There was also a main effect of lineup procedure on correct rejection responses [F(1, 586)= 7.91, p= .005]. There were more correct rejections when participants completed simultaneous [M(SD)= 0.78 (0.42)] compared to when they completed sequential lineups [M(SD)= 0.67 (0.47)]. A nonparametric Mann-Whitney U test confirmed this effect [U= 37807.00, z= -2.23, p= .003, r= .113]. For false positive responses where a foil (and not the target replacement) was selected, there was a main effect of lineup procedure [F(1, 114]
There were more false positive responses in the sequential than simultaneous lineups \([M(SD)= 0.28 (0.45) \text{ vs } 0.19 (0.39)]\). This result was confirmed by a nonparametric independent samples Mann-Whitney U test \([U= 46168.00, z= 2.58, p= .005, r= .106]\). No other main effects or interactions were significant. Furthermore, there was a main effect of target position on false positives in target present sequential lineups \([F(1, 354)= 4.76, p= .030; U= 17310.000, z= 2.180, p= .029, r= .115]\). There were more false positives made in sequential lineups when the target was presented late \([M(SD)= 0.19 (0.39)]\) than when it was presented early \([M(SD)= 0.10 (0.31)]\).

The present experiment confirmed findings from Experiment 1, indicating that accuracy was superior in simultaneous than sequential lineups. This finding was extended to target absent lineups. Participants responded with more correct rejections of the lineup as well as less false identifications of foils when the target was absent when they viewed simultaneous lineups compared to when they viewed sequential lineups. The present experiment also confirmed findings from Experiment 1 demonstrating that late target position in sequential lineups may have a negative impact on accuracy. These results are inconsistent with other reported findings in the literature [for example, Lindsay & Wells (1985), Clarke (2005), McQuiston-Surrett and colleagues (2006), Steblay and colleagues (2001); Steblay, Dysart & Wells (2011) and Banks (1970)].

These data are also consistent with explanations posited for Experiment 1 data that participants may engage in a process of comparison across sequential lineups, where they attempt to find the ‘best fit’ among lineup members across sequential lineups. When a target is presented late in the lineup, false identifications are likely to be as a result of identifications made prior to seeing the target face, and when the target is absent from the lineup participants may have felt some pressure to choose before running out of faces in the lineup (Clark, 2005; Clark & Davey, 2005; Clark et al., 2008). Although the current difference in accuracy may have been due to some internal processes, such as anticipation of how many faces were to come, this possibility was not assessed in the current experiment and is therefore speculative. Further, it must be noted that due to no participant responses in some cells of the design, interactions of target position and photograph condition with lineup procedure were not able to be investigated for numbers of false positive responses where the target was present. As such, interpretations are based on findings investigating these differences in only sequential lineups, and although these
did not appear to occur in simultaneous lineups, this could not be tested and thus interpretation is limited in this respect. Further investigation of the interaction between lineup type and target position on accuracy will be undertaken in Experiment 3. Theoretical implications of current findings will be discussed further in Chapter 6.

4.2.1.2 Photograph condition (same versus different)

There was a main effect of photograph condition on the number of hits [F(1, 586)= 130.76, p< .001]. There were more hits when the same compared to when a different photograph of the target person was viewed in the lineup [M(SD)= 0.88 (0.33) vs 0.48 (0.50)]. This result was confirmed by a nonparametric independent samples Mann-Whitney U Test [U= 26606.00, z= -10.31, p< .001, r= .423]. There was also a main effect of photograph condition on the number of participants’ miss responses [F(1, 586)= 148.06, p< .001]. There were less misses when the same compared to when a different photograph of the target person was viewed in the lineup [M(SD)= 0.03 (0.17) vs 0.41 (0.49)]. This result was confirmed by a nonparametric independent samples Mann-Whitney U test comparing same to different photographs [U= 60983.00, z= 11.16, p< .001, r= .458]. In target absent lineups, there was a main effect of photograph condition on correct rejection responses [F(1, 586)= 8.19, p=.004]. There were more correct rejections when a different background than that originally observed behind the target person was presented in the lineup [M(SD)= 0.77 (0.42)] than when the same background was presented [M(SD)= 0.66 (0.48)]. A nonparametric Mann-Whitney U test also confirmed this effect [U= 49159.00, z= 3.09, p=.001, r=.127].

The impact of presenting the same photograph of the target person in the lineup as was originally viewed on accuracy when the target was present in the current experiment was consistent with Experiment 1. Participants were more accurate when they viewed the same photograph than when they viewed a different photograph of the target person in the lineup as that on original observation. However, it should be noted that both experiments 1 and 2 used a same or different photograph and did not vary similarit, perse. It may be that different cognitive processes are used for judging similarity than for judging whether a photograph is the same or different. Arguably, if the same photograph, or image, is used on test and retest, this may not test face recognition but rather image matching ability. Thus, the current findings may not be generalisable to similarity, when different photos of
varying similarities are tested. Experiment 3 further investigated the impact of visual similarity using only different photographs of the targets, in order to clarify this possibility. Current findings also suggest that the context in which a given target person is viewed has an impact on subsequent recognition of that person in a lineup. This impact of context is consistent with previous research which has shown that consistency of the context in which a person is viewed originally, compared to when they are viewed upon identification, facilitates recognition of unfamiliar people (Thomson et al., 1982). Similarly, and in line with present data, previous research found that presenting a new person in an old, or previously observed context, led to decreases in correct rejections of the lineup. In other words, placing a person not previously seen, in a context that was familiar led to decreases in accuracy (Thomson et al., 1982). The current results support these findings, suggesting that consistency in context across observations may not only facilitate recognition but also impair accuracy when a different person is placed in a previously observed context. Thus, while reinstating the context in which a perpetrator was originally seen may facilitate correct identification of that perpetrator in a lineup, placing a suspect in the same context as the perpetrator may also decrease the likelihood of correct rejection of the lineup, resulting in increased false identifications.

4.2.2 Confidence

Levene’s equality of error variances test was violated for the analysis on confidence in hit responses as well as that comparing confidence in hit versus false positive responses. Given very small sample sizes in several conditions of the design it was also not appropriate to run the planned ANOVA including both target position and photograph condition as factors for miss responses. As such, these factors were included separately and the analysis was conducted first collapsing across target position and with photograph condition as a factor, and then collapsing across photograph condition including target position as a factor. When target position was collapsed and photograph condition was included, sample sizes were still very small, with only three misses when participants viewed simultaneous lineups and the same photograph was viewed. Therefore, it was not appropriate to conduct the ANOVA. When photograph condition was collapsed and target position was included, sample sizes were sufficient to conduct the ANOVA, with the condition with the smallest sample size being 25.
Due to no participant responses in some of the cells of the design (meaning that no confidence ratings were available for analysis), confidence data for false positive responses where the target was present could not be analysed. This would not have been aided by collapsing successively across the variables of target position and photograph condition as there would still be some conditions in the design that had no, or very low numbers. Data for false positive responses where the target was present are presented here for descriptive purposes, however no definitive conclusions can be drawn from data relating to this response type.

Furthermore, given only one participant response with no standard deviation in two cells of the design for the analysis for miss versus correct rejection responses, it was not appropriate to run the planned ANOVA with photograph and target position included as factors. Thus, interpretation of the analysis for miss versus correct rejection responses was limited by small numbers of participant responses in some conditions of miss responses. Given the absence of any effects or interactions involving photograph or target position in the individual response type analyses and the focus of comparing miss with correct rejection responses for this analysis, these factors were collapsed and only lineup procedure was included as a between-subjects factor. Table 4.5 displays participant mean confidence ratings for hit, miss and false positive responses in lineups where the target was present.

Due to no participant responses in one cell (meaning that no confidence ratings were available for analysis) as well as standard deviations of zero in some cells of the design (meaning that all participants contributing to the mean gave the same response), confidence data for where the target replacement was selected in lineups when the target was absent, could not be analysed using confidence in identifications of the target replacement as the dependent variable. Instead, analysis of all false positive responses in target absent lineups was conducted and whether or not the target replacement was selected was included as an independent variable. This was in order to ascertain whether or not there were interactions between the relevant independent variables and whether or not the target replacement was selected impacting on the confidence of false positive responses.

Table 4.6 displays participant mean confidence ratings for correct rejection and false positive responses when the target replacement was selected and when a different lineup foil was selected in lineups where the target was absent.
Table 4.5
Mean (SD) confidence levels for hit, miss and false positive responses when the target was present in sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: n/a refers to conditions where there was no standard deviation due to a sample size of 1; Nil refers to conditions in which no responses were given so no confidence ratings existed.

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Target Position</th>
<th>Lineup procedure</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>Early</td>
<td>Simultaneous</td>
<td>4.73 (0.58)</td>
<td>2.00 (1.41)</td>
<td>Nil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>4.43 (0.87)</td>
<td>3.00 (n/a)</td>
<td>2.50 (0.97)</td>
</tr>
<tr>
<td></td>
<td>Late</td>
<td>Simultaneous</td>
<td>4.57 (0.70)</td>
<td>5.00 (n/a)</td>
<td>Nil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>4.50 (0.78)</td>
<td>1.60 (0.89)</td>
<td>2.24 (0.83)</td>
</tr>
<tr>
<td>Different</td>
<td>Early</td>
<td>Simultaneous</td>
<td>2.76 (1.23)</td>
<td>3.21 (1.25)</td>
<td>2.60 (0.89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>2.90 (0.93)</td>
<td>2.60 (1.26)</td>
<td>2.25 (0.89)</td>
</tr>
<tr>
<td></td>
<td>Late</td>
<td>Simultaneous</td>
<td>3.27 (1.42)</td>
<td>3.58 (1.14)</td>
<td>3.00 (0.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.00 (1.09)</td>
<td>3.00 (1.05)</td>
<td>2.12 (0.60)</td>
</tr>
</tbody>
</table>

4.2.2.1 Lineup procedure

There was a main effect of lineup procedure on confidence in miss responses [F(1, 133)= 9.50, p= .003]. Confidence in misses was higher when participants viewed simultaneous [M(SD)= 3.37 (1.23)] than when they viewed sequential lineups [M(SD)= 2.72 (1.18)]. There was also a main effect of lineup procedure on confidence of correct rejections [F(1, 415)= 11.46, p= .001]. Reported confidence in correct rejections was higher for simultaneous than for sequential lineups [M(SD)= 3.43 (1.21) vs 3.03 (1.24)]. There was a main effect of lineup procedure [F(1, 156)= 9.07, p= .003] on confidence in false positive responses when the target was absent. There was higher confidence in false positive responses in sequential compared with simultaneous lineups when the target was absent [M(SD)= 2.57 (0.89) vs 2.26 (0.96)]. There was also an interaction between lineup procedure and whether or not the target replacement was selected from the lineup [F(1, 156)= 5.52, p= .020].
Confidence ratings were higher in sequential compared with simultaneous lineups, but this difference was only significant when the replacement was selected from the lineup [replacement selected: $t(27)= 2.69, p=.006, r=.472, M(SD)= 2.70 (0.92)$ vs $1.78 (0.67)$; replacement not selected: $t(140)= 1.06, p=.145, M(SD)$ simultaneous= $2.36 (0.99)$, sequential= $2.54 (0.89)$]. No other main effects or interactions were significant.

**Table 4.6**

*Mean (SD) confidence levels for correct rejection and false positive responses where either the target replacement or a different lineup foil was chosen and when the target was absent in sequential and simultaneous lineups in the same or different photo condition for early and late photograph presentations. NB: Nil refers to conditions in which no responses were given so no confidence ratings existed.*

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Target Position</th>
<th>Lineup procedure</th>
<th>Correct Rejections</th>
<th>False Positives (foil)</th>
<th>False Positives (replacement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>Same</td>
<td>Simultaneous</td>
<td>3.49 (1.22)</td>
<td>2.46 (0.69)</td>
<td>1.75 (0.96)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.10 (1.28)</td>
<td>2.62 (0.73)</td>
<td>3.00 (0.82)</td>
</tr>
<tr>
<td></td>
<td>Late</td>
<td>Simultaneous</td>
<td>3.53 (1.16)</td>
<td>2.64 (0.93)</td>
<td>2.00 (0.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>2.80 (1.21)</td>
<td>2.50 (0.88)</td>
<td>2.73 (1.01)</td>
</tr>
<tr>
<td>Early</td>
<td>Different</td>
<td>Simultaneous</td>
<td>3.25 (1.35)</td>
<td>1.92 (1.00)</td>
<td>Nil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.16 (1.21)</td>
<td>2.63 (1.20)</td>
<td>2.00 (1.00)</td>
</tr>
<tr>
<td></td>
<td>Late</td>
<td>Simultaneous</td>
<td>3.43 (1.14)</td>
<td>2.43 (1.32)</td>
<td>1.50 (0.71)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.00 (1.26)</td>
<td>2.44 (0.87)</td>
<td>3.00 (0.00)</td>
</tr>
</tbody>
</table>

Findings for confidence are inconsistent with the notion that there may be some pressure to choose late in sequential lineups. If this were the case, some change in confidence would have been expected when the target or target replacement was identified late in sequential lineups, compared with early in these lineups. However, no such change in confidence was seen in current results. The current findings suggest that confidence of identification decisions may be a more
reliable predictor of accuracy for simultaneous than sequential lineups when the
target is absent from the lineup. Confidence in correct rejections of the lineup were
higher in simultaneous than sequential lineups whereas confidence in false
identifications when the target replacement was falsely identified was higher in
sequential than simultaneous lineups. These are novel findings that have not been
documented elsewhere within the literature.

One explanation for the finding that confidence may be a more reliable
indicator of accuracy for correct rejections of simultaneous lineups than that of
sequential lineups is that there are differences between lineups that may have
affected participants’ appraisals of the lineups. The higher confidence in correct
rejections for simultaneous rather than that for sequential lineups may have been due
to participants’ metacognitions about the differences between sequential and
simultaneous lineup procedures. In other words, given that participants who rejected
sequential lineups progressed through all eight lineup members, there was likely to
be significantly more time elapsed and more faces presented between the original
observation of the target person and the viewing of the last lineup member in
sequential when compared to simultaneous lineups. This difference would likely
have been greater than in other response types where a decision was made before all
eight lineup members were viewed. As such, participants may have perceived the
sequential lineup procedure as more difficult or arduous given that more time had
elapsed between original observation of the target person and their viewing of all of
the lineup members. This may have lowered confidence due to participants’
evaluation of, or thoughts about, sequential lineups as very difficult and their
subsequent appraisal that their answer may not be reliable. This metacognitive
process may not have occurred for participants who viewed simultaneous lineups due
to the relatively short period of time between original observation of the target
person and presentation of the entire lineup. The current investigation, however, did
not assess participants’ evaluations of lineup types and thus this interpretation
remains speculative. Additionally, given that the current findings have not been
observed previously within the literature, further research would be needed to
confirm the robustness of the current findings. Thus, Experiment 3 will examine
differences in confidence and accuracy between lineup types in order to add to the
current findings.

The finding that confidence was higher for sequential lineups compared with
simultaneous lineups when the target was absent from the lineup and the target
replacement was falsely identified is difficult to explain. This finding is particularly notable as it suggests that confidence may be influenced by system variables that are independent of accuracy. Similarly, these variables may influence accuracy independent of confidence. While differences in confidence may align quite closely with differences in accuracy under some conditions, differences in accuracy between sequential and simultaneous lineups did not align well with differences in confidence in the current experiment. One explanation for this finding is that some of the differences between sequential and simultaneous lineups in general may have contributed to such a result. However it is unclear how these factors may have eventuated in higher confidence in false identifications of target replacements for sequential than simultaneous lineups. Although the reasons for these findings remain unclear, they are nevertheless pertinent to consider if the courts wish to rely on confidence as an indicator of accuracy. Therefore, further research on differences in accuracy and confidence between simultaneous and sequential lineups is needed. Thus, Experiment 3 will further investigate this issue.

4.2.2.2 Photograph condition (same versus different)

There was a main effect of photograph condition on participant confidence in hit responses \([F(1,393)=267.22, p< .001]\). Confidence was higher when the same compared with when a different photograph of the target person as that viewed originally was presented in the lineup \([M(SD)= 4.55 (0.76) vs 2.98 (1.17)]\). This effect was confirmed by an independent samples nonparametric Mann-Whitney U test \([U= 5594.00, z= -12.45, p< .001, r= .662]\). No other main effects or interactions were significant.

Findings relating to the impact of photograph condition on participant confidence are consistent with findings relating to accuracy as well as those from Experiment 1. Both accuracy in terms of correct identifications, and confidence in correct identifications, increased when the same photograph of the target person as was originally observed was viewed in the lineup. This adds further support to findings that similarity between original observation of a target and observation of that target in the lineup increases the likelihood of correct recognition as well as confidence in the identification decision. This issue will be discussed further in Experiment 3. However Experiments 1 and 2 used the same or a different photograph to that originally presented as the target and thus it is not known what impact varying levels of similarity would have on accuracy and confidence.
Experiment 3 will investigate this issue under varying levels of photograph similarity.

4.2.2.3 Confidence for hit versus false positive responses

Table 4.7 displays mean confidence levels for hit and false positive responses in simultaneous and sequential lineups where the target or target replacement was presented early or late in same or different photograph conditions.

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Target Position</th>
<th>Lineup procedure</th>
<th>n</th>
<th>Hits</th>
<th>n</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>Early</td>
<td>Simultaneous</td>
<td>60</td>
<td>4.73 (0.58)</td>
<td>15</td>
<td>2.27 (0.80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>93</td>
<td>4.43 (0.87)</td>
<td>32</td>
<td>2.63 (0.79)</td>
</tr>
<tr>
<td>Late</td>
<td>Simultaneous</td>
<td>51</td>
<td>4.57 (0.70)</td>
<td>17</td>
<td>2.53 (0.88)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>85</td>
<td>4.50 (0.78)</td>
<td>28</td>
<td>2.42 (0.83)</td>
<td></td>
</tr>
<tr>
<td>Different</td>
<td>Early</td>
<td>Simultaneous</td>
<td>33</td>
<td>2.76 (1.23)</td>
<td>17</td>
<td>2.12 (0.99)</td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>75</td>
<td>2.90 (0.93)</td>
<td>43</td>
<td>2.44 (1.09)</td>
<td></td>
</tr>
<tr>
<td>Late</td>
<td>Simultaneous</td>
<td>34</td>
<td>3.27 (1.42)</td>
<td>11</td>
<td>2.36 (1.21)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>70</td>
<td>3.00 (1.09)</td>
<td>55</td>
<td>2.34 (0.78)</td>
<td></td>
</tr>
</tbody>
</table>

There was a significant main effect of response type on confidence. Reported confidence in hits (range= 1-5) was higher than confidence in false positive responses (range= 1-5) overall [F(1, 620)= 44.761, p< .001; M(SD)= 3.98 (1.19) vs 2.42 (0.78)]. There was also an interaction between response type and photograph [F(1, 620)= 22.39, p< .001] that did not drive the main effect of response type. The difference between response types was exaggerated for the same compared with the different photograph condition overall. This was confirmed by nonparametric Mann-
Whitney U tests [same: U= 1828.00, z= -15.16, p< .001, r= .771; M(SD) hits= 4.55 (0.76), false positives= 2.49 (0.82); different: U= 5009.00, z= -4.16, p< .001, r= .226; M(SD) hits= 2.79 (1.17), false positives= 2.33 (0.95)]. (See Table 4.7 below.)

There were also significant main effects and interactions reflective of findings reported in the individual response type analyses. There was a main effect of photograph condition [F(1, 620)= 81.47, p< .001]. Confidence was higher in the same than different photograph conditions overall [U= 24390.00, z= -10.55, p< .001, r= .420]. There was also an interaction between lineup procedure and target presence [F(1, 620)= 4.85, p= .032]. Reported confidence was higher in simultaneous compared with sequential lineups when the target was present [U= 33823.00, z= -4.21, p< .001, r= .173], but higher in the sequential compared with simultaneous lineups when the target was absent [U= 36600.50, z= -2.83, p= .005, r= .116]. No other significant main effects or interactions were observed.

The comparison of confidence in correct versus incorrect identifications is consistent with findings from Experiment 1, as well as findings in the wider literature indicating that confidence is positively associated with accuracy, particularly for choosers (Brewer & Wells, 2006; Fleet et al., 1987; Sauer et al., 2010; Weber & Brewer, 2004, 2006). On the surface, results also appear consistent with the propositions of the USA Supreme Court relating to confidence (Neil v Biggers, 1972; Wise & Safer, 2004).

