

How interviewing techniques and  
temperament affect children as eyewitnesses  
and jurors' perceptions.

Ben F. Cotterill, BA (Hons), MSc, AFHEA

A thesis submitted in partial fulfilment of the  
requirements of Edinburgh Napier University,  
for the award of Doctor of Philosophy

August 2021

## **Declaration**

---

I hereby declare that the work presented in this thesis has not been submitted for any other degree or professional qualification, and that it is the result of my own independent work.

Ben F. Cotterill

---

Full Name Goes Here (Candidate)

28/04/2021

---

Date

## **Abstract**

---

The aim of this research was to better understand how the temperament of a child witness may affect their eyewitness performance and the perception of mock-jurors during different interview conditions. In Study One, a new self-report method (the Temperament Assessment Tool for Children; TATC) was developed for assessing temperament traits in 4- to 8-year-old children. Internal consistency and test-retest reliability was found to be satisfactory overall with 202 participants. In Study Two, the same 4- to 8-year-old children watched a video of a theft and were interviewed using open, closed, or misleading questions about what they remembered. The children also completed the TATC. Witnesses in the open-ended condition were significantly more accurate overall than those in the other interview conditions. Temperament characteristics that limit attention (i.e., distractibility) predicted an overall higher number of errors during the interview, while misleading questions were particularly detrimental to the accuracy of less adaptable and less persistent children. In Study Three, a sample of 82 mock-jurors read transcripts of interviews with child witnesses (6 to 8 years old) from the previous study. Mock-jurors were randomly assigned to one of three interview conditions (open-ended, closed-ended, or misleading). The child witnesses varied in levels of shyness (i.e., high, moderate, and low) according to self-reports, gathered in the previous study. Contrary to the researcher's hypothesis, the findings indicated that interview condition had no impact on the mock-jurors' perceived reliability of witnesses. Furthermore, the shyness of the child witness overall did not affect the perception of mock-jurors. Mock-juror perceptions were, however, influenced by their own personality traits. For example, those higher in withdrawal, an aspect of neuroticism, were more likely to rate child witnesses favourably. Overall, the results of this research provide some support for the assertion that certain temperament traits can impact the eyewitness performance of child witnesses and that closed-ended and misleading questions are unwise due to an overall decrease in accuracy.

## Publications associated with this research

---

### AWARDS

The SIPR and Police Scotland Student Award Sponsored by Palgrave for The Best Presentation 2019

### REFEREED PUBLICATIONS

Cotterill, B. F. (2020). Gender and suggestibility. *Forensic Update*, 134, 22-24.

### REPORTS

Cotterill, B. F. (2020). *How temperament and interviewing techniques affect children as eyewitnesses*. In The Thirteenth Scottish Institute for Policing Research Annual Report 2019/20. Retrieved from <http://www.sipr.ac.uk/publications/research-papers-reports>

### REFEREED CONFERENCE PUBLICATIONS

Cotterill, B. F. (2020, June). *How temperament and interviewing techniques affect children as eyewitnesses*. Presentation at the British Psychological Society's Division of Forensic Psychology Annual Conference, Solihull, UK. (Conference cancelled)

Cotterill, B. F. (2020, February). *How temperament and interviewing techniques affect children as eyewitnesses*. Poster at the Society for Personality and Social Psychology Annual Convention, New Orleans, LA, USA.

### INVITED PRESENTATIONS AND COLLOQUIA

Cotterill, B. F. (2020, September). *Child witnesses: the role of individual differences*. Invited Guest Lecturer for the MSc in Applied Forensic Psychology and Criminology, Edinburgh Napier University, Edinburgh, UK.

Cotterill, B. F. (2020, February). *Children as eyewitnesses*. Presentation to the School of Applied Sciences – Discovery Seminar, Edinburgh Napier University, Edinburgh, UK.

Cotterill, B. F. (2019, December). *How temperament and interviewing techniques affect children as eyewitnesses and the decision making of jurors*. Presentation at the Scottish Institute for Policing Research (SIPR) Postgraduate Symposium, Edinburgh, UK.

Cotterill, B. F. (2019, May). *How temperament and interviewing techniques affect children as eyewitnesses*. Poster at the School of Applied Science Postgraduate Conference, Edinburgh Napier University, Edinburgh, UK.

*Publications associated with this research*

Cotterill, B. F. (2019, May). *How interviewing techniques and temperament affect children as eyewitnesses and the decision making of jurors*. Presentation at the Scottish Institute for Policing Research (SIPR) Postgraduate Roundtable Researching and Networking Event, Edinburgh, UK.