When further consideration of these data takes place, however it can be noted that variability of confidence levels are different between response types. Where the confidence of hit responses appears positively skewed with many more responses at the top end of the confidence scale of five, responses for the other response categories appear more evenly spread amongst response types. Therefore, while confidence ratings for false positive responses appeared lower in general, some people still responded as highly confident, at a four or a five on the 5-point scale. This implies that while the mean confidence for hits was higher than that of false positive, less is known about the proportion of individuals who may have rated themselves as highly confident for both hit and false positive responses. This may be the reason for differences between the current findings and findings examining a general relationship between confidence and accuracy, and that of the Devlin Report (Devlin, 1976), as much of this report was dedicated to examining individual cases. As such, current data would be difficult to apply within the courtroom setting, as elevated confidence is relative to the general confidence of the individual, so a
baseline level of confidence would be needed to clarify whether a certain confidence rating is relatively high for any individual person. In other words, some people are very confident in general whereas others may be less confident, and so a high confidence rating may either reflect a tendency to be very confident in general, or a statement of high confidence in a particular decision in comparison to other decisions. As such, an assessment of whether or not confidence for any given witness is likely to be a reliable indicator of accuracy would be aided by the attainment of a baseline measure of confidence for that witness. It must be noted, however, that in a courtroom setting, no baseline of an eyewitness’s confidence is taken. Such a baseline may be highly relevant to the question of reliability of any given eyewitness’s testimony, such as that in seminal cases reviewed in the Devlin Inquiry (Devlin, 1976).

4.2.2.4 Confidence for miss versus correct rejection responses

Table 4.8 displays means confidence levels for miss and correct rejection responses in simultaneous and sequential lineups where the target replacement was presented early or late in same or different photograph conditions.

There was no main effect of response type on reported confidence [F(1, 553)= 2.22, p=.137]. There was no significant difference between the reported confidence of miss (range= 1-5) and correct rejection (range= 1-5) responses [M(SD) misses= 2.97 (1.24), correct rejections= 3.21 (1.24)]. This was the case even when all other variables were collapsed [F(1, 555)= 3.65, p=.057].

There was a main effect of lineup procedure [F(1, 553)= 17.87, p< .001]. This effect reflected the higher confidence in simultaneous compared with sequential lineups in the correct rejection response type. This could also be observed from descriptive statistics in the miss response type. Differences between the different conditions can be seen in the means (Table 4.8) and these were generally consistent with differences observed in the analyses of the individual response types. No other significant main effects or interactions were observed.
Table 4.8
Mean (SD) confidence levels for miss and correct rejection responses in simultaneous and sequential lineups where the target/target replacement was presented early or late in same or different photograph conditions. NB: n/a refers to conditions where there was no standard deviation due to a sample size of 1.

<table>
<thead>
<tr>
<th>Photograph Condition</th>
<th>Target Position</th>
<th>Lineup procedure</th>
<th>n misses</th>
<th>Misses (SD)</th>
<th>n CRs</th>
<th>Correct Rejections (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>Early</td>
<td>Simultaneous</td>
<td>2</td>
<td>3.00 (1.41)</td>
<td>45</td>
<td>3.51 (1.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>1</td>
<td>3.00 (n/a)</td>
<td>59</td>
<td>3.10 (1.28)</td>
</tr>
<tr>
<td>Late</td>
<td>Simultaneous</td>
<td>1</td>
<td>5.00 (n/a)</td>
<td>45</td>
<td>3.53 (1.16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>5</td>
<td>1.60 (0.89)</td>
<td>49</td>
<td>2.80 (1.21)</td>
<td></td>
</tr>
<tr>
<td>Different</td>
<td>Early</td>
<td>Simultaneous</td>
<td>24</td>
<td>3.21 (1.25)</td>
<td>40</td>
<td>3.25 (1.35)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>40</td>
<td>2.60 (1.26)</td>
<td>73</td>
<td>3.16 (1.21)</td>
</tr>
<tr>
<td>Late</td>
<td>Simultaneous</td>
<td>24</td>
<td>3.58 (1.14)</td>
<td>53</td>
<td>3.43 (1.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>37</td>
<td>3.00 (1.05)</td>
<td>59</td>
<td>3.00 (1.26)</td>
<td></td>
</tr>
</tbody>
</table>

The finding that there was no difference in participant confidence of incorrect compared with correct rejection responses indicates that the positive relationship between confidence and accuracy may only be present for choosers, or people who identify a person from the lineup. This finding should be interpreted cautiously given that there were no incorrect rejections of the lineup, and therefore no confidence ratings for these responses, in some cells of the design. Results, however, are consistent with Sporer (1995) who found that the correlation between confidence and accuracy was consistently and reliably higher for choosers, or correct versus false identifications, than for non-choosers, or correct versus incorrect rejections of the lineup. The current results confirm and add strength to previous findings which suggest that choosing is an important variable which influences the relationship between confidence and accuracy (Brigham, 1988; Sporer, 1993; Sporer et al., 1995). In other words, that for people who make an identification decision, confidence may
be a more reliable predictor of accuracy than for those who do not. Given the limitations of the current experiment with respect to incorrect rejections, Experiment 3 will further investigate differences in confidence between correct and incorrect rejections.

Current data also suggest that confidence and accuracy may be more highly related for non-choosers under certain circumstances. Specifically, the present study found that higher confidence in correct rejection responses when participants viewed simultaneous lineups aligned with more correct rejections of simultaneous lineups. As such, it is possible that confidence may be more related to accuracy for non-choosers in simultaneous lineups than for sequential lineups.

The difference in findings between choosers and non-choosers may indicate that the process of choosing, or identifying a target person from a lineup may be a different process to that of identifying that a target person is not in a lineup. This possibility would be consistent with findings reported by Clark and colleagues (2008), who found a completely different pattern of diagnosticity for identifications than non-identifications, suggesting that a different process may take place for decisions than non-decisions and different factors may contribute to this decision. Differences in memory processes for various memory tasks has also been seen in other literatures such as the literature which suggests that people are superior when asked to recognise a previously seen item, than when asked to recall this item without any stimulus (Anderson & Bower, 1972; Gillund & Shiffrin, 1984; Jacoby, Craik, & Begg, 1979; Tversky, 1973). These differences have often been suggested to be due to differential encoding and retrieval processes (Jacoby et al., 1979; Tversky, 1973). It is therefore possible that identifying a target that is present in a lineup, which is essentially a facial recognition task, may involve different processes to identifying that a target is not in a lineup when that target is absent from the lineup.

Previous research on visual memory suggests that visual recognition of a previously seen item involves activating sets of memory representations based on the match between the information in the retrieval cue and the information in the memory cue (Atkinson & Juola, 1974). A sufficient match should eventuate in recognition of the item. However, this process may be more complicated when considering the decision that a previously seen item is not present. For example, the inverse of the recognition process may occur, and there may be insufficient information in the retrieval cue that matches the information in the memory cue.
This would presumably result in a failure to recognise an item, or in the case of eyewitnesses, a rejection of the lineup based on insufficient similarity between the memory of the target person and the lineup members. However, a second possibility is that a person may reject an item based on elements in the retrieval cue which are particularly dissimilar when compared with information in the memory cue. The decision to reject may not be based on insufficient matching information, but rather on information that mismatches with information in the memory cue. For example, a particular feature of a lineup member, such as hair or skin colouring, may be particularly different from the previously observed target. This would result in a lineup rejection based on the marked dissimilarity of the feature between the target and the lineup member. Thus, rejecting the lineup may involve different memory processes to identifying a person from a lineup which decreases the relationship between confidence in lineup rejections and accuracy. In other words, the nature of memory processes for identifying a previously observed person may differ from that for identifying the absence of that person. This would be consistent with current findings, and deserves further investigation. Thus, differences in confidence for accurate and inaccurate responses in target present and target absent lineups will be further investigated in Experiment 3.

4.3 Summary

The major findings central to the hypotheses and aims of Experiment 2 are as follows:

- There were more hits when participants viewed simultaneous compared with sequential lineups.
- There were generally more correct rejections of the lineup when participants viewed simultaneous compared with sequential lineups, although this was not the case when a different photograph was presented early in the lineup.
- There were more hits when the same photograph as was originally presented as the target was viewed in the lineup.
- There were less misses when the same photograph as was originally presented as the target was viewed in the lineup.
- There were more false positives made in sequential lineups when the target was presented late than when it was presented early—no firm conclusions could be drawn about simultaneous lineups.
- There were less correct rejections of lineups when target replacements were viewed with the same background as that originally presented behind the target person, compared to when a different background was presented.
- There was higher confidence in false positive responses when participants viewed sequential lineups compared with simultaneous lineups when the target was absent from the lineup and where the target replacement was falsely identified from the lineup.
- There was higher confidence in correct rejection responses for simultaneous compared with sequential lineups.
- Confidence in hits was higher than confidence in false positive responses.

Many of the findings from Experiment 2 are consistent with those of Experiment 1. More specifically, findings of no evidence of a sequential superiority effect was observed and results suggested that a close match between two photographs of an unfamiliar target person increases accuracy and confidence in terms of correct identifications. However, some novel findings were observed. The findings of Experiment 2 suggested that confidence and accuracy may better align when participants viewed simultaneous than sequential lineups and the target was absent. Particularly, while confidence in correct rejections of the lineup was higher in simultaneous lineups, confidence in false identifications when the target replacement was falsely identified was elevated for sequential lineups. Given that these findings have not been observed elsewhere in the literature, further study testing the replicability of these effects would be need to confirm the reliability of findings. Experiment 3 will aim to test the reliability of results observed in Experiment 2. Additionally, Experiments 1 and 2 indicated a marked effect of photograph condition on accuracy. However, Experiments 1 and 2 only used the same or a different photograph to that originally presented as the target and thus it is not known what impact varying levels of similarity would have on accuracy and confidence. Experiment 3 aims to build on Experiment 2 by examining the impact of varying levels of similarity between observations of the target, and possible interactions of similarity with lineup type and target position, on patterns of accuracy and confidence.
CHAPTER 5

Experiment 3: Visual Similarity

The results of the first two experiments conflict with previous research, finding no evidence of a sequential superiority effect, but rather providing some evidence of a simultaneous superiority effect. On the other hand, these experiments have continued to support past literature suggesting that consistency in clothing, background, lighting and orientation between original observation of the perpetrator and observation of the alleged suspect in the lineup enhances accuracy and increases confidence in correct identifications. However, different levels of visual similarity were not created, with experiments examining only the effect of presenting a same or different photograph in the lineup. For example, differences between observations could be minor in that only clothing is changed, very major including differences in hairstyle, clothing, facial hair and lighting, or somewhere in between.

The effects of familiarity and visual similarity on eyewitness accuracy have been previously investigated by empirical psychological research. The idea that a familiar person is more easily recognised than an unfamiliar person receives some support within the empirical psychological literature (Thomson, 2003; Thomson et al., 1982). Differential processing of familiar and unfamiliar faces has been demonstrated by a number of researchers (Bruce, 1982; Ellis et al., 1979; Thomson, 2003). For example, Bruce (1982) examined the effect of changing faces on recognition for familiar and unfamiliar faces across two different experiments. The first experiment tested 36 participants on their ability to recognise photographs of people unfamiliar to them when the photograph either remained the same or was changed, in terms of angle, expression, or both upon testing. The second experiment tested a different 24 participants who were presented photographs of faces, half of which were highly familiar to them and half of which were unfamiliar to them, and upon recognition testing the photographs either remained the same or were changed in both angle and expression. Results indicated that overall, familiar faces were recognised more accurately than unfamiliar faces, although false positive rates were similar for familiar and unfamiliar faces. Further, unfamiliar faces were recognised more slowly and less accurately if changed in angle and expression on test, while familiar faces were recognised more slowly but just as accurately if changed on test. On the basis of these findings, it appeared that familiar faces were generally
recognised more accurately than unfamiliar faces and changes to unfamiliar faces reduced accuracy whereas recognition of familiar faces was preserved when changed in angle and expression.

Similarly, Thomson et al. (1982) conducted a comprehensive study on the role of context and familiarity in recognition of people. Results indicated that changing the pose, clothing and context of the person viewed drastically reduced the likelihood that unfamiliar people would be identified as having been seen in the earlier series of slides, but these changes had little or no impact on identification of familiar people. Additionally, unfamiliar people were much more likely to be falsely identified if their pose and clothing was similar to someone in the previous series of slides compared with familiar people. Further, research suggests that recognition is facilitated when the target person is of the same race or gender as the observer, which has been suggested to be due to increased familiarity with people of the same race and gender as oneself (Brigham & Malpass, 1985; Sporer & Horry, 2011; Wright & Sladden, 2003). These findings together suggest that recognition of unfamiliar people is a much more fragile process than that of familiar people, and is affected by changes in context, hairstyle, clothing, ethnicity, and pose of the individual originally observed (Brigham & Malpass, 1985; Sporer & Horry, 2011; Thomson et al., 1982; Wright & Sladden, 2003). More recent research has also found that the length of time a face is examined in simultaneous lineups is related to how familiar a face is, or how similar to a previously seen face the examined face is, suggesting that similarity or familiarity of faces play an important part in decision making in eyewitness identification tasks (Flowe & Cottrell, 2010).

While familiarity and visual similarity of the perpetrator, in general, represents an estimator variable, visual similarity in terms of context and external features can represent a system variable because it can be manipulated in an identification lineup. For example, if the perpetrator wore a black jacket and a red cap, the lineup administrators could conduct the lineup with all members wearing a black jacket and a red cap. As such, research into the effect of visual similarity on eyewitness accuracy is pertinent to recommendations relating to lineup procedures within the legal system. However, little research has examined the interaction of visual similarity between original observation and observation in the lineup with other system variables that may influence accuracy, such as lineup procedure (Amendola & Wixted, 2015a; Clark, 2012; Clark et al., 2011; Flowe, 2011; Memon,
Given that in a real life eyewitness situation, there would seldom be an exact match between the features of the target seen during commission of the crime and those viewed in the lineup, investigating the impacts of different levels of similarity between observations would be highly valuable. Further, some inconsistencies in findings relating to confidence are apparent across the first two experiments, with data suggesting confidence may be higher in sequential compared with simultaneous lineups. Thus, Experiment 3 aimed to extend the findings of the first two experiments, aiming to further examine the impact of lineup procedure, but also examining the impact of different levels of visual similarity, on accuracy and confidence in both target absent and target present lineups. In light of previous literature and the previous two experiments it was hypothesised that:

1. No evidence of a sequential superiority effect would be observed. There will be no reduction in the occurrence of false identifications in sequential compared to simultaneous lineups.

2. Visual similarity between the originally observed photograph and the photograph observed in the lineup in target present lineups will impact upon accuracy and confidence. In other words, greater visual similarity will lead to more correct identifications and less incorrect rejections of the lineup as well as higher confidence in correct identifications.

3. There will be higher confidence in correct than false identifications, and that there will be higher confidence in correct than incorrect rejections of the lineup but to a lesser extent than the difference seen between correct and false identifications. Further, if the proposition of the USA Supreme courts is accurate, then: 1) confidence in making correct identifications will be higher than confidence in making false identifications; 2) patterns of accuracy will align with patterns of confidence such that under conditions where increased accuracy was observed, higher confidence will also be observed. If the proposition of the courts in the UK is more accurate, then either the level of confidence in correct identifications will not differ from that in false
identifications, or conditions where increased accuracy was observed will not align with conditions where higher confidence was observed, or both would occur.

5.1 Method

Experiment 3 was a second independent study and utilised different participants to the other two experiments. Thus, Experiment 3 was an individual experiment and not part of a larger study.

5.1.1 Participants

Participants were 138 (85 women) volunteers aged 18-70 years, with a mean age of 36.09 years (SD= 12.45 years). Participants were recruited through advertisements to complete an anonymous online survey involving photographic lineup judgments. Participants were excluded from the study if they did not complete the survey or were under the age of 18 years, due to ethical considerations with conducting research in this age group. On the basis of these criteria 84 participants (11 who were under 18 years of age and 73 who did not complete the survey) were excluded. Recruitment numbers prior to exclusion were 222. After exclusion the effective sample was 138.

The protocol conformed to the Declaration of Helsinki and had prior approval of the local Human Subjects Ethics Committee. All subjects gave informed consent online prior to their participation in the survey. A copy of the advertisement and the online plain language statement with consent checkbox can be seen in Appendix A. Participants were given no incentive to participate in the study. Procedures were approved by Deakin University’s Human Ethics Committee (#DUHREC-H92_2013; Appendix B).

5.1.2 Design

Patterns of accuracy and confidence were studied using a 2 (sequential vs simultaneous lineups) x 2 (target present vs target absent) x 3 (highly similar, moderately similar, or least similar photograph of the target compared to the originally seen target photograph) x 2 (early vs late position of the target or target replacement) mixed model design.
Dependent variables were the same as in Experiment 2: participant accuracy in target present lineups (hits, misses or false positives of a lineup foil); participant accuracy in target absent lineups (correct rejections, false identification of the target replacement, and false identifications of lineup foils); and participant confidence in these decisions, measured on a 5-point Likert scale.

Whether or not the target replacement was selected was also added as an additional independent variable where confidence of false positive responses were investigated.

5.1.3 Procedures

All lineup procedures were the same as for previous experiments with the exception of the number of lineups. The format of the online survey was the same general format as in Experiments 1 and 2. The online survey consisted of 12 lineups with eight photographs in each lineup and participants were randomly allocated to either sequential or simultaneous lineup procedures. On average, participants took 12.9 (SD= 4.33) minutes to complete the survey when viewing simultaneous lineups and 16.98 minutes (SD= 4.18) minutes when viewing sequential lineups.

5.1.4 Materials and apparatus

5.1.4.1 Photographs

Photographs were identical in orientation, lighting and appearance as in Experiments 1 and 2 (section 3.1.4). All target photographs had plain blue backgrounds with the face in the foreground (Figure 3.1). All lineup photographs (both targets and foils) had a high backed chair in the background with the face in the foreground (Figure 5.1). Appearance of the lineups were identical to previous experiments in all other aspects.
Figure 5.1. Example of a lineup photograph with high backed chair background as was used for all lineup photographs in experiment four. Taken from Gross et al. (2010).

5.1.4.2 Lineup composition

All lineup targets were Caucasian men. Participants viewed 12 eight-person lineups all of which had different targets, six where the target was present and six where the target was absent. Lineup foils were balanced across lineups based on level of similarity to the target. Target to foil similarity was not, however, manipulated as an independent variable. To judge the level of similarity between the target photograph and target observed in the lineup, as well as between lineup targets and foils, a panel of 11 independent participants (five female) completed two sets of similarity ratings on an 8-point scale ranging from zero to seven: one between the target photograph and 12 different photographs of the target person, and a second between the target photograph and 27 different lineup foils.

Balancing foil to target similarity

All lineups included at least two of the seven foils judged by participants to be most similar to the target photograph (mean ratings for these foils ranged from 1.80 to 2.73 across the 12 lineups). Lineups were then filled with foils of a similar age, and possessing similar colour and features to the target person, as judged by two experimenters. In target present lineups, target photographs were counterbalanced in terms of target position, such that half of the targets were displayed early in the lineup (at position two or three) and half were displayed late in the lineup (at position six or seven). For target absent lineups, the foil judged as most similar to the target photograph was chosen as the ‘target replacement’ and counterbalanced in terms of position in the same way as target photographs in target present lineups.
Development of similarity levels for the independent variable of ‘similarity’

Similarity of the target photograph observed in the lineup to the originally observed target photograph was varied using the similarity ratings provided by the 11 participants mentioned above. Target present lineups contained three photograph conditions with two lineups in each condition: the highly similar condition (HS) included target photographs in the lineups judged by participants as ‘most similar’ to the originally observed target photographs (mean similarity ratings between 6.0 and 7.0); the least similar condition (LS) included target photographs that participants rated as ‘least similar’ to the originally observed target photographs (mean similarity ratings between 3 and 4.2); and the moderately similar condition (MS) had target photographs in the lineup that had participant ratings that were closest to halfway between the HS and MS condition for that target (mean similarity ratings between 4.5 and 5.5). Presentation order of the 12 lineups was counterbalanced using a Latin squares design.

5.1.5 Data reduction & statistical analyses

Participant answers were collected via Qualtrics™ (Provo, USA) and participant data for accuracy and confidence ratings were collated in the same way as for Experiment 2. Analysis was similar to Experiments 1 and 2. Several analyses were conducted in order to investigate differences in confidence and accuracy as a function of the independent variables, and these are outlined below. Where Mauchly’s test of sphericity was violated, the Huynh-Feldt test statistic with adjusted degrees of freedom was reported.