Cotterill, B. F. (2018, April). *How shyness and leading questions affect children as eyewitnesses*. Presentation at the School of Applied Science Postgraduate Conference, Edinburgh Napier University, Edinburgh, UK.

---

## **Acknowledgements**

---

I would like to extend my deepest gratitude to my three supervisors (Faye Skelton, Rory MacLean, and Kathy Charles), not only for their unwavering belief in me and this project, but also for their consistent availability and willingness to provide advice, feedback, and support. They have truly gone above and beyond.

I would also like to thank the participants who took part in this project and the schools that granted me access. Without them, this project would not have been possible.

Lastly, thank you to my wonderful family who continue to support me with everything I do.

## Table of contents

---

Declaration .....	i
Abstract .....	ii
Publications associated with this research .....	iii
Acknowledgments .....	v
Table of contents .....	vi
List of figures .....	xiii
List of tables .....	xiv
<b>1. Chapter 1: Introduction .....</b>	<b>1</b>
<b>1.1. The scope of children in the legal system .....</b>	<b>3</b>
<b>1.2. Day care sexual abuse panic and early research .....</b>	<b>4</b>
<b>1.3. An introduction to suggestibility .....</b>	<b>7</b>
<i>1.3.1. Source monitoring .....</i>	<i>11</i>
<b>1.4. Suggestibility and individual differences .....</b>	<b>12</b>
<i>1.4.1. Cognitive and personality factors .....</i>	<i>15</i>
<b>1.5. Juror perceptions .....</b>	<b>18</b>
<b>1.6. Summary .....</b>	<b>19</b>
<b>2. Chapter 2: Literature review .....</b>	<b>20</b>
<b>2.1. Memory and development .....</b>	<b>20</b>
<i>2.1.1. Externally-driven and internally-driven false memories .....</i>	<i>21</i>
<i>2.1.2. Developmental differences in memory .....</i>	<i>24</i>
<i>2.1.2.1. Development of autobiographical memory .....</i>	<i>25</i>
<i>2.1.3. Other age-related differences in memory recall .....</i>	<i>25</i>
<i>2.1.3.1. Knowledge differences .....</i>	<i>25</i>
<i>2.1.3.2. Language differences .....</i>	<i>25</i>
<i>2.1.3.3. Theory of mind .....</i>	<i>27</i>

2.1.3.4. <i>Own-age bias</i> .....	27
2.1.3.5. <i>Social influences</i> .....	28
2.1.3.6. <i>Other influences</i> .....	29
2.1.3.7. <i>Children's ability to lie</i> .....	29
2.1.4. <i>Summary: Memory</i> .....	30
2.2. <b>Children's describing and identifying capabilities</b> .....	31
2.2.1. <i>Children answering nonsensical questions</i> .....	31
2.2.2. <i>Forensically-relevant conversations with children</i> .....	32
2.2.3. <i>Accuracy of child eyewitnesses at describing</i> .....	35
2.2.4. <i>Accuracy of child eyewitnesses at identifying</i> .....	38
2.2.5. <i>Summary: Children's describing and identifying capabilities</i> .....	39
2.3. <b>External factors of suggestibility</b> .....	40
2.3.1. <i>Question types</i> .....	40
2.3.2. <i>Question repetition</i> .....	43
2.3.3. <i>Other external factors of suggestibility</i> .....	45
2.3.3.1. <i>Delay and repeated interviews</i> .....	45
2.3.3.2. <i>Interviewer bias</i> .....	46
2.3.3.3. <i>Interviewer status</i> .....	48
2.3.3.4. <i>Emotional tone</i> .....	48
2.3.4. <i>Current interviewing formats</i> .....	49
2.3.4.1. <i>The National Institute of Child Health and Human Development</i> <i>(NICHD) investigative interview protocol</i> .....	49
2.3.4.2. <i>Achieving Best Evidence (ABE)</i> .....	51
2.3.4.3. <i>Stepwise interview</i> .....	51
2.3.5. <i>Summary: External factors of suggestibility</i> .....	52
2.4. <b>Internal factors of suggestibility</b> .....	52

## Contents

2.4.1. <i>Temperament</i> .....	54
2.4.1.1. <i>Activity</i> .....	63
2.4.1.2. <i>Adaptability</i> .....	64
2.4.1.3. <i>Distractibility</i> .....	64
2.4.1.4. <i>Emotionality</i> .....	65
2.4.1.5. <i>Persistence</i> .....	66
2.4.1.6. <i>Shyness</i> .....	66
2.4.1.7. <i>Measuring temperament</i> .....	67
2.4.2. <i>Summary: Internal factors of suggestibility</i> .....	68
2.5. <b>Perceptions of children's credibility as eyewitnesses</b> .....	69
2.5.1. <i>Effects of witness age</i> .....	69
2.5.2. <i>Effectors of interviewing technique</i> .....	70
2.5.3. <i>Effects of witness confidence and shyness</i> .....	71
2.5.4. <i>Effects of juror gender</i> .....	71
2.5.5. <i>Effects of juror age</i> .....	72
2.5.6. <i>Effects of juror personality</i> .....	72
2.5.7. <i>Summary: Perceptions of children's credibility as eyewitnesses</i> .....	73
2.6. <b>Research objectives</b> .....	73
2.6.1. <i>Research objective 1</i> .....	74
2.6.2. <i>Research objective 2</i> .....	74
2.6.3. <i>Research objective 3</i> .....	74
3. <b>Chapter 3: Study one – The temperament assessment tool for children (4 to 8 years old): a self-report technique</b> .....	76
3.1. <b>Introduction</b> .....	76
3.1.1. <i>Measurements of temperament</i> .....	77
3.1.2. <i>The difficulty with caregiver questionnaires</i> .....	81

Contents

3.1.3. <i>Children's self-reports</i> .....	82
3.1.4. <i>Gender differences in temperament</i> .....	84
3.1.5. <i>Research aims</i> .....	86
<b>3.2. Methods</b> .....	<b>87</b>
3.2.1. <i>Design</i> .....	87
3.2.2. <i>Participants</i> .....	88
3.2.3. <i>Materials</i> .....	88
3.2.3.1. <i>The Temperament Assessment Tool for Children (TATC)</i> .....	88
3.2.3.2. <i>Scoring of the TATC</i> .....	90
3.2.3.3. <i>TATC Parent form</i> .....	91
3.2.3.4. <i>Distraction task</i> .....	91
3.2.3.5. <i>Persistence task</i> .....	92
3.2.4. <i>Procedure</i> .....	93
<b>3.3. Results</b> .....	<b>94</b>
3.3.1. <i>Reliability</i> .....	94
3.3.2. <i>Validity</i> .....	95
3.3.3. <i>Individual differences</i> .....	97
3.3.3.1. <i>Temperament and gender</i> .....	97
3.3.3.2. <i>Temperament and age</i> .....	98
<b>3.4. Discussion</b> .....	<b>100</b>
3.4.1. <i>Reliability</i> .....	100
3.4.2. <i>Validity</i> .....	102
3.4.3. <i>Gender differences in temperament</i> .....	105
3.4.4. <i>Strengths and weaknesses</i> .....	106
<b>3.5. Summary</b> .....	<b>108</b>

<b>4. Chapter 4: Study two – How temperament and interviewing technique affect children as eyewitnesses .....</b>	<b>109</b>
<b>4.1. Introduction .....</b>	<b>109</b>
<b>4.1.1. Question types .....</b>	<b>109</b>
<b>4.1.2. Age-related differences in suggestibility .....</b>	<b>111</b>
<b>4.1.3. Temperament .....</b>	<b>111</b>
<b>4.1.4. Research aims .....</b>	<b>113</b>
<b>4.2. Methods .....</b>	<b>114</b>
<b>4.2.1. Design .....</b>	<b>114</b>
<b>4.2.2. Participants .....</b>	<b>114</b>
<b>4.2.3. Materials .....</b>	<b>115</b>
<b>4.2.3.1. Stimulus .....</b>	<b>115</b>
<b>4.2.3.2. Interview schedules .....</b>	<b>115</b>
<b>4.2.3.3. Temperament .....</b>	<b>116</b>
<b>4.2.4. Procedure .....</b>	<b>117</b>
<b>4.2.5. Coding .....</b>	<b>117</b>
<b>4.3. Results .....</b>	<b>119</b>
<b>4.3.1. Interview format and age .....</b>	<b>119</b>
<b>4.3.2. Temperament .....</b>	<b>122</b>
<b>4.4. Discussion .....</b>	<b>123</b>
<b>4.4.1. Interview format and age .....</b>	<b>123</b>
<b>4.4.2. Temperament .....</b>	<b>125</b>
<b>4.4.3. Conclusions .....</b>	<b>127</b>
<b>5. Chapter 5: Study three – How temperament and interviewing techniques affect jurors' perceptions of witness credibility .....</b>	<b>129</b>
<b>5.1. Experiment one: Introduction .....</b>	<b>129</b>