5.1.5.1 Accuracy measures

A series of 2 (sequential vs simultaneous lineups) x 3 (photograph condition) x 2 (target position early vs late) mixed model ANOVAs were conducted on dependent measures of hits, misses and false positives in target present lineups, with target position and photograph condition being treated as within-subjects factors. Similarly, in target absent lineups, a series of 2 (sequential vs simultaneous lineups) x 2 (target replacement position early vs late) mixed model ANOVAs were conducted on dependent measures of correct rejections, false positives of foils and false positives of target replacements, with target position being treated as the within-
subjects factor. Appropriate t-tests and post hoc tests were carried out to assist interpretation of any interaction effects.

5.1.5.2 Confidence ratings

Differences in confidence were analysed across lineups and as such, all factors were treated as between-subjects factors. This was due to the limited numbers of participants responding with a single response type in all levels of the photograph similarity variable. Patterns of confidence were investigated using nine separate analyses.

A series of 2 (sequential vs simultaneous lineups) x 3 (photograph condition) x 2 (target position early vs late) univariate ANOVAs were conducted to assess differences in confidence for hits, misses and false positives to foils in lineups where the target was present. Similarly, where the target was absent, a series of 2 (sequential vs simultaneous lineups) x 2 (target replacement position early vs late) univariate ANOVAs were conducted to assess differences in confidence for correct rejections, false positives to foils, and false positives to target replacements. Differences in confidence for false positive responses were also analysed including both target absent and present lineups in the analysis to examine the effects of target presence on confidence of false positive responses, and these were assessed using a 2 (sequential vs simultaneous lineups) x 2 (absent vs target present) x 2 (target position early vs late) x 2 (replacement selected vs replacement not selected) univariate ANOVA.

In addition, confidence ratings to hits and false positives (including those in target present and target absent lineups) were compared using a 2 (hit vs false positive responses) x 2 (sequential vs simultaneous lineups) x 2 (target/target replacement position early vs late) univariate ANOVA. Differences in confidence between miss responses in target present lineups and correct rejection responses in target absent lineups were assessed using a 2 (miss vs correct rejection responses) x 2 (sequential vs simultaneous lineups) x 2 (target/target replacement position early vs late) univariate ANOVA. Independent samples t-tests were carried out post hoc to assist interpretation of any interaction effects.
5.2 Results and discussion

In order to investigate differences in confidence and accuracy between lineup procedure and photograph condition for hit, miss, false positive and correct rejection responses, 14 separate analyses were conducted. Two additional analyses were conducted to investigate differences in confidence between hit and false positive responses, and between miss and correct rejection responses. All analyses for accuracy in Experiment 3 were assessed using the within-subjects variables of photograph condition and target position. Analyses for confidence, however, were assessed across lineups, rather than within individuals. This was due to limited numbers of response types in all levels of the within-subjects variable. There were 1656 lineups in total. Response type frequencies are listed in the relevant sections below.

5.2.1 Accuracy

There were 405 hit, 288 miss, 391 false positive responses in lineups when the target was present. For all three analyses when the target was present Box’s test of equality of covariance matrices, and Mauchly’s test of sphericity for the interaction effect were violated. Levene’s test of equality of error variances was also violated for analyses on hit, miss and false positive responses in target present lineups - for hit responses, violations occurred for HS and LS early and HS late target presentation; for miss responses, violations occurred for HS and MS late target presentation; and for false positive responses, violations occurred for HS early and HS and MS late target presentations. Table 5.1 displays total lineup sample sizes and participants’ mean numbers of responses for hit, miss and false positive responses when the target was present.

There were 572 correct rejection and 256 false positive responses in total (45 when the target replacement was selected and 211 when a foil was selected) when the target was absent. For the analysis on the number of the target replacements being selected in target absent lineups, Levene’s test of equality of error variances was violated for early target replacement presentation and Box’s test of equality of covariance matrices was also violated.
Table 5.1
*Mean (SD) numbers of responses and sample sizes for hits, misses, and false positive responses in lineups where the target was present for sequential and simultaneous lineups in highly similar (HS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late target presentations.*

<table>
<thead>
<tr>
<th>Target Position</th>
<th>Photograph Condition</th>
<th>Lineup Type</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HS</td>
<td>Simultaneous (n=64)</td>
<td>0.66 (0.48)</td>
<td>0.27 (0.45)</td>
<td>0.08 (0.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.55 (0.50)</td>
<td>0.30 (0.46)</td>
<td>0.15 (0.36)</td>
</tr>
<tr>
<td>Early (n=138)</td>
<td>MS</td>
<td>Simultaneous (n=64)</td>
<td>0.66 (0.48)</td>
<td>0.31 (0.47)</td>
<td>0.03 (0.18)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.65 (0.48)</td>
<td>0.31 (0.47)</td>
<td>0.04 (0.20)</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>Simultaneous (n=64)</td>
<td>0.28 (0.45)</td>
<td>0.61 (0.49)</td>
<td>0.11 (0.32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.18 (0.38)</td>
<td>0.66 (0.48)</td>
<td>0.16 (0.37)</td>
</tr>
<tr>
<td>Late (n=138)</td>
<td>HS</td>
<td>Simultaneous (n=64)</td>
<td>0.78 (0.42)</td>
<td>0.14 (0.35)</td>
<td>0.08 (0.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.64 (0.48)</td>
<td>0.03 (0.16)</td>
<td>0.34 (0.48)</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>Simultaneous (n=64)</td>
<td>0.48 (0.50)</td>
<td>0.36 (0.48)</td>
<td>0.16 (0.37)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.43 (0.50)</td>
<td>0.26 (0.44)</td>
<td>0.31 (0.47)</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>Simultaneous (n=64)</td>
<td>0.32 (0.47)</td>
<td>0.42 (0.50)</td>
<td>0.25 (0.44)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential (n=74)</td>
<td>0.27 (0.45)</td>
<td>0.51 (0.50)</td>
<td>0.22 (0.42)</td>
</tr>
</tbody>
</table>

For false positive responses where the target was absent and a foil (and not the target replacement) was selected Levene’s Test of equality of error variances was violated for early target replacement presentation. Table 5.2 displays total lineup sample sizes and participant mean numbers of correct rejection and false positive responses when the target was absent and either the target replacement or a lineup foil was selected.
Table 5.2
Mean (SD) frequencies and sample sizes for correct rejection and false positive responses where either a lineup foil or target replacement was selected in lineups where the target was absent for sequential and simultaneous lineups for early and late target replacement presentations.

<table>
<thead>
<tr>
<th>Target Position</th>
<th>Lineup Procedure</th>
<th>Correct Rejections</th>
<th>False Positives (foil)</th>
<th>False Positives (replacement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early (n= 138)</td>
<td>Simultaneous</td>
<td>2.18 (0.91)</td>
<td>0.61 (0.80)</td>
<td>0.20 (0.41)</td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>2.28 (0.87)</td>
<td>0.41 (0.60)</td>
<td>0.08 (0.28)</td>
</tr>
<tr>
<td>Late (n= 138)</td>
<td>Simultaneous</td>
<td>1.92 (0.95)</td>
<td>0.92 (0.84)</td>
<td>0.16 (0.37)</td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>1.90 (0.97)</td>
<td>0.96 (0.9)</td>
<td>0.15 (0.40)</td>
</tr>
</tbody>
</table>

5.2.1.1 Lineup procedure

There was a main effect of lineup procedure on the occurrence of participants’ hit responses [F(1, 136)= 4.04, p= .046]. There were more hits in simultaneous than sequential lineups [M(SD)= 3.19 (1.53) vs 2.72 (1.22)]. This effect was confirmed by a nonparametric independent samples Mann-Whitney U Test [U= 1850.00, z= -2.27, p= .023, r= .194]. There was a main effect of lineup procedure on false positive responses when the target was present [F(1, 136)= 6.76, p= .010]. There were more false positive responses in sequential compared with simultaneous lineups. This was confirmed by a nonparametric Mann-Whitney U test [U= 2954.00, z= 2.67, p= .004, r= .227; M(SD)= 1.22 (1.231) vs 0.70 (1.064)]. There was an interaction between photograph condition and lineup procedure [F(2, 272)= 3.97, p= .020]. There were more false positives in the MS compared with the LS photograph condition when participants viewed simultaneous lineups [HS vs MS, T= 45.50, z= 0.58, p= .564; MS vs LS, T= 187.00, z= 2.20, p= .014, r= .034; early M(SD) HS= 0.08 (0.27), MS= 0.03 (0.18), LS= 0.11 (0.31); late M(SD) HS= 0.08 (0.27), MS= 0.16 (0.37), LS= 0.25 (0.43)], but not when they viewed sequential
lineups [HS vs MS, T= 96.00, z= -1.66, p=.097; MS vs LS, T= 264.00, z= 0.35, p=.730; early M(SD) HS= 0.15 (0.36), MS= 0.04 (0.20), LS= 0.16 (0.37); late M(SD) HS= 0.34 (0.48), MS= 0.31 (0.47), LS= 0.22 (0.41)]. There were also significantly more false positive responses in sequential compared with simultaneous lineups in the HS [U= 3034.00, z= 3.59, p< .001, r= .306; M (SD)= 0.49 (0.625) vs 0.16 (0.407)] and MS [U= 2754.00, z= 2.16, p= .015, r= .184; M (SD)= 0.35 (0.508) vs 0.19 (0.432)] photograph conditions, but not in the LS photograph condition [U= 2336.00, z= 0.17, p= .443; sequential M (SD)= 0.38 (0.635) vs 0.36 (0.545)]. Thus, the interaction between photograph condition and lineup procedure partially drove the main effect of lineup procedure.

Additionally, there was an interaction between target position, lineup procedure and photograph condition impacting on the number of false positive responses [F(1.88, 256.20)= 3.45, p= .036]. For simultaneous lineups when the target was presented early, there was no significant difference between HS, MS and LS conditions [HS vs MS, T= 3.00, z= -1.34, p= .090; MS vs LS, t(64)= 1.93, p= .058; M(SD) HS= 0.08 (0.27), MS= 0.03 (0.18), LS= 0.11 (0.31)], however when the target was presented late, there was a significant increase in false positives between HS and LS, but not between the other similarity points [HS vs MS, T= 35.00, z= 1.67, p= .096; MS vs LS, T= 114.00, z= 1.41, p= .078; HS vs LS, T= 104.00, z= 2.84, p=.003, r= .351; M(SD) HS= 0.08 (0.27), MS= 0.16 (0.37), LS= 0.25 (0.44)]. When participants viewed sequential lineups and the target person was presented early, there were significantly more false positives in the HS and LS than the MS condition, [HS vs MS, T= 13.00, z= -2.31, p= .011, r= .268; MS vs LS, t(74)= 2.59, p= .012; M(SD) HS= 0.15 (0.36), MS= 0.04 (0.20), LS= 0.16 (0.37)] but there were no significant differences when the target was presented late [M(SD) HS= 0.34 (0.48), MS= 0.31 (0.47), LS= 0.22 (0.41)]. There were also more false positive responses in sequential compared with simultaneous lineups in the HS [U= 2983.00, z= 3.68, p< .001, r= .313; M(SD)= 0.34 (0.48) vs 0.08 (0.27)] and MS [U= 2734.00, z= 2.115, p= .017, r= .180; M(SD)= MS= 0.31 (0.47) vs MS= 0.16 (0.37)] photograph conditions when the target was presented late in the lineup, but not in the LS photograph condition [M(SD)= MS= 0.22 (0.41) vs MS= 0.25 (0.44)] or when the target was presented early in the lineup [M(SD) sequential vs simultaneous HS= 0.15 (0.36) vs 0.08 (0.27); MS= 0.04 (0.20) vs 0.03 (0.18); LS= 0.16 (0.37) vs 0.11 (0.31)]. No other interactions or main effects were significant.
The current results are consistent with findings from Experiments 1 and 2, demonstrating that participants responded with more correct identifications when they viewed simultaneous lineups compared to when they viewed sequential lineups. Also consistent with the findings of Experiment 1, there were more false identifications of a lineup member when the target was present in sequential than simultaneous lineups when targets were presented late compared to early in the lineup and where lineup photographs were highly or moderately similar to originally observed photographs of the target person. This finding lends further support to the possibility raised by Experiments 1 and 2 that simultaneous lineups may lead to increased accuracy compared to sequential lineups when the target is presented late in the lineup, particularly when the the photograph of the target person originally viewed is highly similar to that presented in the lineup. These results are inconsistent with other reported findings [for example, Lindsay and Wells (1985), Clarke (2005), McQuiston-Surrett and colleagues (2006), Steblay and colleagues (2001), Steblay, Dysart & Wells (2011) and Banks (1970)]. On the basis of current findings, participants may be comparing across sequential lineups to find the lineup member who best fits their mental representation of the target person. Thus, target position and photograph similarity have marked and consistent effects on accuracy across sequential lineups.

Inconsistent with Experiments 1 and 2, results also showed a three-way interaction between lineup procedure, visual similarity of originally observed target photograph to target photograph observed in the lineup, and target position. This interaction indicated that for participants who viewed simultaneous lineups, there were no differences in false identifications when the target was presented early, but an increase in false identifications with decreasing visual similarity between target photographs when the target was presented late. For participants completing sequential lineups, however, when the target was presented early, there were more false identifications when target photographs were highly similar and least similar compared to the originally observed target photograph, but there were no differences when the target was presented late. The reasons for this pattern of results are unclear, particularly given the lack of previous literature to suggest that target position should affect accuracy in simultaneous lineups (Clark, 2005; Clark & Davey, 2005; Clark et al., 2008). There were relatively low numbers of false identifications in target present lineups, particularly when the target was presented
early. Furthermore, the current experiment may not have been sufficiently powerful to reliably detect three-way interactions. Thus, it is likely that the three-way interaction between lineup procedure, visual similarity, and target position is a type one error and the interaction is erroneous.

5.2.1.2 Visual similarity of photographs and target position

There was a main effect of photograph on the number of hit responses \(F(2,272)= 58.47\), \(p<.001\). As photograph similarity increased, the number of hits also increased \([M(SD)\ HS= 1.30 (0.71); MS= 1.11 (0.73); LS= 0.52 (0.62)]\). Additionally, there was an interaction between target position and photograph condition \([F(1.94, 263.45)= 8.97, p< .001\] which drove the main effect of photograph condition. Similarity of the photograph only had an impact when the target person was presented late in the lineup \([HS vs MS, T= 354.00, z= -4.46, p< .001, r= .016; MS vs LS, t(137)= 2.97, p= .002; M(SD) HS= 0.70 (0.46); MS= 0.45 (0.50); LS= 0.30 (0.46)]\). When the target was presented early, the rate of hits in the HS and MS conditions were not significantly different, however there were significantly more hits in both of these conditions than in the LS condition \([HS vs MS, T= 648.00, z= 1.02, p=.307; MS vs LS, T= 451.00, z= -6.56, p< .001, r=.56; HS vs LS, T= 680.00, z= -5.67, p< .001, r=.483; M(SD) HS= 0.60 (0.49), MS= 0.65 (0.48), LS= 0.22 (0.42)]\).

There was a significant main effect of photograph condition on the number of participants’ miss responses \([F(2, 272)= 58.15, p< .001]\). There were less misses as similarity between the originally observed target photograph and target photograph presented in the lineup increased \([early M(SD) HS= 0.28 (0.45), MS= 0.31 (0.46), LS= 0.64 (0.48); late M(SD) HS= 0.09 (0.28); MS= 0.31 (0.46); LS= 0.48 (0.50)]\). There was a significant main effect of target position \([F(1, 136)= 21.06, p< .001]\). There were more misses when the target person was presented early than when he was presented late in the lineup \([M(SD)= 1.23 (0.890) vs 0.86 (0.892); Table 5.1]\). There was also an interaction between target position and photograph condition \([F(1.91, 259.36)= 4.58, p= .012\] which drove the main effect of photograph condition. Photograph similarity only had an impact when the target person was presented late in the lineup \([HS vs MS, T= 594.00, z= 5.24, p< .001, r=.020; MS vs LS, T= 962.00, z= 3.22, p=.001, r=.012; M(SD) HS= 0.09 (0.28); MS= 0.31 (0.46); LS= 0.48 (0.50)]\). When the target was presented early, the rate of misses in the HS
and MS conditions were not significantly different, however there were significantly less misses in both of these conditions than in the LS condition [HS vs MS, t(137)= 0.62, p= .270; MS vs LS, t(137)= 5.87, p< .001; HS vs LS, t(137)= 6.324, p< .001; M(SD) HS= 0.28 (0.45), MS= 0.31 (0.46), LS= 0.64 (0.48)]. This interaction also partially drove the main effect of target position. There were more misses when the target was presented early than when they were presented late in the HS [T= 74.00, z= -4.67, p< .001, r= .281; M(SD) early= 0.29 (0.45), late= 0.09 (0.28)] and LS [t(137)= 3.14, p= .001; M(SD) early= 0.64 (0.48), late= 0.48 (0.50)] photograph conditions, but not in the MS [T= 552.00, z= -0.15, p= .442; M(SD) early= 0.31 (0.46), late= 0.31 (0.46)] photograph condition.

There was a main effect of target position [F(1, 136)= 28.62, p< .001]. There were more false positive responses in target present lineups when the target was presented late than when they were presented early [M(SD)= 0.69 (0.870) vs 0.29 (0.556); Table 5.1]. The aforementioned interaction between target position, lineup procedure and photograph condition partially drove this main effect as there were more false positive responses for late compared with early target presentations when simultaneous lineups were viewed for MS [U= 49.50, z= 2.53, p= .006, r= .316] and LS [U= 140.00, z= 2.07, p=.020, r=.258] photograph conditions and when sequential lineups were viewed for HS [U= 34.42, z= 2.75, p=.003, r=.319] and MS [U= 275.00, z= 4.08, p< .001, r=.475] photograph conditions, but not in sequential lineups in the LS condition [U= 85.00, z= 1.00, p=.159] or simultaneous lineups in the HS condition [U= 18.00, z= 0.000, p>.99]. No other interactions or main effects were significant.

Visual similarity between the originally observed target photograph and target photograph viewed in the lineup had a significant impact on accuracy and this was consistent with findings observed in the same or different photograph condition in Experiments 1 and 2. With higher visual similarity between the originally observed target photograph and target photograph viewed in the lineup in the current experiment, there were more correct identifications and less incorrect rejections. These findings are consistent with the wider literature which demonstrates that visual similarity in features between the original observation and subsequent observation of a target in a lineup has a marked impact on accuracy (Thomson et al., 1982).

The current results extend findings from Experiments 1 and 2, suggesting that varying levels of visual similarity between the original observation and lineup
observation of the target can have a marked impact on accuracy. While Experiments 1 and 2 have suggested that observing the same photograph in the lineup as that seen originally can enhance accuracy, the current investigation suggests that this effect also occurs for varying levels of visual similarity, none of which include the same photograph as that originally seen. This finding increases the generalisability of findings relating to visual similarity between observations of the target. Given that in a real life scenario, the target person is highly unlikely to look exactly the same in the lineup as upon original observation, the current findings suggest that visual similarity between these observations can still enhance accuracy. These data are consistent with previous literature which has highlighted the impact of visual similarity in features between original observation and subsequent observation of a target in a lineup on accuracy (Thomson et al., 1982).

Further, there were some instances in the current experiment where the difference in accuracy between the highly similar and moderately similar photograph condition was smaller than the difference between these two categories and the least similar photograph condition. This was the case for correct identifications and incorrect rejections of the lineup when the target was presented early in the lineup. Further, the finding of more false identifications in sequential than simultaneous lineups was present in both the same photograph condition in Experiment 1 and both the highly and moderately similar photograph conditions in the current experiment. These results add to previous literature on visual similarity in eyewitness identification indicating that there may be a threshold effect, whereby a certain threshold of visual similarity, or familiarity, must be reached for people to accurately recognise a previously seen face. In other words, rather than visual similarity affecting accuracy in a linear fashion, there may be some threshold for visual similarity between information in the the retrieval cue and the memory cue required for successful recognition of a target (Atkinson & Juola, 1974; Atkinson & Shiffrin, 1971). In the current experiment, it would seem that the highly similar and moderately similar conditions were more likely to reach this threshold than the least similar photograph condition. Thus, individuals may be able to withstand some changes in appearance and still remain accurate on facial identification tasks, however once a certain threshold is reached, in terms of differences between observations, people are much less likely to be accurate. This explanation would be consistent with signal detection theory which suggests that targets in the least similar
photograph condition were unlikely to reach participants’ familiarity threshold whereas those in the other two conditions were more likely to reach this threshold for choosing (Banks, 1970; Meissner et al., 2005).

The precise level of the threshold of visual similarity required for successful recognition, however, is not clear from the current experiment, although it would appear that while recognition was robust to some differences in the highly and moderately similar photograph conditions, the least similar photograph condition certainly contained differences sufficient to disrupt recognition accuracy. This threshold effect may also interact with other variables that impact upon accuracy, such that performance may be more impacted by differences between observations in the presence of other variables that reduce accuracy. This possibility would be consistent with current findings showing that the difference between the highly similar and moderately similar photograph conditions was negligible when targets were presented early in the lineup, but significant when targets were presented late in lineups. It must be noted, however, that the visual similarity manipulation was based on subjective ratings of an independent group of participants that may have differed somewhat from the subjective judgment of those participating in the study. As such, it may not be surprising that there was some overlap in patterns of accuracy between the highly and moderately similar photograph conditions. (See Chapter 6 for further discussion.)

The current experiment also found that when the target was presented early in the lineup, the impact of visual similarity on accuracy was reduced such that there was sometimes no significant difference in accuracy between highly similar and the moderately similar photograph conditions. This was the case for accuracy both in terms of correct identification and incorrect rejections of the lineup. In other words, the change in visual similarity between the highly similar and moderately similar photograph conditions only impacted on accuracy, in terms of correct identifications and incorrect rejections of the lineup, when targets were presented late in the lineup. The reasons behind this pattern of results is unclear, particularly given that it was seen in both sequential and simultaneous lineup procedures. Although it is possible that some change in memory or in criterion thresholds for choosing may have occurred across sequential lineups that may have impacted upon accuracy (Altmann & Gray, 2002; Gold, Murray, Sekuler, Bennett, & Sekuler, 2005; Jenkins, Earle-Richardson, Slingerland, & May, 2002; Pavlik & Anderson, 2005) it is unclear why
similar changes should occur across simultaneous lineups where all lineup members are presented at once. Indeed, most researchers suggest that the presentation order of target should not impact accuracy in simultaneous lineup procedures (Clark & Davey, 2005; Clark et al., 2008). The current results, on the other hand, indicate that it is possible that late presentation of targets may have a negative effect on accuracy in both sequential and simultaneous lineups.

5.2.1.3 Target replacement position

For correct rejection responses, no violations of the parametric assumptions were observed. There was a main effect of target replacement position \( [F(1, 136) = 15.51, p < .001] \). There were more correct rejection responses when the target replacement was presented early than when it was presented late \( [M(SD) = 2.25 (0.88) \text{ vs } 1.91 (0.97)] \). There was also a main effect of target replacement position on false positive responses where a foil was selected \( [F(1, 136) = 38.928, p < .001] \). There were more false positive responses when the target replacement was presented late than when it was presented early \( [M(SD)= 0.94 (0.87)] \text{ vs } 0.50 (0.67)] \). This was confirmed by a nonparametric related samples Wilcoxon signed-rank test \( [T= 2132.00, z= 5.59, p< .001, r= .337] \). No other interactions or main effects were significant.

Results demonstrate that there were more correct rejections of the lineup and less false identifications in target absent lineups when the target replacement was presented early than when it was presented late. This finding has not been observed in previous experiments in the current investigation but has been observed in previous literature only for sequential lineups (Clark & Davey, 2005). The current results are not consistent with these previous findings as an order effect was seen for both sequential and simultaneous lineups. The reasons for this discrepancy are difficult to understand particularly given the majority of literature suggesting that target replacement position should not affect accuracy for judgments in simultaneous lineups (Clark & Davey, 2005). There were numerically more participants who completed sequential than simultaneous lineups (74 compared to 64) in the current experiment. As such, it is possible that the effect for sequential lineups may be partially driving the main effects despite no interaction with lineup effect being observed. However, while the difference in the number of correct rejections of the lineup and false identification of a lineup member between early and late
presentations was slightly less for participants who viewed simultaneous lineups, a consistent difference in simultaneous lineups was still present. As such, it cannot be concluded that the significant main effects are due to a type two error to detect the two-way interaction. It is possible that this finding indicates that there was a portion of participants who completed the simultaneous lineups as if it were a sequential lineup (or going through each face sequentially to find a match), which may explain an effect of target position in both lineup procedures. Nevertheless, despite the mechanisms for the effect remaining unclear, current data indicate that target position may have an impact on accuracy for simultaneous as well as sequential lineups. This suggests that lineup administrators should be conscious of suspect positioning for simultaneous as well as sequential lineups.

5.2.2 Confidence
Table 5.3 displays sample sizes for hit, miss, false positive and correct rejection confidence ratings. Table 5.4 displays participant mean confidence ratings for hit, miss and false positive responses in lineups where the target was present. Given very small sample sizes in some conditions of the design for miss and false positive responses when the target was present it was not appropriate to run the planned ANOVA including both target position and photograph condition as factors (Table 5.5). As such, these factors were included separately and the analysis was conducted first collapsing across target position and with photograph condition as a factor, and then collapsing across photograph condition including target position as a factor.

When target position was collapsed and photograph condition was included for, sample sizes were sufficient to conduct the ANOVA with the condition with the smallest sample size being 24 for miss responses and 10 for false positive responses. When photograph condition was collapsed and target position was included for miss responses, sample sizes were sufficient to conduct the ANOVA, with the condition with the smallest sample size being 59 for miss responses and 31 for false positive responses.
Table 5.3
Sample sizes for hit, miss, correct rejection and false positive confidence ratings for target absent and present sequential and simultaneous lineups in highly similar (HS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late photograph presentations

<table>
<thead>
<tr>
<th>Lineup Procedure</th>
<th>Target Position</th>
<th>Target Presence</th>
<th>Photograph Condition</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives (Replacement Selected)</th>
<th>Correct Rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simultaneous</td>
<td>Early (n=384)</td>
<td>Present (n=192)</td>
<td>HS (n=64)</td>
<td>42</td>
<td>17</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MS (n=64)</td>
<td>42</td>
<td>20</td>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS (n=64)</td>
<td>18</td>
<td>39</td>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absent (n=192)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>52 (13)</td>
<td>140</td>
</tr>
<tr>
<td>Late (n=384)</td>
<td>Present (n=192)</td>
<td>HS (n=64)</td>
<td>50</td>
<td>9</td>
<td>5</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MS (n=64)</td>
<td>31</td>
<td>23</td>
<td>10</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LS (n=64)</td>
<td>21</td>
<td>27</td>
<td>16</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absent (n=192)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>69 (10)</td>
<td>123</td>
</tr>
<tr>
<td>Sequential</td>
<td>Early (n=444)</td>
<td>Present (n=222)</td>
<td>HS (n=74)</td>
<td>41</td>
<td>22</td>
<td>11</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MS (n=74)</td>
<td>48</td>
<td>23</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LS (n=74)</td>
<td>13</td>
<td>49</td>
<td>12</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absent (n=222)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>53 (11)</td>
<td>169</td>
</tr>
<tr>
<td>Late (n=444)</td>
<td>Present (n=222)</td>
<td>HS (n=74)</td>
<td>47</td>
<td>2</td>
<td>25</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MS (n=74)</td>
<td>32</td>
<td>19</td>
<td>23</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LS (n=74)</td>
<td>20</td>
<td>38</td>
<td>16</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absent (n=222)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>82 (11)</td>
<td>140</td>
</tr>
</tbody>
</table>
Table 5.4
Mean (SD) confidence levels for hit, miss and false positive responses when the target was present in sequential and simultaneous lineups in highly similar (HS), moderately similar (MS) and dissimilar (LS) photograph conditions for early and late photograph presentations.

<table>
<thead>
<tr>
<th>Target Position</th>
<th>Photograph Condition</th>
<th>Lineup Procedure</th>
<th>Hits</th>
<th>Misses</th>
<th>False Positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>HS</td>
<td>Simultaneous</td>
<td>3.57 (1.02)</td>
<td>3.00 (0.79)</td>
<td>2.00 (1.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.36 (1.16)</td>
<td>2.82 (1.33)</td>
<td>1.73 (0.65)</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>Simultaneous</td>
<td>3.12 (1.02)</td>
<td>3.25 (0.97)</td>
<td>2.50 (0.71)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.23 (1.08)</td>
<td>2.48 (1.24)</td>
<td>2.33 (0.58)</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>Simultaneous</td>
<td>2.78 (1.06)</td>
<td>3.00 (1.05)</td>
<td>2.29 (0.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>2.31 (0.86)</td>
<td>2.86 (1.34)</td>
<td>2.08 (1.17)</td>
</tr>
<tr>
<td>Late</td>
<td>HS</td>
<td>Simultaneous</td>
<td>3.78 (1.20)</td>
<td>2.67 (1.12)</td>
<td>2.60 (0.89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.81 (1.01)</td>
<td>3.50 (0.71)</td>
<td>2.08 (0.86)</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>Simultaneous</td>
<td>2.74 (1.32)</td>
<td>3.04 (1.33)</td>
<td>1.60 (0.67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>3.44 (0.95)</td>
<td>2.63 (1.26)</td>
<td>2.65 (0.98)</td>
</tr>
<tr>
<td></td>
<td>LS</td>
<td>Simultaneous</td>
<td>2.38 (1.02)</td>
<td>2.67 (1.00)</td>
<td>1.63 (0.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sequential</td>
<td>2.500 (1.05)</td>
<td>2.47 (1.25)</td>
<td>2.43 (0.89)</td>
</tr>
</tbody>
</table>

Table 5.5 displays participant mean confidence ratings for correct rejection and false positive responses when the target replacement was selected and when a different lineup foil was selected in lineups where the target was absent.

5.2.2.1 Lineup procedure

When photograph condition was collapsed and target position was included for miss responses, there was a main effect of lineup procedure \[F(1,287)= 3.97, p= .047\]. There was slightly higher confidence in misses when participants viewed simultaneous \[M(SD= 2.96 (1.06))\] than when they viewed sequential lineups \[M(SD)= 2.68 (1.28)] overall.
Table 5.5
Mean (SD) confidence levels in for correct rejection and false positive responses where either the target replacement or a different lineup foil was chosen and when the target was absent in sequential and simultaneous lineups for early and late photograph presentations.

<table>
<thead>
<tr>
<th>Target Position</th>
<th>Lineup Procedure</th>
<th>Correct Rejections (Replacement)</th>
<th>False Positives (Replacement)</th>
<th>False Positives (Foil)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>Simultaneous</td>
<td>3.17 (1.24)</td>
<td>1.54 (0.66)</td>
<td>2.10 (1.10)</td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>3.26 (1.27)</td>
<td>2.64 (0.67)</td>
<td>2.48 (0.92)</td>
</tr>
<tr>
<td>Late</td>
<td>Simultaneous</td>
<td>2.88 (1.23)</td>
<td>2.00 (0.82)</td>
<td>2.20 (0.96)</td>
</tr>
<tr>
<td></td>
<td>Sequential</td>
<td>3.15 (1.27)</td>
<td>2.09 (0.94)</td>
<td>2.25 (0.86)</td>
</tr>
</tbody>
</table>

There was also an interaction between lineup procedure and target position impacting on confidence in hit responses [F(1,393)=3.96, p= .047]. However, post hoc t-tests found no significant differences. When target position was collapsed and photograph condition was included for false positive responses, there was an interaction between lineup procedure and photograph similarity [F(1, 134)= 3.66, p= .029]. The HS photograph condition [M(SD)= 2.30 (0.95)] made little difference to confidence in false positive responses compared to the MS [M(SD)= 1.75 (0.75); t(20)= 1.52, p=.145] and LS [M(SD)= 1.826 (0.89); t(31)= 1.38, p=.177] conditions when participants viewed simultaneous lineups, however the HS condition decreased confidence in comparison to the MS condition when sequential lineups were viewed [t(60)= 2.88, p=.005; M(SD)= 1.97 (0.81) vs 2.62 (0.94)]. There was no significant difference in confidence between the HS and LS photograph conditions [t(62)= 1.38, p=.174; M(SD) HS= 1.97 (0.81), LS= 2.29 (1.013)]. When photograph condition was collapsed and target position was included for false positive responses, there was an interaction between target position and lineup procedure [F(1,134)= 5.62, p= .019]. However, upon post hoc testing no significant differences were revealed. For
confidence in false positive responses in target absent lineups where replacements were selected, there was a main effect of lineup procedure \[F(1,41)= 6.54, p= .014\]. There was higher confidence in false positive responses when the target replacement was selected in sequential compared with simultaneous lineups [M(SD)= 2.36 (0.84) vs 2.04 (0.85)]. There was also an interaction between target replacement position and lineup procedure which drove the main effect of lineup procedure \[F(1,41)= 4.70, p= .036\]. Confidence in choosing target replacements was higher in sequential compared with simultaneous lineups when the target replacement was presented early \[t(22)= 4.02, p= .001; M(SD)= 2.64 (0.67) vs 1.54 (0.66)\] but not when it was presented late \[t(19)= 0.235, p= .408; sequential vs simultaneous M(SD)= 2.09 (0.94) vs 2.00 (0.81)\]. No other interaction or main effects were significant.

Results indicated that participant confidence in false identifications was higher in sequential than simultaneous lineups when the target was absent from the lineup and the target replacement was presented early in the lineup. This was despite there being no difference between lineup procedures in the number of false identifications of the lineup when the target was absent. Results further indicated that participant confidence in their decisions when they selected the target replacement was higher when they viewed sequential than simultaneous lineups. This effect was only seen when the target replacement was presented early, but not when it was presented late. These findings demonstrate that participant confidence was elevated when viewing the target replacement early in sequential lineups compared with simultaneous lineups when the target was absent irrespective of the accuracy of those responses.

The findings relating to confidence are only partially consistent with those from Experiments 1 and 2. Experiment 2 demonstrated that participant confidence in false identifications was higher in sequential lineups when the target replacement was falsely identified from the lineup. Experiment 1 found that participant confidence in false identifications was higher when completing sequential lineups than simultaneous lineups in general. Experiment 3 demonstrated the same effect as in Experiment 2, but only when the target replacement was presented early in sequential lineups. Previous literature has seldom found a difference between sequential and simultaneous lineup procedures for confidence in lineup decisions. Where differences have been found, findings have indicated that confidence of sequential and simultaneous lineups is equivocal (Lindsay & Wells, 1985) and when
it does differ, that confidence in sequential lineups is inflated for nonchoosers (i.e.
mistaken responses), but not for choosers (i.e. hit and false positive responses; Sporer,
1993). However, the current experiment has demonstrated inflated confidence in
sequential lineups compared to simultaneous lineups for false identifications when
the target was absent and when the target replacement was identified early in the
lineup.

The reasons why the current findings, with respect to confidence, are
inconsistent with previous literature and with Experiments 1 and 2 are unclear.
There appears to be no single explanation that can account for the pattern of accuracy
and confidence obtained in the current experiment. Further, there appears no
plausible explanation in relation to differences in protocol as to why the current
experiment would not have observed effects typically seen in the literature, such as
the sequential superiority effect. As such, more detailed discussion of the factors that
may have contributed to differences in accuracy and confidence between lineup
procedures across experiments will take place in later sections. (See General
Discussion, Chapter 6.)

While the mechanisms behind the present results remain unclear, current
findings demonstrate that sequential lineups may lead to higher confidence
irrespective of accuracy of responses, particularly confidence in false identifications
of lineup members. This elevation in confidence occurred in the current experiment
in situations in which a target replacement was falsely identified early in the lineup.
Thus, confidence may be more likely to be elevated in sequential when compared to
simultaneous lineups where a witness falsely identifies an alleged suspect early in a
lineup. The most dire miscarriages of justice occur when the perpetrator is absent
from the lineup and where an alleged suspect is falsely identified. Furthermore,
eyewitness confidence is heavily relied upon by jurors when making decisions about
guilt or innocence (Fox & Walters, 1986; Lindsay et al., 1989; Whitley & Greenberg,
1986). Thus, on the basis of the current results, in situations where eyewitness
confidence is relied upon heavily, conducting simultaneous lineups may be
preferable to conducting sequential lineups when considering the prevention of
wrongful convictions.
5.2.2.2 **Visual similarity and target/target replacement position**

For hit responses there was a main effect of photograph condition on confidence [F(2,393)=29.06, p< .001]. Confidence increased with increasing photograph similarity [M(SD) HS= 3.64 (1.11); MS= 3.14 (1.10); LS= 2.50 (1.01)]. Post hoc t-tests indicated that confidence in hits was higher in the HS than the MS condition [t(331)= 4.12, p< .001], and in the MS than the LS condition [t(223)= 4.20, p< .001].

Results for confidence were in line with findings relating to accuracy, suggesting that increasing visual similarity between the originally observed target photograph and target photograph viewed in the lineup increases correct identifications as well as confidence in correct identifications. This finding is consistent with previous experiments and suggests that visual similarity is related to the way identification decisions are made. The current results also extend the findings from Experiments 1 and 2 suggesting that greater similarity between photographs, none of which include the same photograph of the target person, increases the relationship between confidence and accuracy. Thus, on the basis of current data, visual similarity may represent a tool by which people make identifications decisions. The marked impact of visual similarity on accuracy and confidence across experiments suggests that eyewitnesses may utilise similarity between the previously seen target and lineup members as a basis for making identification decisions. In other words, identification decisions may be highly based in a judgment of similarity, or comparison of lineup members across simultaneous and sequential lineups in order to find the ‘best fit’, or best match with a person’s memory of the target.

5.2.2.3 **Confidence in accurate versus inaccurate responses**

Levene’s equality of error variances test was violated for analyses comparing accurate with inaccurate responses. There was a main effect of response type for the analysis comparing confidence in hits with that in false positive responses [F(1, 788)= 200.34, p< .001]. Reported confidence for hits (range= 1-5) was higher compared to false positive responses (range= 1-5) overall [M(SD)= 3.25 (1.16) vs 2.19 (0.93)]. This effect was confirmed by a within-subjects nonparametric Wilcoxon signed-rank test [T= 26.50, z= -6.73, p< .001, r= .560]. When only false positives where the target replacement was selected were included in the analysis,
Levene’s equality of error variances test was violated. There was still a main effect of response type \([F(1, 442)= 44.124, p< .001]\). There was higher confidence in hit compared to false positive responses when the target replacement was selected \([M(SD)= 3.25 (1.16) vs 2.04 (0.85)]\). This result was confirmed by a nonparametric Mann-Whitney U test \([U= 3928.00, z= -6.45, p< .001, r= .304]\).

There was also a main effect of response type for the analysis comparing confidence in miss with that in correct rejection responses \([F(1, 852)= 12.312, p< .001]\). Confidence in correct rejection responses was higher than that in miss responses \([M(SD)= 3.13 (1.26) vs 2.81 (1.19)]\). This result was confirmed by a nonparametric Mann-Whitney U test \([U= 70134.00, z= -3.65, p< .001, r= .125]\). There was also an interaction between lineup procedure and response type that drove the main effect of response type \([F(1, 852)= 6.58, p= .010]\). There was higher confidence in correct rejection than miss responses when participants completed sequential lineups \([U= 18099.50 z= -4.21, p< .001, r= .196; M(SD)= 3.21 (1.27) vs 2.68 (1.28)]\) but not when they completed simultaneous lineups \([U= 17149.50, z= -0.57, p= .284; M(SD) correct rejections= 3.03 (1.24) vs 2.96 (1.06)]\). (See Table 5.6) There was also a main effect of target position \([F(1, 852)= 5.577, p= .018]\). Confidence was higher when the target or target replacement was presented early than when it was presented late overall \([M(SD)= 3.10 (1.24) vs 2.92 (1.25)]\).

There were also significant main effects reflective of findings reported in the individual response type analyses. There was a main effect of lineup procedure \([F(1, 788)= 5.68, p= .017]\). There was higher confidence in sequential compared with simultaneous lineups however, this result was not consistent with the nonparametric Mann-Whitney U test \([U= 82100.50, z= 1.05, p= .148]\). No other significant main effects or interactions were observed.
### Table 5.6

Mean (SD) confidence levels for miss and correct rejection responses in simultaneous and sequential lineups where the target replacement was presented early or late in the lineup.