Contents

5.1.1. Questioning style .....	130
5.1.2. Perceived confidence and shyness of witnesses .....	131
5.1.3. Research aims .....	132
<b>5.2. Experiment one: Methods .....</b>	<b>132</b>
5.2.1. Design .....	132
5.2.2. Participants .....	133
5.2.3. Materials .....	133
5.2.4. Procedure .....	134
<b>5.3. Experiment one: Results .....</b>	<b>135</b>
5.3.1. Manipulation checks .....	135
5.3.1.1. Perceived bias of interview questions .....	135
5.3.1.2. Perception of shyness levels .....	135
5.3.2. Interview format and shyness .....	136
5.3.2.1. Shyness: Open-ended condition .....	139
5.3.2.2. Shyness: Closed-ended condition .....	140
5.3.2.3. Shyness: Misleading condition .....	140
5.3.2.4. Perceived shyness .....	140
5.3.3. Summary: Results .....	141
<b>5.4. Experiment one: Discussion .....</b>	<b>141</b>
5.4.1. Questioning style .....	141
5.4.2. Shyness .....	143
<b>5.5. Experiment two: Introduction .....</b>	<b>145</b>
5.5.1. Personality and the Big Five model .....	145
5.5.2. Juror personality .....	150
5.5.2.1. Openness .....	151
5.5.2.2. Conscientiousness .....	152

Contents

5.5.2.3. <i>Extraversion</i> .....	153
5.5.2.4. <i>Agreeableness</i> .....	154
5.5.2.5. <i>Neuroticism</i> .....	155
5.5.3. <i>Research aims</i> .....	156
<b>5.6. Experiment two: Methods</b> .....	<b>156</b>
5.6.1. <i>Design</i> .....	156
5.6.2. <i>Participants</i> .....	156
5.6.3. <i>Materials</i> .....	157
5.6.4. <i>Procedure</i> .....	157
<b>5.7. Experiment two: Results</b> .....	<b>157</b>
<b>5.8. Experiment two: Discussion</b> .....	<b>158</b>
<b>5.9. Conclusions</b> .....	<b>159</b>
<b>6. Chapter 6: General discussion</b> .....	<b>162</b>
<b>6.1. Research objectives</b> .....	<b>162</b>
6.1.1. <i>Research objective 1</i> .....	162
6.1.2. <i>Research objective 2</i> .....	163
6.1.3. <i>Research objective 3</i> .....	163
<b>6.2. Impactors of witness performance</b> .....	<b>164</b>
<b>6.3. Methodological considerations</b> .....	<b>168</b>
<b>6.4. Future recommendations</b> .....	<b>169</b>
<b>6.5. Conclusions</b> .....	<b>172</b>
<b>7. List of References</b> .....	<b>173</b>
<b>8. Appendices</b> .....	<b>264</b>

## List of figures

---

<b>Figure 1.1. Temperament dimensions, according to models by Martin (1988) and Rothbart (2007), as developmental precursors for DeYoung's (2006) Big Five personality traits .....</b>	<b>61</b>
<b>Figure 3.1. Examples of TATC items .....</b>	<b>90</b>
<b>Figure 3.2. Distraction condition of distraction task (Forster &amp; Lavie, 2016) .....</b>	<b>92</b>
<b>Figure 3.3. Wooden puzzle .....</b>	<b>93</b>
<b>Figure 5.1. Interaction between interview on condition and shyness level on mean reliability .....</b>	<b>137</b>
<b>Figure 5.2. Interaction between interview on condition and shyness level on mean confidence .....</b>	<b>138</b>
<b>Figure 5.3. Interaction between interview on condition and shyness level on mean compelling .....</b>	<b>139</b>

## List of tables

---

<b>Table 1.1. Interview extracts from Clarke-Stewart et al. (2004)</b> .....	<b>13</b>
<b>Table 3.1. TATC Definitions</b> .....	<b>87</b>
<b>Table 3.2. Reliabilities for temperament categories of TATC</b> .....	<b>94</b>
<b>Table 3.3. Mean item score (1-5) within each temperament dimension for self-reports and parent forms, and the interrater consistency correlation</b> .....	<b>96</b>
<b>Table 3.4. Summary of reliability and validity scores for each subscale</b> .....	<b>97</b>
<b>Table 3.5. Mean score (between 8 and 40) for each temperament trait for both females and males</b> .....	<b>98</b>
<b>Table 3.6. Mean score (between 8 and 40) for each temperament trait for both younger and older children</b> .....	<b>99</b>
<b>Table 3.7. Reliability and validity scores for younger and older children</b> .....	<b>99</b>
<b>Table 4.1. List of central and peripheral descriptors</b> .....	<b>118</b>
<b>Table 4.2. Mean correct descriptors, errors, and overall accuracy based upon interview format and age</b> .....	<b>121</b>
<b>Table 5.1. Interviews with child witnesses used for the study</b> .....	<b>134</b>
<b>Table 5.2. Perceptions of child witnesses based upon interview condition</b> .....	<b>136</b>
<b>Table 5.3. Perceptions of child witnesses based upon shyness levels</b> .....	<b>136</b>
<b>Table 5.4. Perceptions of child witnesses based upon shyness levels in the open-ended condition</b> .....	<b>139</b>
<b>Table 5.5. Perceptions of child witnesses based upon shyness levels in the closed-ended condition</b> .....	<b>140</b>
<b>Table 5.6. Perceptions of child witnesses based upon shyness levels in the misleading condition</b> .....	<b>140</b>
<b>Table 5.7. Big Five Aspect Scale means and internal consistencies</b> .....	<b>158</b>
<b>Table 8.1. TATC correlation matrix</b> .....	<b>273</b>