<table>
<thead>
<tr>
<th>Target Position</th>
<th>Lineup procedure</th>
<th>n misses</th>
<th>Misses (SD)</th>
<th>n CRs</th>
<th>Correct Rejections (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simultaneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early</td>
<td>76</td>
<td>3.07 (0.97)</td>
<td></td>
<td>140</td>
<td>3.17 (1.23)</td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>2.76 (1.31)</td>
<td></td>
<td>169</td>
<td>3.26 (1.27)</td>
</tr>
<tr>
<td>Late</td>
<td>59</td>
<td>2.81 (1.15)</td>
<td></td>
<td>123</td>
<td>2.88 (1.23)</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>2.56 (1.24)</td>
<td></td>
<td>140</td>
<td>3.15 (1.27)</td>
</tr>
</tbody>
</table>

The finding that participants’ confidence of correct identifications of the target person was higher than that of false identifications of a lineup member is consistent with Experiments 1 and 2, as well as some findings in the wider literature and the propositions of the USA Supreme Court relating to confidence (Neil v Biggers, 1972; Wise & Safer, 2004). However, these data appear inconsistent with the suggestion from the courts in the UK that confidence is not a reliable indicator of accuracy (R v Turnbull, 1976; Devlin, 1976). As discussed in preceding sections, this may be due to the use of the comparison of group means, rather than a comparison of confidence of certain individuals, which may have importantly different implications for the findings. However, it must be noted that in the current investigation a more even spread of scores across the confidence range was observed for hit responses in comparison to Experiments 1 and 2.

Consistent with Experiment 1 and 2, findings within the separate response types indicated that the relationship between confidence and accuracy may be more relevant under some conditions than others. Particularly, differences in accuracy between sequential and simultaneous lineups did not align with differences in confidence. In particular, sequential lineups led to decreased accuracy but higher levels of confidence in comparison to simultaneous lineups when the target was absent. This was irrespective of the accuracy of those decisions. On the other hand,
confidence in correct responses for simultaneous lineups did not differ from that of sequential lineups. As such, the current results indicate that although confidence may generally be related to accuracy for eyewitness identifications, accuracy when people view simultaneous lineups may align more closely with their levels of confidence, particularly when a lineup member or suspect is falsely identified as the target person who is absent from the lineup. On the basis of current findings, the use of sequential lineups in cases where the court relies on confidence as an indicator of accuracy may be more risky than the use of simultaneous lineups. This possibility is particularly salient when considering the occurrence of higher confidence for false identifications of a target replacement in sequential lineups in the current study.

Finally, the finding that there was higher confidence in correct rejections than incorrect rejections of the lineup in only sequential lineups is inconsistent with Experiment 2 as well as findings from past research. Generally, the association between confidence and accuracy has been found to be lower for non-choosers than for choosers (Brigham, 1988; Sporer, 1993; Sporer et al., 1995). In other words, that for people who make an identification decision, confidence may be a more reliable predictor of accuracy than for those who do not. However, current data indicate that confidence for correct rejections of the lineup were significantly higher than that for incorrect rejections of the lineup, but only for participants who completed sequential lineups. Differences in the relationship between confidence and accuracy between the different lineup procedures have rarely been investigated or reported in the literature. Although Sporer (1993) found a significant relationship between confidence and accuracy for non-choosers (albeit weaker than that for choosers), and that confidence in decisions for non-choosers was higher for sequential lineups than simultaneous lineups, they reported no interaction between the confidence-accuracy relationship and lineup procedure. As such, it is difficult to explain the current finding that participants’ confidence in correct rejections of the lineup were higher than that in incorrect rejections of the lineup but only in participants who completed sequential lineups. It is possible that these findings may relate to changes in memory or in participants’ criterion threshold for identifying a lineup member. These changes may have been contributed to or been impacted upon by differences between lineup procedures in general. For example, it is possible that there were some changes in memory or in criterion threshold that occurred across sequential lineups,
but not across simultaneous lineups. However, how these changes may have occurred to produce the current results in this experiment and not other experiments in this investigation is unclear. As such, further study is needed to clarify the reasons for inconsistent findings throughout the literature and among experiments in the present investigation relating to differences in confidence for incorrect compared to correct rejections.

5.3 Summary

The major findings central to the hypotheses and aims of Experiment 3 are as follows:

- There were more correct identifications of the target from the lineup in simultaneous than in sequential lineups and less false identifications in simultaneous than sequential lineups overall due to more false positive responses in sequential than simultaneous lineups in the highly and moderately similar photograph conditions, where the target was presented late in the lineup.
- With decreasing visual similarity conditions, there were less hit and more miss responses, as well as more false positive responses in simultaneous lineups.
- There was higher confidence in hit responses with increasing visual similarity between the originally observed photograph and the photograph observed in the lineup.
- Reported confidence in false positive responses when the target was absent and the target replacement was selected early in the lineup was higher for sequential than simultaneous lineups overall.
- Confidence in correct identifications was higher than that of incorrect identifications.
- Confidence in correct rejections of the lineup was higher than that of incorrect rejections of the lineup.

Many of the findings from Experiment 3 are consistent with those of Experiments 1 and 2. Consistent with the first two experiments, confidence in correct identifications was higher than that of incorrect identifications. Confidence of correct rejections of the lineup, however, was also higher than that of incorrect
rejections of the lineup, which has not been previously observed. Further, findings of no evidence of a sequential superiority effect was observed and results suggested that a close match between two photographs of an unfamiliar target person increases accuracy and confidence in terms of correct identifications. Further consistent with findings of Experiment 2, Experiment 3 data suggested that confidence and accuracy may better align when participants viewed simultaneous than sequential lineups and the target was absent. Particularly, confidence in false identifications when the target replacement was falsely identified early in the lineup was elevated for sequential lineups. On the basis of current findings, confidence where an innocent suspect is identified in sequential lineups may be a particularly unreliable indicator of accuracy. Given that these findings have not been observed elsewhere in the literature and false identification of an innocent suspect can lead to wrongful conviction, explanations for these patterns of findings should be explored. Theoretical explanations and mechanisms underlying the pattern of results observed across all three experiments will be discussed in Chapter 6.
CHAPTER 6  
General Discussion

The present investigation had two primary aims. First, it aimed to clarify patterns of accuracy across different visual similarity conditions for simultaneous and sequential lineup judgments. There has been some debate in the literature as to whether sequential lineups are superior in terms of eyewitness accuracy. Some studies suggest this procedure yields benefits in terms of reducing the number of false identifications without a cost to the number of correct identifications obtained, whereas others indicate that sequential lineups are only advantageous in terms of false identifications with simultaneous lineups being superior in terms of correct identifications. Secondly, the present investigation aimed to characterise how patterns of confidence in lineup judgments related to patterns of accuracy of those judgments in different photograph conditions for sequential and simultaneous lineups. While much literature has been dedicated to examining the degree of association between confidence and accuracy in general, no study to date has examined the effect of different lineup procedures and different photograph conditions on patterns of confidence in the different response types. Investigation of these effects is essential to evaluating the practicalities of the use of eyewitness confidence as an indicator of eyewitness accuracy, as is routine in USA Supreme Courts.

Drawing on what is known about eyewitness lineup judgments, it was hypothesised that lineup procedure would have an impact on accuracy and sequential lineups would lead to less false positive identifications. It was also expected that higher visual similarity between the originally observed target and the target observed in the lineup would lead to increased accuracy and higher confidence in accurate responses. Drawing on what is known about the relationship between confidence and accuracy, it was predicted that patterns of confidence in the different response types would mirror patterns of accuracy such that conditions in which more accurate responses occurred would also be related to higher confidence. Further, it was expected that there would be higher confidence in correct than false identifications, and that there would be higher confidence in correct than incorrect rejections of the lineup but to a lesser extent than the difference seen between correct and false identifications.
Contrary to expectations, no benefit of sequential lineups in terms of false identifications was observed. Across experiments, more correct identifications and less false identifications in simultaneous when compared with sequential lineups was consistently observed. These findings are inconsistent with previously reported findings within the eyewitness literature. Results supported the hypothesis that visual similarity between the originally observed target photograph and that seen in the lineup would increase accuracy and confidence in accurate responses. Across experiments, results indicated a consistent increase in correct identifications and enhancement in confidence when the same photograph as was originally observed was presented in the lineup, and where there was higher visual similarity between the originally observed and lineup target photograph. Finally, in line with predictions, the current experiments consistently indicated that confidence in correct identifications was higher than that of false identifications. However, the same pattern was not observed for correct and incorrect rejections of the lineup. Inconsistent with predictions, patterns of confidence within the separate response types did not mirror patterns of accuracy indicating that higher confidence may not be related to higher accuracy under all circumstances. Particularly, the current investigation consistently indicated that confidence in false identifications was higher in sequential than simultaneous lineups. The differences observed between the present findings and previous studies support the theoretical reasoning behind the present study that clarification of past research on eyewitness accuracy and confidence was needed. The unique findings observed in the current study, in this respect, support the investigation’s rationale, which noted inconsistencies in the literature on eyewitness accuracy, as well as unstudied connections between accuracy, confidence and lineup procedure.

6.1 Impact of lineup procedure on accuracy and confidence: Integration of findings, theoretical implications and practical application

The current investigation demonstrated consistent findings across three different experiments with respect to the impact of lineup procedure on eyewitness accuracy. Across three experiments, data indicated that there were more hits, or correct identifications, and less false positives, or false identifications, of the target in simultaneous lineups compared to sequential lineups. Although this effect was
confined to instances that the target was presented late in the lineup for the last experiment, these findings collectively provide strong evidence of a simultaneous superiority effect rather than the commonly reported sequential superiority effect in previous literature.

The current pattern of results conflicts with prominent theoretical explanations for the differences in accuracy between simultaneous and sequential lineups. More specifically, this pattern of results is inconsistent with the proposition that the sequential lineup procedure encourages absolute, rather than relative, judgment processes (Wells, 1984). Researchers have proposed in the past that when people complete simultaneous lineups, their tendency is to compare across lineup members and choose the lineup member that is most similar to the target person, or the suspect. However, when they complete sequential lineups, they are not able to engage in this direct comparison of faces, and so they must rely on their original memory of the target, or suspect, in order to make absolute judgments about whether each lineup member matches their memory of the target (Steblay et al., 2001; Wells, 1984, 1993). According to this assertion, absolute judgments occurring in sequential lineups protect against false identifications of the suspect. In other words, instead of choosing the most similar of the lineup members presented, they only choose a lineup member if that member’s appearance is a sufficient match to their mental representation of the original target. In other words, instead of being encouraged to compare each lineup member to their mental representation of the target person in their mind and choosing the most similar lineup member relative to the others, eyewitnesses would instead be encouraged to compare each member of the lineup to their mental representation of the target, or suspect, in their mind and only choose a lineup member if a sufficient match was found - if no match occurred, they would reject the lineup. Conceptually, in the former case of making a relative judgment the task is analogous to a forced choice protocol, whereby an eyewitness will make a choice, and they will do this based on a decision about which lineup member is most similar to the suspect. In the latter case of making an absolute judgment, rejection of the lineup is possible if no sufficient match with the mental image of the perpetrator is present in the lineup. This proposition also assumes that no change in the memory trace of the target person, and no shift in criterion threshold, would occur across sequential lineups in comparison to simultaneous lineups. Logically, then, it would be expected that if sequential lineups encouraged absolute over relative judgments,
and a fixed matching threshold were applied over sequential lineups, there would be more rejections of the lineup as well as less false identifications - the mental representation of the target face may not match sufficiently with any of the lineup members, eventuating in rejection of the lineup. More stringent absolute judgments would lead to a lower probability that the target face presented in the lineup would match sufficiently with an eyewitness’s original memory of the suspect. Encouraging relative judgments would, in contrast, encourage eyewitnesses to make a choice of the most similar lineup member relative to the other lineup members, regardless of the degree of match between the image of the perpetrator and the lineup member identified. This method would eventuate in less rejections of the lineup, more correct identifications and also more false identifications of lineup members.

The findings of the current investigation indicated that sequential lineup procedures led to more false identifications with no difference in incorrect rejections of the lineup. These findings are in direct conflict with the proposition that sequential lineups encourage absolute judgment processes rather than relative judgment processes. As such, the current results lend support to studies that have questioned the accuracy of the assertion that sequential lineups encourage absolute judgments whereas simultaneous lineups encourage relative judgments (Clark, 2005; Flowe & Ebbesen, 2007).

The current pattern of results is also in contradiction with proposed explanations in line with signal detection theory (Ebbesen & Flowe, 2002; McQuiston-Surrett et al., 2006; Meissner et al., 2005). In terms of signal detection theory, recognition of a face depends on whether or not the familiarity of the face meets the threshold of familiarity, or criterion threshold set by the individual making the judgment (Banks, 1970). It has previously been contended that, rather than encouraging a certain kind of judgment, sequential lineups may simply encourage an individual to be more conservative when making their lineup decisions. That is, that sequential lineups encourage the use of a stricter criterion threshold, in comparison to simultaneous lineups, which stays constant across the course of the lineup (Clark & Davey, 2005; Dobolyi & Dodson, 2013; Ebbesen & Flowe, 2002; Flowe & Ebbesen, 2007; Gronlund, 2004; Malpass, 2006; Malpass, Tredoux, & McQuiston-Surrett, 2009). In other words, the degree of match between the mental representation of the target and the lineup member required for participants to choose a lineup member
must be much closer for sequential than simultaneous lineups. This theory assumes that the memory of the target person would be equivalent between lineup procedures, with only the criterion for choosing being different, and that this criterion would not shift across the lineup. If a stricter criterion threshold was elicited by sequential lineups, this too would result in less false identifications of the suspect. However, there would also be less correct identifications and more incorrect rejections of the lineup because the stricter criterion for choosing simply results in the eyewitness being more conservative, or choosing less often. Current results, however, are inconsistent with this theory, as if stricter criteria for choosing were being elicited for one of the lineup procedures (namely sequential lineups), then that lineup procedure should be associated with less false identifications, less correct identifications and more incorrect rejections of the lineup. However, the present investigation has consistently demonstrated that simultaneous lineups lead to less false identifications and more correct identifications with no difference in the number of incorrect rejections of the lineup. If the memory of the target was equal for both lineup procedures and a strict matching threshold was applied over sequential lineups, then the present investigation would not have observed an increase in false identifications in sequential lineups, or an increase in false identifications with late target presentation in sequential lineups. Further, Experiment 1 of the investigation demonstrated that with late target presentation, participants who viewed sequential lineups responded with more false identifications, rather than more incorrect rejections of the lineup. These results suggest that participants who viewed sequential lineups did not tend to choose less often and were not more conservative, but rather showed a tendency towards falsely identifying a lineup member rather than rejecting the lineup. This is contrary to the theoretical proposition that sequential lineups make people more conservative, or elicit a stricter criterion threshold, than simultaneous lineups, as has been posed as a potential explanation in past research.

Given that the current investigation calls into question previous theoretical arguments explaining the difference in accuracy between sequential and simultaneous lineups, alternative explanations should be explored. Some previously reported data have been consistent with the current results, finding that sequential lineups are not always superior to simultaneous lineups in terms of accuracy (Carlson et al., 2008; Dobolyi & Dodson, 2013; Ebbesen & Flowe, 2002; Flowe & Ebbesen, 2007; Gronlund, 2005; Gronlund et al., 2009; McQuiston-Surrett et al., 2006; Mickes
et al., 2012; Zimmerman et al., 2006). For example, one study which sought to replicate the original findings of a sequential advantage reported more false identifications of the innocent suspect under sequential lineup procedures (Carlson et al., 2008). They also found that the sequential advantage only occurred for unfair lineups where the suspect stood out and was enhanced when the target was placed late in the lineup (Carlson et al., 2008). Another large study on differences in accuracy between sequential and simultaneous lineup procedures also found that there was no evidence for an advantage of either sequential or simultaneous lineup procedures under real world conditions where other factors, such as photograph quality, vary randomly (Gronlund et al., 2009). Further, research has found that discrimination of whether a guilty target is present or absent in the lineup is inferior in sequential lineups (Mickes et al., 2012). From these previous and current data, it appears possible that some points of difference between sequential and simultaneous lineup procedures may result in sequential lineups leading to inferior performance both in terms of correct and false identifications.

The most obvious point of difference in the sequential lineup procedure is the sequential presentation of lineup members. On face value, it is perhaps not surprising that this procedure led to decreased accuracy in the current investigation. Indeed, the sequential lineup procedure appears a more challenging task for several reasons. First, a longer time elapses over the lineup, and thus there is more time between observation of the original target and the presentation of that target in the lineup. As others have noted, sequential lineups require more, and possibly more detailed, information from memory storage as the basis for an identification response (Dobolyi & Dodson, 2013; Ebbesen & Flowe, 2002). Secondly, the procedure does not allow participants to form a baseline for visual similarity, or for forming their criterion threshold for choosing, and this may eventuate in a moving baseline and a shifting criterion threshold across the lineup. In other words, when participating in a sequential lineup a person does not know what faces are to come, or if they will view a more similar face later in the lineup to the one that they are viewing. Thus, they are forced to either decide that a face occurring early in the lineup is the best match that they will see and risk making a false identification early in the lineup, or hold out for more similar faces and risk a false rejection of the lineup. As such, it is perhaps logical that there may be some change in either the quality, or integrity, of
the memory of the target person, or in participants’ criterion threshold over sequential lineups. Indeed, if there was no change in either the integrity of the memory trace or participants’ criterion threshold for choosing over sequential lineups, an increase in false identifications in sequential lineups would not have been observed. As such, explanations for such changes across sequential lineups should be considered.

Although previous research has often assumed that the length of the lineup and position of the target person within the lineup has minimal effects on accuracy, this assumption may be in error (Clark & Davey, 2005; Lindsay & Wells, 1985). Conceptualisations of visual memory suggest that a pictorial code, or record of a visual element is created in memory when a person sees an unfamiliar face, and successful recognition of that face will occur when the information in the retrieval cue matches the information in the memory cue sufficiently to activate sets of memory representations (Atkinson & Juola, 1974; Bruce & Young, 1986). In simultaneous lineups, information that matches the memory cue can be searched for across the lineup, allowing the lineup participants to quickly determine whether a lineup member sufficiently activates the relevant sets of memory representations. In the sequential lineup procedure, matching information must be searched for in each lineup member over a longer period of time until information matches sufficiently enough to activate the relevant memory sets, or until the lineup ends. Therefore, in the sequential procedure, given its longer length, there may be more opportunity for interruptions or disruptions to the memory trace to occur. There may also be more opportunity for participants’ decision criterion threshold to change across the lineup. The current investigation has found some evidence that accuracy may be impacted upon by late presentation of the target in sequential lineups. For example, in the first experiment there were more false identifications and less hits and misses in sequential lineups when the target was positioned late in the lineup, and in the final experiment of the investigation, the difference in false identifications of a lineup member was due to more false identifications when the target was presented late in the lineup. This suggests that changes may be occurring over the course of sequential lineups which negatively impact accuracy. These changes may represent changes in memory, or changes in the criteria against which the judgment is measured.
There is much literature demonstrating that delays in time between an observed target and subsequent observation of that target can decrease the integrity of a given memory trace (Altmann & Gray, 2002; Gold et al., 2005; Jenkins et al., 2002; Pavlik & Anderson, 2005). In terms of visual memory, many researchers make the distinction between visual short term, or working, memory and visual long term memory (Downing, 2000; Hollingworth, 2004; Hollingworth, Richard, & Luck, 2008; Olson & Jiang, 2002; Phillips, 1974; Phillips & Baddeley, 1971; Zhang & Luck, 2009). Short term visual memory is thought to have a capacity of three or four items, lasting at least several seconds, although the precise maximum length has not been defined and is rarely discussed (Hollingworth, 2004). Visual long term memory maintains visual representations in a similar format to visual short term memory but has a remarkably large capacity and robust storage. As such, visual short term and long term memory are both very similar forms of memory, both being sensitive to object token, orientation, and the structural relationship between object parts, but being relatively insensitive to absolute size and precise object contours (Hollingworth, 2004; Hollingworth et al., 2008; Luck & Hollingworth, 2008). They are also both impacted by time delays, although the magnitude of the delay needed to impact memory appears to vary, with some research indicating that delays of one to nine seconds can decrease the integrity of visual short term memories (Alvarez & Cavanagh, 2004; Elliott & Anderson, 1995; Jenkins et al., 2002; Luck & Hollingworth, 2008; Luck & Vogel, 1997; Phillips, 1974; Vogel, Woodman, & Luck, 2001). Some research has also shown that the capacity of visual memory is dependant on both the number of items in storage and the information load of each item (Alvarez & Cavanagh, 2004). Therefore, if items are particularly complex, then fewer items will be able to be held in short term visual memory.

Given that faces contain more features and are more complex than stimuli often used in visual memory experiments (such as colours), it is possible that the different lineup procedures may challenge the capacity of visual memory differentially. Some research has shown that decision processes may be different for sequential than simultaneous lineups, with foils being analysed for a longer period of time in sequential lineups (Flowe, 2011). Sequential lineups may therefore represent a greater challenge if holding multiple faces in visual short term memory challenges visual memory capacity. This may make these memories more sensitive to
interferences due to small increases in the time delay, or viewing other distractor items, before seeing the target. If the time delay in sequential lineups had an impact on the integrity of visual memories or the capacity of the visual memory system, this may have resulted in an increased numbers of errors in general. This may have manifested as increases in false identifications and decreases in correct identifications of the target person. If this were the case, this effect would be more likely to be observed when the target person is presented late in the lineup, but may occur even when the target person is presented early if the particular recognition task is particularly sensitive to very small delays in time, or particularly complex and likely to overwhelm the capacity of visual short term memory. These effects were currently observed in the present investigation, as consistently more errors occurred in sequential lineups, and the number of these errors were exaggerated by late presentation of the target in some experiments. However, these findings were not always consistent for both same and different photograph conditions (e.g. Experiment 1), which would not be expected if a time delay explanation could be applied. Further, experiments on the VIPER system in England and Wales, have shown that lineup performance is superior when participants are required to view sequential lineups twice before making a decision, compared to when they make a decision on first viewing (Valentine et al., 2007). These findings are also inconsistent with the notion that increased time before making a decision impacts performance negatively. Additionally, while features of sequential lineups, such as a time delay or multiple decisions, are possible explanations for why accuracy was decreased in the current experiment, they are unlikely to explain differences between this and previous experiments within the literature. (See below for further discussion.)

Another possible theoretical explanation for why accuracy may have been decreased in sequential lineups in the current investigation can be understood in terms of signal detection theory (Banks, 1970). In other words, rather than a change in memory, some internal process, such as a desire to identify a face from the lineup, may have resulted in a downward shift in lineup participants’ criterion thresholds with serial presentation of lineup members. If participants had an increasing desire to identify a face from the lineup as the lineup went on, causing a downward shift in criterion threshold, a shift in confidence in choices across the lineup may also be anticipated. The expected direction of such a shift depends on how one believes
confidence of decisions relates to accuracy in general. There has been much discussion of this issue in the literature (Stretch & Wixted, 1998; Van Zandt, 2000). Although this debate has not been settled, there are several possibilities.

A very straightforward theory present within the literature is that confidence judgments are reached in much the same way as recognition decisions are reached (Balakrishnan & Ratcliff, 1996; Stretch & Wixted, 1998). In other words, the higher the criterion threshold, the higher a witness will estimate their confidence, the lower the threshold, the lower the confidence estimate. This theory is based on the assumption that a witness is, to some extent, aware of the shift in their criterion threshold. Other models suggest that confidence ratings may either diverge, or converge depending on the strength of evidence along the familiarity axis (Stretch & Wixted, 1998). The first model is considered to be a very useful conceptualisation of understanding confidence under unbiased conditions (Balakrishnan & Ratcliff, 1996; Stretch & Wixted, 1998; Van Zandt, 2000). If this model was correct and participants purposefully lowered their criterion threshold as the lineup progressed, out of an increasing desire to choose a face, a decrease in confidence across sequential lineups would be expected. Experiments 1 and 3 in the current investigation reported position effects for confidence. Experiment 1 data indicated that confidence in hits (correct choices) was higher when the target was presented early than late in the lineup. Thus, assuming that confidence judgments are made in the same way that recognition decisions are made, this finding partially supports the possibility that participants may have lowered their criterion threshold across lineups. Given that hit responses are always correct, lowered confidence when this choice was made late in the lineup may represent some evidence that participants were aware of lowering their criterion for this choice. However, this result was seen for sequential as well as simultaneous lineups, without an interaction effect suggesting the effect may not be restricted to sequential lineups, whereas increases in false positives across the lineup were only seen for sequential lineups. Findings from the third experiment, indicated that confidence in choosing target replacements was higher in sequential lineups when the target replacement was presented early but not when it was presented late. As such, findings when the target was absent are consistent with findings from Experiment 1, indicating that false identifications may have increased because participants lowered their criterion thresholds (which
therefore resulted in lower confidence in decisions) for choosing out of a desire to choose a face from the lineup.

Although it has previously been assumed that the sequential lineup procedure leads to a higher absolute threshold for familiarity than the simultaneous lineup procedure leading to a decrease in false positive responding (Clark & Davey, 2005; Clark et al., 2008), based on the current data the present author suggests that this is not the case. Previous research has generally assumed that the criterion threshold remains the same across sequential lineups. However, it is possible that this assumption is in error, and with each presentation of a lineup member in sequential lineups, a person’s criterion threshold may change. Simultaneous lineups, where all faces are presented at once, may provide participants with an opportunity to form a baseline, or to calibrate their familiarity threshold based on the range of visual similarity viewed across the lineup. However, in sequential lineups, this opportunity is not available and a person does not know whether they will view a highly similar face in the lineup prior to actually viewing that face. As such, a person’s threshold for recognising a target person may shift and change as different faces are presented to them in a sequential lineup. For example, if a person anticipates that there may not be many more faces presented as a sequential lineup progresses, they may feel an increased pressure to choose and lower their criterion threshold in order to choose a lineup member. This process may result in an increase in false identifications and a decrease in correct identifications, particularly if the pressure to choose could be felt by some participants fairly early in the lineup. All three of the current experiments indicated that false identifications were more likely to occur when targets or target replacements were presented late in sequential lineups, which is clearly consistent with a downward shift in criterion threshold across sequential lineups. This explanation would also be consistent with Clark and Davey’s (2005) Experiment 2 data which indicated that when target replacements were presented late in target absent lineups, there were increases in false identifications of the replacement and decreases in rejection rates. In other words, results indicated that some witnesses did exhibit a downward shift in their criterion threshold in order to identify a lineup member in lineups where the target person was absent. In other words, some participants’ desires to make an identification decision were sufficiently strong that they lowered their criterion threshold in order to do so. The current results strengthen and extend these findings, indicating that this downward shifting in
criterion may also occur and result in an increase in false identifications when the
target person is present in the lineup.

However, data from the third experiment when the target was present is not
consistent with this pattern of findings. These data indicated that there was no
difference in confidence in sequential lineups when the target was presented late
compared to when it was presented early. The most obvious difference between the
first and third experiments (where the target was present) was the similarity
manipulation, whereby photographs of the target were of low, medium or high
similarity to the target photograph. Manipulating task difficulty has been shown to
have an effect on confidence (Griffin & Tversky, 1992). Namely, people make
confidence estimates based on the balance of evidence available to them for and
against their judgment, and more difficult tasks often result in overconfidence in
decisions, particularly when the strength of evidence is perceived to be strong
(regardless of the reliability of this evidence). The presence of the similarity
manipulation in Experiment 3 may have made decisions more difficult on average
when compared to the first experiment, when no similarity manipulation occurred.
Further, when the target was presented late in the lineup, the participant had an
opportunity to collect evidence about whether the lineup members presented were
similar to the target they viewed earlier. Targets for the lower similarity condit-
ions may have appeared quite similar to some of the foils, thus the task may have been
quite difficult for participants. Further, when the target was presented late in the
lineup, several foils were likely to have been viewed before a decision was made.
Participants may have chosen the target or another foil because it was relatively
much more similar to the target than the previously seen foils (i.e. after looking at
many dissimilar foils, a very similar foil or target may have been viewed). This may
have eventuated in some increase in confidence in comparison to Experiment 1, or
reliance on the strength of the evidence they had obtained across the lineup (i.e.
belief that this must be the right choice because of the high degree of similarity
relative to the other foils seen), regardless of whether the choice they made was
correct or incorrect.

It must be noted that both time delay and criterion threshold explanations are
somewhat inconsistent with the current data for confidence ratings. On balance,
confidence data are more consistent with a criterion threshold explanation than the
time delay explanation, however some of these data are inconsistent with both explanations. One might expect that if the increased time delay in sequential lineups, represented a challenge for participants, this may have impacted upon confidence. If participants were aware that their memory of the target person was weakened, confidence may be decreased for sequential lineups, whereas if they were unaware of this effect confidence may remain unchanged. Similarly, if criterion thresholds were lowered as sequential lineups progressed and participants were aware of this, one might expect consistently lower confidence in false identifications in sequential lineups, when the target was presented late in the lineup, however this was only found for Experiment 1 and target absent lineups in Experiment 3. In other words, the pattern of confidence ratings appears relatively mixed across experiments. For example, the first experiment found increases in confidence across correct and false identifications in sequential lineups, whereas Experiment 2 found decreased confidence in sequential lineups for correct rejections of the lineup, but increases in confidence for false identifications when the target was absent. Experiment 3 found increases in confidence sequential lineups when the target was absent from the lineup. Although all experiments found some effect on confidence, the direction of these effects is inconsistent. Further study would aid the clarification of such effects.

Of note, however, is the consistent finding across experiments of higher confidence in false identifications in sequential lineups compared to simultaneous lineups. Particularly, this effect was seen for false identifications specifically of target replacements in Experiments 2 and 3. This is similar to the findings of Dobolyi and Dodson (2013) who found decreased accuracy in sequential lineups accompanied by higher confidence in false identifications. However, findings in this previous study were explained in terms of a criterion threshold account, as evidence for a higher criterion threshold in sequential lineup judgments. Given that no evidence of a higher criterion threshold in sequential lineups was found in the current study, this explanation is not likely to apply to the current data. The current findings, however, indicate that if participants’ memories of the target person were decreased across sequential lineups, they were unlikely to be aware of it, particularly when false identifications were made. Further, if participants were consistently lowering their criterion thresholds as sequential lineups progressed, they were not likely to have been consistently aware of this, otherwise one may have expected consistently
lower confidence in identification judgments when they did so. Confidence ratings, however, were inconsistent across experiments in this respect.

Based on the current results, it appears that theoretical explanations for differences in confidence may be unrelated to theoretical explanations for differences in accuracy between lineup procedures. For example while time delay or criterion threshold explanations may explain findings relating to accuracy, there may be other factors associated with sequential lineups that tend to increase confidence. Multiple decisions and repeated questioning have been found in previous experiments to increase confidence in responses due to participants’ feeling a pressure to choose in order to conform with perceived experimenter expectations (Hastie et al., 1978; Milgram, 1965; Shaw, 1996). This has been found independent of any changes in accuracy. While these experiments have focussed on the effect of repeated questioning in regards to a decision previously made by a participant, rather than repeated questioning about different decisions, the repeated questioning and multiple decisions that occur in the context of sequential lineups may nevertheless have some effect on confidence. Such possibilities deserve further investigation so that a coherent picture of the effect of lineup procedure on confidence and accuracy can be developed. At this stage, however, patterns of confidence, particularly those observed currently, appear unclear and inconsistent with most existing theoretical explanations.

While the aforementioned possibilities (changes in memory or in criterion threshold) for why sequential lineups may have resulted in lower accuracy than simultaneous lineups appear logical, it is not clear why these effects would not have been seen in previous experiments. Some differences exist between the current investigation and previous experiments. For example, the current investigation was designed as a memory paradigm, rather than a simulated real life lineup identification paradigm. As such, participants gave multiple lineup responses in a variety of different conditions in contrast to previous research in which only one lineup judgment per participant has generally been given. Further, many of the previous studies on eyewitness identification, have included a time delay of anywhere between five minutes and several days or months between observation of the target or suspect, and commencement of the lineup (Clark & Davey, 2005; Lindsay et al., 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b;
Wells et al., 1979). In the current experiments, however, observation of the target took place followed by the lineup procedure only a matter of 5 to 10 seconds afterwards. Further, previous experiments have tended to use live staging of crimes or video footage of staged crimes as the original observation of the target, whereas the current experiment used close up portrait photographs as targets (Clark & Davey, 2005; Lindsay et al., 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; Wells et al., 1979). These previous experiments, therefore, would be akin to the different photograph (or least similar photograph condition) in the current experiment. While it is possible that these factors may have had differential effects on accuracy and confidence in sequential and simultaneous lineups, it is unlikely that these differences would have eventuated in the current results. Particularly, many of the features of the current experiments would simply have served to decrease error in comparison to previous studies, making it more, rather than less, sensitive to detecting changes in memory processes.

While it could be argued that differences between the current and previous studies may be due to the difference in paradigms, other studies using similar lineup identification paradigms as in the majority of previous research have also failed to find a sequential advantage (Carlson et al., 2008; Gronlund et al., 2009; Mickes et al., 2012) and studies using multiple lineup trials have replicated findings of studies that have used a singular lineup test (Flowe & Ebbesen, 2007; Meissner et al., 2005). As such, it appears unlikely that the difference in paradigm was entirely responsible for differences in findings. There appears little reason why the current experimental paradigm would not have revealed the previously reported sequential superiority effect if one was present. As an investigation designed to examine the memory processes behind lineup judgments, the multiple responses given by participants over multiple lineup procedures should have increased experimental power increasing the likelihood of detecting such effects. Particularly, if the advantage of sequential over simultaneous lineups in terms of false identifications seen in past experiments was due to participants making absolute rather than relative judgments, or due to a stricter criterion threshold in sequential lineups, this advantage should have been observed in the current experiment. In other words, there appears no feasible explanation why the current investigation would not have revealed differences consistent with either absolute judgment processes or a stricter decision criterion threshold if such theoretical explanations were accurate. Contrary to these explanations, however,
current data find evidence for a simultaneous superiority effect for correct and false identifications which rebuts all major theoretical explanations for differences in eyewitness decision making under different lineup procedures. Therefore, to date, the results of the current investigation in light of previous research indicate that no firm conclusions regarding the mechanisms behind the impact of lineup procedure on accuracy and confidence can be drawn. Results do strongly suggest, however, that previously reported differences in accuracy between lineup procedures are not likely to be explained by relative and absolute judgment processes or by a higher criterion threshold in one type of lineup.

Regardless of the mechanism behind the present results, the current investigation demonstrates no clear benefit of sequential over simultaneous lineups in terms of false identifications or in terms of correct identifications and rejections. While the paradigm utilised in the current investigation is very different from that of previous experiments on eyewitness identification, the information it provides about the memory system in general remains relevant to memory processes that occur in all eyewitness identification situations. This information is inconsistent with the assertion that sequential lineups encourage absolute judgment processes whereas simultaneous lineups encourage relative judgment processes (Lindsay & Wells, 1985; Wells, 1984) and is also in conflict with major explanations involving signal detection theory (Banks, 1970; Clark & Davey, 2005). In contrast, the current experiments find some evidence of a simultaneous superiority effect, which calls into question the robustness of the previously reported sequential superiority effect (Sporer, 1993; Steblay et al., 2001; Steblay, Dysart & Wells, 2011). This has also been called into question by some previous studies utilising paradigms more consistent with the majority of the wider eyewitness literature (Clark & Davey, 2005; Ebbesen & Flowe, 2002; Flowe & Ebbesen, 2007; Gronlund, 2005; Gronlund et al., 2009; McQuiston-Surrett et al., 2006; Zimmerman et al., 2006). Further, late target position in the lineup is a factor that appears to have had markedly negative effects on accuracy particularly for sequential lineups across the current experiments. This finding is consistent with assertions made by McQuiston-Surrett and colleagues (2006) that variables independent of lineup procedure, such as target position, have an effect on accuracy and have been conflated with the sequential-simultaneous distinction in the past due to failures to control for such variables. The current
results add strength to these previous findings and suggest that further work is needed to determine precisely what factors may contribute to the observation of a sequential or a simultaneous superiority effect.

When considering the practical implications of current data, the problem concerning the conflicting findings becomes clear. Research showing the sequential superiority effect has been used to make recommendations in some jurisdictions to utilise sequential lineups in preference to simultaneous lineups (Klobuchar, Steblay, & Caligiuri, 2006; Lindsay, Mansour, Beaudry, Leach, & Bertrand, 2009; Wells et al., 2000; Wells et al., 1998). However, others have argued that making firm recommendations and rewriting policy surrounding the type of lineups utilised in these jurisdictions may be premature (Gronlund et al., 2009; Malpass, 2006; Malpass et al., 2009; McQuiston-Surrett et al., 2006). The current investigation supports the latter position and suggests very strongly that lineup administrators should question the validity and robustness of the sequential superiority effect. Present findings show no clear benefit of sequential over simultaneous lineups. The present experiments provide evidence of a simultaneous superiority effect which is contrary to the majority of other findings reported in the literature (Banks, 1970; Lindsay & Wells, 1985; McQuiston-Surrett et al., 2006; Sporer, 1993; Wells, 1984), but consistent with some (Carlson et al., 2008; Gronlund et al., 2009). Given that simultaneous lineups are also easier to administer from a practical point of view, as they are less lengthy and involve less manipulation by lineup administrators (Victoria Police, 1998), current results imply that simultaneous lineups may be more desirable than sequential lineups for ease of practical administration. As such, until inconsistencies in data concerning the effect of lineup procedure on accuracy are resolved, practical recommendations advocating or requiring the use of sequential lineups over simultaneous lineups continue to be premature.

6.2 The impact of visual similarity between the originally observed target and target in the lineup: Integration of findings, theoretical implications and practical application

The present investigation demonstrated a clear and consistent effect of visual similarity between the originally observed target and target observed in the lineup on eyewitness accuracy and confidence. Across all experiments, viewing either the same photograph, or a more similar photograph, of the target person in the lineup as
was originally observed, increased correct identifications and decreased false identifications and incorrect rejections of the lineup. This was accompanied by increases in confidence for correct identifications. Further, there was evidence to suggest that instating a target replacement in the context in which the original target was observed increased false identifications.

These data are consistent with a range of previous literature suggesting that consistency of context, clothing and other features between observations of a given suspect is highly advantageous to successful recognition of that suspect (Baddeley, 2014; Patterson & Baddeley, 1977; Thomson, 2003; Thomson et al., 1982). This is thought to be due to the way information is encoded and processed in memory. It is thought that both context (i.e. external features and context) and item information (i.e. faces) is encoded into memory upon observation of the item (Atkinson & Juola, 1974; Murnane, Phelps, & Malmberg, 1999). In other words, a pictorial code, or a record of a particular static visual element, is created in memory when a person sees an unfamiliar face (Bruce & Young, 1986). This includes contextual elements as well as the target item in question (Oliva & Torralba, 2007). The subsequent recognition of the item involves activating sets of memory representations based on the match between the information in the retrieval cue and the information in the memory cue (Atkinson & Juola, 1974). Thus, context information can enhance or facilitate recognition of the item depending on the match between what is seen and the representation in memory. Theorists also believe that implicit learning of contextual cues may occur in humans and this results in objects presented in familiar contexts being faster to recognise and localise (Oliva & Torralba, 2007). The current investigation lends support to these findings, implicating a strong effect of similar (or the same) context and features on accuracy and confidence. Higher visual similarity in context and external features of the target between observations leads to a better match between the retrieval and memory cues, and thus higher accuracy and increased confidence in correct responses. Further, current data lend additional support to the notion that a familiar context can impair accuracy if presented with an unfamiliar person, such that the facilitation effect of context may be sufficiently strong as to result in false identification of a similar target person. This is consistent with, and lends further strength to previous research findings that instating an unfamiliar person in a familiar context can lead to increases in false identifications of
that unfamiliar person (Thomson et al., 1982). Based on these findings, collectively, lineup participants may make lineup decisions based on similarity judgments - the more similar the lineup member to the previously seen target, the more likely that lineup member will be chosen. As such, it may be similarity, rather than identity, that eyewitnesses are making a decision about when asked to make an identity judgment. Further research on the construct of similarity and how it underpins lineup judgments in real life eyewitness situations may elucidate to what extent this construct may be responsible for differences in accuracy of lineup judgments.

The current results in relation to visual similarity indicate that moderate differences across observations of the target may not impair recognition, particularly if the target is viewed shortly after original observation, but that drastic differences are far more likely to disrupt recognition accuracy. The third experiment in the current investigation found that there was some overlap between highly similar and moderately similar photograph conditions in terms of accuracy, however that the least similar photograph condition consistently yielded decreased accuracy in recognition of the target. Further, results of this experiment also indicated that late presentation of a target in the lineup may interact with visual similarity to negatively impact eyewitness accuracy. In other words, moderate changes in appearance between original observation of the perpetrator and observation of the perpetrator in the lineup may not affect accuracy when the suspect is presented early, but may negatively impact accuracy when the suspect is presented late in the lineup. These findings collectively indicate that there may be a visual similarity threshold, whereby people are able to remain accurate despite some differences across observations of the target, however if the magnitude of the difference between observations reaches a certain threshold, then recognition is likely to be compromised. This threshold may be more difficult to meet if the target is presented late in the lineup, thus moderate differences may have more of an impact on accuracy. This threshold effect may logically be due to the target item not providing a sufficient match between the information in the retrieval cue and the information in the memory cue to activate the relevant sets of memory representations (Atkinson & Juola, 1974). On the basis of the current data, rather than differences across observations of a given target person degrading recognition accuracy in a linear fashion, there may be some threshold of matching information needed in the retrieval cue in order to activate the relevant memory representations and facilitate recognition of the item. If this threshold is not
met, however, it may be that the relevant memory sets are not activated and recognition of the item cannot take place. Further, this process of matching may be negatively affected by late presentation of the target in the lineup. This possibility, however, is currently speculative and further study would be needed to confirm such a possibility.

The observed effects of visual similarity between original observation of a target and observation in a lineup on accuracy have significant practical implications for lineup administrators. Strong evidence of the facilitation effect of visual similarity in external features and context suggests that correct identification is more likely to occur if the context and external features of a suspect is matched between original observation of the perpetrator and subsequent observation of the suspect in the lineup. A good match between these observations should lead to higher rates of correct recognition of a suspect. However, the collective findings also demonstrate that lineup administrators should ensure that the match in features and context between original observation and observation in the lineup is equal among lineup members. In other words, if a perpetrator was seen, for example, wearing a red hat and sunglasses, all lineup members should be displayed wearing a red hat and sunglasses. If the external features and context between original observation and observation in the lineup is only matched for the suspect, current data and previous research suggests that false identification of that suspect is highly likely in the case that the suspect is not the perpetrator of the crime. Likewise, this collective evidence further emphasises the dangers of matching context and features between original observation and observation upon identification if conducting a ‘show up’ or single person identification. Such a procedure has been identified in previous cases to be highly dangerous (Alexander v R, 1981; Davies and Cody v R, 1937; R v Williams, 1983), particularly when the suspect resembles the perpetrator (Steblay, Dysart, Fulero, & Lindsay, 2003) and the current research suggests that providing a familiar context in this scenario would likely add to this danger. Further, matching features between observations may be particularly important if the target is presented late in the lineup. On the basis of the current results, where a moderate level of difference between the originally observed perpetrator and suspect observed in the lineup is expected, early presentation of the target is likely to enhance the chances of successful recognition of the suspect, and may therefore be preferable to late
presentation. This appears to be the case for both simultaneous and sequential lineup procedures, suggesting that suspect position in the lineup should be an area for consideration for administrators of both sequential and simultaneous lineups. It must be noted, however, that the current investigation did not examine whether matching external features such as clothing across all lineup members facilitates recognition of a target from the lineup. Thus, this possibility remains speculative and warrants further research.

6.3 The relationship between confidence and accuracy: Integration of findings, theoretical implications and practical application

Current data add to the information available regarding the relationship between confidence and accuracy within the context of eyewitness identification. Results from the present investigation consistently support the argument that, on average, confidence of correct identifications of a target was higher than that of false identifications of a lineup member or target replacement. Further, there was some evidence indicating that confidence of correct rejections was higher than that of incorrect rejections of the lineup but this difference was only significant in sequential lineups in the final experiment of the investigation. In simultaneous lineups, this difference was not significant. However, the patterns of confidence in the different response types indicated that the relationship between confidence and accuracy was perhaps more complicated than the group averages implied. Conditions where accuracy was high did not always match with conditions where confidence was high. For example, there was consistent evidence across experiments that accuracy of simultaneous lineups was superior to that of sequential lineups, however confidence in incorrect responses, particularly false identifications was consistently higher in sequential than in simultaneous lineups. Therefore, although the current experiments may provide some evidence of a relationship between confidence and accuracy, there are conditions under which this relationship may not be present.

A central aim of the current investigation was to examine the relationship between confidence and accuracy, as whether or not eyewitness confidence should be regarded as a measure of accuracy represents a major conflict between the court rulings in the USA and rulings in the UK (Neil v Biggers, 1972; R v Turnbull, 1976). The current experiments demonstrate a consistent difference between the confidence of correct and false identifications, such that on average confidence of correct
identifications is higher than that of false identifications. Current results, however, do not indicate a consistent difference between the confidence of correct and incorrect rejections of the lineup, with only the last experiment demonstrating such a relationship only when sequential lineups are viewed. This is consistent with findings in previous research indicating that people who are accurate in their lineup identifications are also likely to be confident (Bradfield et al., 2002; Brewer et al., 2002; Brewer & Wells, 2006; Lindsay et al., 1998; Sporer et al., 1995) but that this relationship is significantly weaker for non-choosers (Sporer et al., 1995).

The finding that the relationship between confidence and accuracy is consistently weaker for non-choosers than for choosers may indicate that there may be different memory processes occurring for recognition of a target, than for recognition of the absence of a target. From previous research on visual memory it is argued that visual information related to an item is encoded into memory upon observation of the item. The subsequent recognition of the item is thought to involve activating sets of memory representations based on the match between the information in the retrieval cue and the information in the memory cue (Atkinson & Juola, 1974). A sufficient match should eventuate in recognition of the item. However, this process may be slightly more complex when considering the decision that a previously seen item is not present. It is possible that there is more than one way in which such a decision may be made. For example, the inverse of the recognition process may occur, and there may be an insufficient match between information in the retrieval cue and the information in the memory cue, resulting in a failure to recognise an item, or in the case of eyewitnesses, a rejection of the lineup. This process may occur when a person views a target person, then views a lineup comprised of people with features similar to the target. The person would reject the lineup if one of the lineup members does not form a sufficient match with their memory of the target. However, a second possibility is that a person may reject an item based on elements in the retrieval cue which are particularly dissimilar when compared with information in the memory cue. In other words, there may be information in the retrieval cue that is markedly different from that same information in the memory cue, representing a mismatch with the information in the memory cue. This process may occur when a person views a target person and is presented with a lineup comprised of members that are of a different colouring or ethnicity to the
target person. The decision to reject may not be based on insufficient matching information, but rather on information that mismatches with information in the memory cue. As such, it is possible that while the decision to recognise a lineup member may involve a process of matching information about a lineup member with information in the retrieval cue about the target person, the decision to reject the lineup may involve several different processes. Thus, rejecting the lineup may involve different, or more complex, memory processes to identifying a person from a lineup which may decrease the relationship between confidence in lineup rejections and accuracy. This possibility, however, is currently speculative and requires further research.

In terms of identifying a lineup member, a relationship between confidence and accuracy has particularly been observed when simple identification tasks are used, or studies have manipulated variables that impair or facilitate accuracy. For example, Lindsay and colleagues (1998) manipulated exposure time to a video tape depicting individuals that participants were later asked to identify. Participants who viewed the video for longer were more accurate and more confident in their identifications. However, the manipulation of exposure time would have been likely to yield higher levels of both accuracy and confidence independent of one another. Similarly, Bradfield and colleagues (2002) found a highly significant relationship between confidence and accuracy, however the task used involved participants watching videos of targets for three minutes in a variety of contexts and positions making the task very basic. Simple identification tasks are known to enhance the relationship between confidence and accuracy (Weber & Brewer, 2004). This is logically due to a high relationship between actual accuracy and belief in accurate identifications. Further, this would be likely to be due to a high degree of consistency or good match between the memory cue and the retrieval cue, which would facilitate accuracy and also enhance confidence due to participant awareness of the high quality match with their encoded memory.

The tasks used in the current experiment varied in difficulty with some conditions being simpler than others. Current data demonstrated that viewing the same, or a more similar, photograph of the target in the lineup as that originally viewed facilitated accuracy and enhanced confidence significantly. The relationship between confidence and accuracy under these conditions may have been independently related to the variable of visual similarity (or same vs different
photograph) which was manipulated across all experiments. As such, both accuracy and confidence may have been independently related to photograph condition, and not necessarily to each other. This may have partially accounted for the overall elevation in confidence and accuracy. Certainly, it was seen in the second experiment that the difference in confidence between correct and false identifications was exaggerated when the same photograph of the target person as was presented in the lineup was viewed. This difference in confidence also appeared to be reduced with decreasing visual similarity in the final experiment. However, the higher confidence for correct identifications was also seen for conditions where a different photograph, or dissimilar photograph, of the target person was viewed, suggesting that the relationship between confidence and accuracy seen within present data may be generalised to more difficult tasks as well. On the basis of current data, it is suggested that many participants may have remained aware of the accuracy of their decisions, or of the integrity of their memory of the target person, despite more difficult lineup procedures. This may have been due to a relatively high degree of perceived visual similarity between their pictorial memory of the target person and the retrieval cue activated by the target person in the lineup despite visual similarity being decreased. These possibilities, however, remain speculative and require further investigation to elucidate.

While superficially, the current results appear consistent with the USA Supreme court’s claim that confidence is a reliable indicator of accuracy (Neil v Biggers, 1972), it must be noted that results pertain only to group averages. In other words, the present results indicate that on average confidence across individuals is higher when they make correct identifications than when they make false identifications. However, the ranges of confidence for correct and false identifications across the three experiments indicate that this does not equate to implying any predictive value of an individual’s confidence rating to their accuracy. Across experiments, it was seen that the range of confidence ratings for both correct and false identifications ranged from one to five, or encompassed the full range of the confidence scale, albeit that higher responses were more often seen for correct identifications. In other words, overlap in confidence ratings between incorrect and correct identifications was significant. As such, while some participants may have had much higher confidence for their correct than false identifications, it is likely that
others rated their confidence in correct identifications equally to that of their false identifications, or rated their confidence in false identifications higher than another person rated their correct identifications. Equally, some participants were highly in their false identifications. This points to the caveat that current observations speak to a general pattern of results across individuals, but cannot be used to imply that high confidence can necessarily predict accuracy on an individual level. Results instead imply that in order to discover whether confidence may be used as an indicator of accuracy for any given individual, a baseline of that individual’s confidence in general must first be established. This, however, is not possible within the context of an eyewitness making an identification in a courtroom where no such baseline is available. As such, due to the range in results observed for confidence currently, in this respect, present data indicate that confidence may not be a reliable indicator of accuracy. Further research would be needed, however, to investigate the extent to which confidence and accuracy may vary together, or are calibrated, as the current investigation did not utilise a calibration approach. Nevertheless, the assertions of the USA Supreme court that confidence is a reliable indicator of accuracy on an individual level appear inappropriate, and the approach of the British courts, which have taken into account several individual cases in which eyewitness certainty was not a reliable indicator of accuracy, may be more appropriate (Devlin, 1976; Neil v Biggers, 1972; R v Turnbull, 1976).

A further complication for the relationship between confidence and accuracy is the consistent observation across current experiments that the system variable of lineup procedure influenced confidence in judgments independently from its effects on accuracy. In other words, confidence of false identifications in sequential lineups is higher than that in simultaneous lineups. This is despite decreased accuracy in sequential lineups in general. This appears to be a valid example of the possibility that confidence can be disconnected from accuracy in some situations. In other words, elevated confidence may be an indicator of the kind of lineup administered rather than the accuracy of the response. On a theoretical level, the reasons for this effect are unclear, particularly given the absence of such an effect in previous experiments. However, the effect of lineup procedure on confidence has rarely been comprehensively assessed within the literature, possibly explaining the failure to identify such an effect in the past (Sporer et al., 1995). It is possible that some elements of sequential lineups may enhance confidence, particularly when a false
identification is likely to occur. In addition to factors examined within the current investigation, factors such as disruption to the memory trace with an increasing time delay between original observation and observation in the lineup; participant fatigue as the lineup progresses; repeated questioning; multiple decisions; and integration of intervening information being incorporated into the original memory trace, are all factors that may interact in some way to elevate confidence in false identifications in sequential lineups (Altmann & Gray, 2002; Ebbesen & Rienick, 1998; Hastie et al., 1978; Jenkins et al., 2002; Loftus, 1979, 2005; Loftus & Greene, 1980; Pavlik & Anderson, 2005; Shaw, 1996). However, at this stage it is unclear how these factors may contribute to, or interact to produce, such an effect. Further research on the impact of lineup procedure on confidence and accuracy would be necessary to determine the mechanisms underlying this effect.

On a practical level, however, the findings relating to the impact of lineup procedure on confidence in false identifications are noteworthy. This is the case particularly given that two of the current experiments provided evidence that participants who viewed sequential lineups exuded more confidence for incorrect identification of target replacements than those who viewed simultaneous lineups. These results suggest that elevated confidence may be an indicator of the kind of lineup administered, rather than being due to increased accuracy. As such, relying on confidence as an indicator of accuracy may be particularly problematic if sequential lineups are utilised. Thus, if confidence is relied upon as an indicator of accuracy, it would be pertinent to warn the jury of the potential impacts of sequential lineups and to advise them of which lineup procedure has been administered. This is relevant to some jurisdictions in the USA, in which official lineup procedures are being rewritten to preference sequential lineups upon research-based recommendations (Klobuchar et al., 2006; Lindsay et al., 2009; Malpass, 2006; Wells et al., 1998) and in which confidence is also viewed as a reliable indicator of accuracy (Neil v Biggers, 1972). Thus, in the courtroom the observation of a highly confident witness may indicate that an accurate witness is being observed or it may indicate the administration of a sequential lineup. Under these circumstances, which possibility is applicable cannot be determined. This is notable, particularly given knowledge gained from previous literature that juror decision making is highly influenced by the confidence of eyewitnesses (Fox & Walters, 1986; Lindsay et al.,
1989; Whitley & Greenberg, 1986). As such, results from the current investigation suggest that the practice of relying on confidence as an indicator of accuracy, particularly if sequential lineups are readily administered, may be dangerous and may lead to increases, rather than decreases, in wrongful conviction rates.

6.4 **Strengths, limitations and future directions**

A number of strengths were present in the current investigation. A major strength of the current investigation is that participants within the experiment provided multiple responses that contributed to a number of different conditions. This has not been the case in the majority of previous studies in which single lineup observations have generally been used (Clark & Davey, 2005; Lindsay et al., 1998; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Malpass & Devine, 1981b; Wells et al., 1979). The current protocol allowed for tighter experimental controls, in terms of controlling for differences in lineup characteristics between lineups, and increases the power of the current investigation and its findings considerably. Additionally, the current experiments used a memory paradigm to investigate the impacts of lineup procedure on accuracy and confidence. Given that the human memory system is central to eyewitness lineup judgments, information about this system is widely applicable to the processes behind, and practice of, eyewitness lineup identifications. Further, with regards to method, in the third experiment, visual similarity ratings were used to manipulate visual similarity of target photographs to those in the lineup. This is rare in the eyewitness literature and provides clear benefits in terms of accounting for subjectivity of the variable over procedures which change different numbers of aspects of a target person presented in the lineup. Another major strength of the present investigation is novelty in its attempt to extend the current literatures related to the examination of confidence and accuracy when different lineup procedures are utilised. The current investigation represents one of the most comprehensive examinations of accuracy and confidence utilising different lineup procedures and across different photograph similarity conditions. Further, these factors were investigated across three experiments, strengthening support for findings that were replicated across experiments. To the author’s knowledge this is the first attempt to examine the impact of all of these factors and their interactions on the separate response types as well as on the relationship between confidence and accuracy in general.
Although this study had a number of notable strengths, it is not without its limitations. The most notable limitation relevant to all of the current experimental studies is the limited ecological validity which may affect the generalisability of findings. There was no viewing of a crime in the current experiments, and as such they were not representative of real life eyewitness identification scenarios. In a real life situation, viewing a crime may induce heightened emotions and other factors that were not accounted for by the current study. Some researchers suggest that the presence of emotional arousal, whether elicited by external stimuli, internal events, or stress hormones, can modulate memory and memory consolidation processes (Mather & Sutherland, 2011; Wagner, Hallschmid, Rasch, & Born, 2006). According to theorists, arousal modulates the strength of competing mental representations, enhancing memory for the most dominant items contesting for selective attention. It is thought that this process begins during perception and continues into long-term consolidation. In other words, arousal biases contesting items in favour of goal-related or perceptually conspicuous stimuli, and then enhances memory consolidation for the most conspicuous or relevant stimuli regardless of whether those stimuli were arousing or not. As such, emotional arousal can enhance memory if the dominant stimuli are emotionally arousing, but interfere with memory if the dominant stimuli are not what has contributed to emotional arousal, or when there are multiple high priority items present at once (Mather & Sutherland, 2011). As such, it is possible that the emotional arousal in real life eyewitness situations has some effect on eyewitness memory that cannot be captured by the present investigation. The present investigation is, however, relevant to the process of memory for faces in general and prioritises investigation of these memory processes in a lineup-like situation. Further, not all eyewitnesses experience emotional arousal at the time of seeing a given perpetrator, and so the current investigation is, perhaps, more relevant to these situations.

Another limitation is the comparison of the difference between means for confidence, rather than taking a calibration approach. In other words, the extent to which confidence is calibrated with accuracy was not examined, only whether or not the mean confidence was significantly higher for accurate than inaccurate responses. This limits interpretation, as it remains unclear the extent to which confidence and accuracy may vary together, regardless of whether the means are significantly
different or not. Further, as no mock witness tests were run for the lineups employed, it is not clear whether some lineups may have been biased, or contained targets or target replacements that stood out more than usual. This would be unlikely, given that identification rates in the current investigation were not high for any lineup. Thus, the results of the current thesis do add valuable information to the literature, which does not typically investigate the impact of a number of variables, and interactions between them, on confidence and accuracy.

Another factor that may affect the generalisability of findings is the use of multiple lineup judgments in the current paradigm. In real life identification judgments only a single lineup judgment occurs whereas in the current experiment participants engaged in multiple lineup decisions. That said, other studies using similar lineup identification paradigms as the majority of previous research have also failed to find a sequential advantage (Carlson et al., 2008; Gronlund et al., 2009; Mickes et al., 2012) and studies using multiple lineup trials have replicated findings of studies that have used a singular lineup test (Flowe & Ebbesen, 2007; Meissner et al., 2005). Therefore, it appears unlikely that the difference in paradigm was entirely responsible for differences in findings.

It must also be noted that while the current experiment used a very different paradigm than has been commonly utilised in previous studies to investigate memory processes in eyewitness lineup judgments, the information that has been provided speaks to the nature of the memory system underlying eyewitness identification processes. This information clearly contradicts major explanations within the literature for differences in accuracy between lineup procedures, for example, that sequential lineup procedures encourage absolute judgment processes. The current experiment, however, compared only pure forms of sequential and simultaneous lineup procedures and did not consider the impact of hybrid lineup procedures on accuracy and confidence. Further studies comparing the impact of hybrid designs with that of purely simultaneous and sequential lineups procedures may shed light onto which specific elements of these procedures may hinder or enhance accuracy and impact confidence.

Another limitation present in the current studies is that for two out of the three experiments (Experiments 2 and 3), only male Caucasian lineup members were utilised. This was because gender and ethnicity were not variables that were relevant to the investigation’s focus, and because the majority of lineup members in nearly all
jurisdictions are predominantly male and the majority of lineup members in Australia are Caucasian (Australian Institute of Criminology, 2013; Office for National Statistics, 2015; Truman & Langton, 2015). Given that ethnicity and gender of lineup members have previously been shown to have an effect on accuracy of identifications, findings from these experiments cannot be generalised to lineups for female suspects or suspects of different ethnicities (Gross, 2009; Wright & Sladden, 2003). In terms of sample characteristics, the present studies included mainly Caucasian participants, and as such the current findings cannot necessarily be generalised to eyewitnesses of different races. Further, the current studies were not strictly controlled in terms of sample sizes, and as such there were unequal sample sizes in each cell of the designs. This may mean that the current experiments were compromised in terms of statistical power for some analyses. However, given that statistical effects relating to the central hypotheses were still observed this was unlikely to have been a major issue for the current studies. It must be noted, however, that failure to find significance for some effects or interactions may have been related to type II errors on account of reduced statistical power.

Another limitation with respect to study measurements, is that throughout the current investigation there was no examination of participant fatigue or effort. Although the online design of the study allowed a larger quantity and variety of respondents in terms of age and characteristics than other recruiting methods, it did limit the ability to assess for the effects of fatigue or effort. There were a substantial number of participants excluded because they did not complete the survey (387, 168 and 75 for Experiments 1, 2 and 3 respectively), however it must be noted that there are reasons other than participant fatigue that can contribute to this number (such as computer or network failure). Although any participants who dropped out of the survey, as well as those who gave the same rating for all answers, were excluded from the studies, it is possible that some participants may have become fatigued or lacked effort towards the end of the study. Additionally, there may have been interference effects as participants had seen more faces by the last lineup than the first one. It is possible that fatigue or interference effects may have been more likely for participants allocated to viewing sequential lineups, such that these participants may have been more fatigued towards the end of these lineups. Completion time was on average between four, and five and a half, minutes longer in sequential than
simultaneous lineups. However, it is unlikely that fatigue or interference effects had a marked impact as no order effects on accuracy were observed for simultaneous or sequential lineups. Nevertheless, future studies on accuracy and confidence of lineup judgments should examine the levels of fatigue across different lineup procedures and whether this may have a differential impact on responding in simultaneous and sequential lineups.

Another limitation with regard to the third experiment, is that the visual similarity manipulation was based on the subjective judgments of eleven independent participants, rather than on objective elements such as numbers of features varied in the photograph. Although this manipulation is a previously mentioned strength, as more objective strategies are not always found to align with participant judgments of visual similarity, it must be noted that the manipulation was based on subjective opinion which may or may not have been consistent with the subjective judgments of the participants. While generally, the trend in accuracy suggested that the manipulation was successful, it is possible that the overlap in terms of accuracy between the highly similar and moderately similar photograph conditions was due to the subjective nature of the manipulation.

While it was a strength of the current investigation to examine accuracy and confidence in different response types, it is a limitation that there was no “I don’t know” option in the current studies. While this is consistent with the literature on eyewitness lineups in general (Clark & Davey, 2005; Clark et al., 2008; Cutler et al., 1987; Lindsay & Wells, 1985; Malpass & Devine, 1981a; Sporer et al., 1995), it does dictate that none of the responses can be differentiated from when participants were guessing or when they rejected the lineup because they were not sure. Confidence ratings could give some indication, however low confidence does not necessarily indicate having no idea of whether a response was accurate or not. It may be useful for future studies to allow participants to indicate whether or not they believed they knew the answer when making lineup judgments.

Finally, in all three experiments there was a higher percentage of females than males in the sample. All three samples were also predominantly Caucasian. Previous psychological research has found accuracy is increased when participants are asked to identify faces of the same race (Wright, Boyd, & Tredoux, 2001) or gender (Wright & Sladden, 2003) as themselves. Although it was beyond the scope of the current research to investigate gender or ethnicity effects, it is possible that
gender and ethnicity could have had some unknown effects or interactions with lineup type, target position, or photograph similarity that may have impacted on patterns of accuracy and confidence in the current study. Future research should investigate the whether the impacts of currently examined system and estimator variables on accuracy and confidence may interact with own-gender or own-race biases.

6.5 Conclusions

The current experiments investigating the effect of lineup procedure and photograph similarity on accuracy and confidence in lineup judgments contribute a variety of unique information to many existing literatures. Specifically, they add important information to literatures related to accuracy and confidence of lineup judgments, as well as that on the relationship between the two variables. The present investigation provides consistent evidence across three experiments of a simultaneous superiority effect, rather than the commonly reported sequential superiority effect, with results indicating less false identifications and more correct identifications when simultaneous lineups were completed. These findings strongly suggest that differences in accuracy between simultaneous and sequential lineups cannot be accounted for by prominent theories within the literature, such as relative versus absolute judgment processes, or a stricter criterion threshold being elicited by one lineup procedure. Further, consistent evidence across all of the current experiments for higher confidence in false identifications in sequential lineups is provided. These two findings are novel within the literature and have wide reaching practical implications for the administration of lineup procedures in criminal matters. Current experiments also strengthen the wealth of literature relating to the effect of similarity in features and context between observation of a suspect and observation of that suspect in the lineup. These studies demonstrated that the facilitative effect of consistency across observations on accuracy is strong. The current investigation extends these findings to confidence, with greater consistency across observations leading to higher confidence in accurate responses. These findings further highlight the benefits of ensuring consistency in external features and context across observations, but also emphasise a need to keep such consistency equal among lineup foils and suspects alike in order to guard against false identifications. Finally, the
present studies provide consistent evidence that while, in terms of group means, confidence of correct identifications is higher than that of false identifications, that there is a great degree of overlap in the confidence in incorrect and correct responses. As such, although a person who is confident is perhaps more likely to be accurate, individual variation dictates that accuracy cannot be predicted reliably from confidence.

The present investigation highlights a need for further examination of the effects of lineup procedure on accuracy and confidence. These experiments suggest that prominent theoretical explanations of the differences in accuracy between simultaneous and sequential lineups may be in error. Furthering our understanding of the mechanisms behind such differences can inform legal policy surrounding the use of different types of eyewitness lineups in criminal cases. Results also suggest that the use of sequential lineups in preference to simultaneous lineups may be highly dangerous and lead to increases in wrongful conviction rates, particularly in jurisdictions that are likely to rely on confidence as a reliable indicator of accuracy (*Neil v Biggers*, 1972). Particularly, the use of sequential lineups may lead to inflated confidence in false identifications, including false identification of an innocent suspect. The practical implications of these results support the proposition that until inconsistencies between this investigation and previous investigations within the literature are resolved, recommendations indicating that sequential lineups should be preferred in criminal proceedings continue to be premature (Malpass, 2006; Malpass et al., 2009; McQuiston-Surrett et al., 2006). The present investigation also highlights a need for further exploration of the relationship between confidence and accuracy, and what system variables may influence confidence irrespective of accuracy. Clarifying these effects would allow research in this area to more effectively inform legal policy, and aid in reducing the rates of wrongful convictions for those falsely identified from a lineup in criminal cases.
REFERENCES


References


References


References


Henkel, L. A. (2017). Inconsistencies across repeated eyewitness interviews: supportive negative feedback can make witnesses change their memory
Reports. Psychology, Crime and Law, 23(2), 97-117. doi: 10.1080/1068316X.2016.1225051


References


References


References


Appendix A: Advertisement, Plain Language Statements and Consent Forms

PARTICIPANTS WANTED

aged 18 to 70 years

TO PARTICIPATE IN A COMPUTER BASED STUDY ON AGE AND IDENTITY JUDGMENTS IN PHOTOGRAPHIC LINEUPS

Requires Approximately Half an Hour to Complete

IF YOU ARE INTERESTED, PLEASE GO TO THE FOLLOWING WEB ADDRESS TO COMPLETE THE STUDY

https://ecuaq.qualtrics.com/SE/?SID=SV_5dJ9UygkIncyngF

Please Note:

PARTICIPATION IS PURELY VOLUNTARY AND IS OUTSIDE THE SCOPE OF ANY ACADEMIC CURRICULUM

School of Psychology, Faculty of Health
Deakin University
This study is an online identification study. The aim of the study is to determine ways to improve identification procedures in the criminal justice system.

If you agree to participate in this study you will be asked to view 10 identification parades, each following the same procedure. First you will be shown a photograph of a target person. In the interval that follows you will be asked to estimate the age of a different person. Finally, you will be shown an identification parade and asked whether you are able to identify the target person.

Please read the statements below before deciding whether to participate in the study.

I understand that:

- Ethical approval has been granted by the Human Ethics Advisory Group, Faculty of Health, Deakin University.
- Participation in the study will take between 20 and 30 minutes.
- My data will be stored in a secure location.
- My data will be strictly anonymous and be treated confidentially.
- My data may be used for associated publications, but no identifiable information will be released.
- My participation is entirely voluntary and there are no penalties for non-participation.
- I have the right to withdraw my participation at any time and for any reason.
- However, I understand that if I withdraw my participation after commencing the study it will not be possible to identify or erase any of the answers I have already provided.

At the conclusion of this study you can obtain a summary of results by emailing the principal researcher at donald.thomson@deakin.edu.au.

Contact inquiries about this research can be made by emailing the principal researcher at donald.thomson@deakin.edu.au.

If you have any complaints about any aspect of this study, the way it is being conducted, or any questions about your rights as a research participant, then you may contact:

The manager, Research Integrity, Deakin University
221 Burwood Highway, Burwood, Victoria 3125
Telephone: (03) 9251-7129 Email: research-ethics@deakin.edu.au
Please quote project number [HEAG-H92-2013]

If you are willing to participate in this study please select ‘I agree to participate’ below and then click next.

☐ I agree to participate
Appendix B | Ethics Approvals

APPENDIX B
Ethics Approval: Experiments 1 and 2

Memo

To: Professor Donald Thomson
School of Psychology

From: Secretary – HEAG-H
Faculty of Health

CC: Julia Chan, Micaiah Zwartz, Carlo Butera, Danielle Drew, Natalie Roberts, Tess Ryan, Alyssa Sisa, Kate Bibby, Ema Blatancic, Blanca Lebenholc, Navjeet Kalsi, Lynn Khoo, Michal Kene, Melissa McDonald

Date: 25 July, 2013

Re: HEAG-H 92_2013: The role of similarity in person identification

Approval has been given for Professor Donald Thomson, of the School of Psychology, to undertake this project for a period of 3 years from 25 July, 2013. The current end date for this project is 25 July, 2016.

The approval given by the Deakin University HEAG-H is given only for the project and for the period as stated in the approval. It is your responsibility to contact the Secretary immediately should any of the following occur:

- Serious or unexpected adverse effects on the participants
- Any proposed changes in the protocol, including extensions of time
- Any events which might affect the continuing ethical acceptability of the project
- The project is discontinued before the expected date of completion
- Modifications that have been requested by other Human Research Ethics Committees

In addition you will be required to report on the progress of your project at least once every year and at the conclusion of the project. Failure to report as required will result in suspension of your approval to proceed with the project.

An Annual Project Report Form can be found at:

This should be completed and returned to the Administrative Officer to the HEAG-H, Pro-Vice Chancellor’s office, Faculty of Health, Burwood campus by Tuesday 19th November, 2013 and when the project is completed. HEAG-H may need to audit this project as part of the requirements for monitoring set out in the National Statement on Ethical Conduct in Human Research (2007).

Good luck with the project!
Appendix B | Ethics Approvals

[Signature]

Steven Sawyer
Secretary
HEAG-H
# Ethics Approval: Experiment 3

## Memo

| To:       | Professor Don Thomson  
| School of Psychology |
| From:     | Secretary – HEAG-H  
| Faculty of Health |
| CC:       | Dr Adrian Scott, Micaiah Zwartz, Julia Chan, Stephanie Ellis, Melissa Allsopp,  
| Taylor Craggill-Haren, Rhiannon Hartney, Zrinka Ilicic, Haley James, Jennifer  
| Nicoll, Mollie Powell, |
| Date:     | 26 May, 2015 |
| Re:       | HEAG-H 92_2013: The role of similarity in person identification |

Approval has been given for Professor Don Thomson, of the School of Psychology, to undertake this project with the modifications that were requested on the 25 May, 2015.

Please note that the current end date for this project is 31 December, 2015.

---

**Steven Sawyer**  
**Secretary**  
**HEAG-H**
APPENDIX C
Example of the Online Survey

Simultaneous: Practice Task 1

For each identification parade, you will be shown a photograph of a target person for five seconds.

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

You will then be shown a photograph of a different person for five seconds. You will be asked to estimate this person’s age and to indicate how confident you are that your estimation is accurate.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Finally, you will be shown an identification parade containing eight photographs of different persons. These photographs will be shown together.

You will be asked to look at each photograph and in your own time indicate whether you are able to identify the target person. You will then be asked to indicate how confident you are that you correctly identified the target person or that the target person was not shown in the photographs.

Practice Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

You have now completed the first practice identification parade. Please click 'Next' when you are ready to begin the second practice identification parade.

Simultaneous: Practice Task 2

Target Person

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.

Appendix C | Example of the Online Survey

Qualtrics Survey Software:

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Practice Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

You have now completed both of the practice identification parades and should be familiar with the procedure. Please click 'Next' when you are ready to begin the main study.

Simultaneous 1.1: Caucasian male, Same, Lineup 1, Position 3

Target Person

Appendix C | Example of the Online Survey

Qualtrics Survey Software:

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

Appendix C  |  Example of the Online Survey

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 1.2: Caucasian female, Same, Lineup 1, Position 3

Target Person

Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 1.3: Asian male, Same, Lineup 1, Position 6

https://example.qualtrics.com/CentralPanelAjax.php?action=GetSurveyPrintPreview&SurveyId=32001201484010 AM}
Appendix C | Example of the Online Survey

Qualtrics Survey Software

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 1.4: Asian female, Same, Lineup 1, Position 6

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.

Appendix C  |  Example of the Online Survey

Quality Survey Software

Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

**Simultaneous 1.5: Caucasian male, Different, Lineup 2, Position 2**

Target Person

These page timer metrics will not be displayed to the recipient.

- First Click: 0 seconds.
- Last Click: 0 seconds.
- Page Submit: 0 seconds.
- Click Count: 0 clicks.

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 1.6: Caucasian female, Different, Lineup 2, Position 2

![Target Person](https://ecnu.qualtrics.com/ControlPanelAjax.php?action=GetSurveyPrintPreview&T=ask47[2001/2014.08.10 AM])

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.

First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
Appendix C | Example of the Online Survey

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident
Slightly confident
Moderately confident
Very confident
Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph.
Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 1.7: Asian male, Different, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Appendix C | Example of the Online Survey

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 1.8: Asian female, Different, Lineup 2, Position 7

These page timer metrics will not be displayed to the recipient.
- First Click: 0 seconds.
- Last Click: 0 seconds.
- Page Submit: 0 seconds.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 1.8: Asian female, Differen, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
- First Click: 0 seconds.
- Last Click: 0 seconds.
- Page Submit: 0 seconds.
- Click Count: 0 clicks.

How old do you think this person is?

[Image of a person]

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 2.1: Caucasian male, Same, Lineup 2, Position 2

Target Person

Appendix C | Example of the Online Survey

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

### Appendix C | Example of the Online Survey

Qualtrics Survey Software

<table>
<thead>
<tr>
<th>How confident are you that your estimation is accurate (give or take two years)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all confident</td>
</tr>
</tbody>
</table>

**Identification Parade**

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

![Identification Parade](image)

<table>
<thead>
<tr>
<th>How confident are you that you have correctly identified the target person?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all confident</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How confident are you that the target person was not shown in the photographs?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not at all confident</td>
</tr>
</tbody>
</table>

**Simultaneous 2.2: Caucasian female, Same, Lineup 2, Position 2**

![Simultaneous 2.2](image)

[Link to online survey](https://example.qualtrics.com/ControlPanelAjax.php?action=GetSurveyPrintPreview?surveyID=3800120148-40.10.10 AM)
Appendix C | Example of the Online Survey

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

[Images of photographs]
- Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 2.3: Asian male, Same, Lineup 2, Position 7

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 2.4: Asian female, Same, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present
Appendix C | Example of the Online Survey

Qualtrics Survey Software:

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 2.5: Caucasian male, Different, Lineup 1, Position 3

![Target Person]

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

Not at all confident •  Slightly confident •  Moderately confident •  Very confident •  Extremely confident •

How confident are you that the target person was not shown in the photographs?

Not at all confident •  Slightly confident •  Moderately confident •  Very confident •  Extremely confident •

Simultaneous 2.6: Caucasian female, Different, Lineup 1, Position 3

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

![Image of face recognition survey](https://example.com/image.png)

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 2.7: Asian male, Different, Lineup 1, Position 6

Target Person

These page timer metrics will not be displayed to the recipient.
- First Click: 0 seconds.
- Last Click: 0 seconds.
- Page Submit: 0 seconds.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 2.8: Asian female, Different, Lineup 1, Position 6

Target Person

Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

Appendix C | Example of the Online Survey

Qualtrics Survey Software

**How confident are you that your estimation is accurate (give or take two years)?**

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**Identification Paradigm**

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

- Not present
- ... (Multiple photographs)

**How confident are you that you have correctly identified the target person?**

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**How confident are you that the target person was not shown in the photographs?**

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**Simultaneous 3.1: Caucasian male, Same, Lineup 1, Position 6**

[Image of a photograph]

<table>
<thead>
<tr>
<th>Target Person</th>
</tr>
</thead>
</table>

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

**How old do you think this person is?**

**How confident are you that your estimation is accurate (give or take two years)?**
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**Identification Parade**
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

**Not present**

**How confident are you that you have correctly identified the target person?**
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**How confident are you that the target person was not shown in the photographs?**
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

**Simultaneous 3.2: Caucasian female, Same, Lineup 1, Position 6**

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Appendix C | Example of the Online Survey

Simultaneous 3.3: Asian male, Same, Lineup 1, Position 3

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 3.4: Asian female, Same, Lineup 1, Position 3

![Target Person]

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

**Appendix C | Example of the Online Survey**

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 3.5: Caucasian male, Different, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident
Slightly confident
Moderately confident
Very confident
Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 3.6: Caucasian female, Different, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?
Appendix C | Example of the Online Survey

How confident are you that your estimation is accurate (give or take two years)?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 3.7: Asian male, Different, Lineup 2, Position 2

Target Person

Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

Example of the Online Survey

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 3.8: Asian female, Different, Lineup 2, Position 2

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 4.1: Caucasian male, Same, Lineup 2, Position 7

Appendix C | Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Appendix C | Example of the Online Survey

Qualtrics Survey Software

Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident
Appendix C | Example of the Online Survey

Simultaneous 4.2: Caucasian female, Same, Lineup 2, Position 7

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.
Appendix C | Example of the Online Survey

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present
Appendix C | Example of the Online Survey

Qualtrics Survey Software:

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 4.3: Asian male, Same, Lineup 2, Position 2

Target Person

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

https://ecuan.qualtrics.com/ControlPanelAjax.php?action=GetSurveyPreview&T=ask&l=2000/12/31/4:00:00 AM
Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?
Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade
If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present
Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 4.4: Asian female, Same, Lineup 2, Position 2

![Target Person]

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.

https://ecntr.qualtrics.com/ControlPanelAjax.php?action=GetSurveyPrintPreview&amp;T=20804201484010AM}
Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

---

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.


281
Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 4.5: Caucasian male, Different, Lineup 1, Position 6

Target Person

These page timer metrics will not be displayed to the recipient.
- First Click: 0 seconds.
- Last Click: 0 seconds.
- Page Submit: 0 seconds.
Appendix C | Example of the Online Survey

Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?
Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that your estimation is accurate (give or take two years)?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?
- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 4.6: Caucasian female, Different, Lineup 1, Position 6

Target Person

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

How old do you think this person is?

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How confident are you that your estimation is accurate (give or take two years)?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Identification Paradox

If you are able to identify the target person please select the corresponding photograph. Alternatively, select ‘Not present’ if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

How confident are you that the target person was not shown in the photographs?

- Not at all confident
- Slightly confident
- Moderately confident
- Very confident
- Extremely confident

Simultaneous 4.7: Asian male, Different, Lineup 1, Position 3

Appendix C | Example of the Online Survey

Qualtrics Survey Software

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph. Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Simultaneous 4:8: Asian female, Different, Lineup 1, Position 3

Appendix C  |  Example of the Online Survey

Qualtrics Survey Software

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.
Last Click: 0 seconds.
Page Submit: 0 seconds.
Click Count: 0 clicks.

These page timer metrics will not be displayed to the recipient.
First Click: 0 seconds.

Appendix C | Example of the Online Survey

Qualtrics Survey Software

Last Click: 0 seconds.  
Page Submit: 0 seconds.  
Click Count: 0 clicks.

How old do you think this person is?

How confident are you that your estimation is accurate (give or take two years)?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

Identification Parade

If you are able to identify the target person please select the corresponding photograph.  Alternatively, select 'Not present' if you do not think the target person is shown in the photographs.

Not present

How confident are you that you have correctly identified the target person?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

How confident are you that the target person was not shown in the photographs?

Not at all confident  Slightly confident  Moderately confident  Very confident  Extremely confident

### Demographic Information

Congratulations, you have now completed the eight identification phases of the main study. Please provide the following demographic information and click 'Next'.

**What is your age?**

- [ ] 18-24
- [ ] 25-34
- [ ] 35-44
- [ ] 45-54
- [ ] 55+

**What is your sex?**
- [ ] Male
- [ ] Female

**What is your ethnicity?**
- [ ] Caucasian
- [ ] African
- [ ] Asian
- [ ] Middle Eastern
- [ ] Aboriginal or Torres Strait Islander
- [ ] Pacific Islander
- [ ] Mixed ethnicity (more than one of the above ethnicities)
- [ ] Other ethnicity (please specify)

**Were you born in Australia?**
- [ ] Yes
- [ ] No
Example of the Online Survey

Qualtrics Survey Software

How long have you lived in Australia?

Survey Powered By Qualtrics