*Contents*

**Table 8.2. Open-ended condition correlation matrix ..... 304**

**Table 8.3. Closed-ended condition correlation matrix ..... 304**

**Table 8.4. Misleading condition correlation matrix ..... 304**

**Table 8.5. Younger children: Open-ended condition correlation matrix ..... 304**

**Table 8.6. Younger children: Closed-ended condition correlation matrix ..... 305**

**Table 8.7. Younger children: Misleading condition correlation matrix ..... 305**

**Table 8.8. Older children: Open-ended condition correlation matrix ..... 305**

**Table 8.9. Older children: Closed-ended condition correlation matrix ..... 305**

**Table 8.10. Older children: Misleading condition correlation matrix ..... 306**

**Table 8.11. Correlation matrix: Juror personality and questionnaire responses ....**

**322**

## **Chapter 1: Introduction**

---

This thesis will examine temperament as an impactor of children's eyewitness performance. To achieve this, a self-report tool for measuring temperament in young children will be constructed and assessed for reliability and validity. A second study will examine the influence of interview techniques and temperament on children's eyewitness performance. Finally, the impact of question types and shyness on perceptions of child witnesses according to mock-jurors will be investigated. First, it is important to outline why impactors of children's eyewitness performance is a topic worthy of investigation.

Faulty eyewitness accounts are the leading cause of wrongful convictions; since the Innocence Project was founded in 1992, it has helped overturn 360 wrongful convictions with post-conviction DNA evidence. Of these 360 cases, witness misidentifications were present in 71% of them (Innocence Project, 2020). To date, the National Registry of Exonerations (2020) has a record of 2,625 cases of men and women wrongfully convicted in the United States. Collectively, these men and women have spent over 23,000 years in prison for crimes they did not commit. Evidence shows that children make poorer eyewitnesses than adults, both remembering less and making more errors at facial recognition (Ceci & Bruck, 1995; Hershkowitz, Lamb, Orbach, Katz, & Horowitz, 2012; Pozzulo & Dempsey, 2006).

Ricardo Rachell is one of these 2,625 cases. In 2002, an 8-year-old boy ran down a street in Houston, crying and screaming. He was taken back to his mother and made a statement to the police that someone had held a knife to his throat and tried to kill him. The next day, he and his family saw neighbour Ricardo Rachell riding his bike in town. The boy identified Rachell as his attacker and added that Rachell had sexually assaulted him. Rachell was charged with sexual assault of a child and was sentenced to 40 years in prison. In 2007, Rachell requested DNA testing after doing some investigating of his own. He argued that the actual person responsible was Andrew Hawthorne (a known serial rapist in the area). In 2008, DNA tests proved that Rachell was innocent and that Hawthorne had committed the crime.

Disclosures of sexual abuse by children are only treated seriously by parents and reported to police in the UK and United States 26% to 30% of the time (Royal College of General Practitioners and National Society for the Prevention of Cruelty to Children, 2014; Snyder, 2000), even though it is estimated that only about 6% of child sexual abuse accusations in the UK or USA are false (Ceci & Bruck, 1995; Everson & Boat, 1989; for a review, see O'Donohue, Cummings, & Willis, 2018). This means many offences against children occur without ever resulting in an arrest. Since the child's testimony is often the only available evidence in cases involving child witnesses (Brewer, Rowe, & Brewer, 1997; Lamb & Brown, 2006), it may be that forensic interviews with children are not eliciting sufficient evidence and this is a contributing factor to low arrest rates. A 14-year-old French girl in 2013 reported to her school counsellor that her father had been sexually assaulting her. She was told that her words would not be evidence enough to convict him, so she set up a camera in her room and filmed the next sexual assault (Goodman, Goldfarb, Chong, & Goodman-Shaver, 2014). The contrasting examples of the American boy and the French girl show the dilemmas that investigators face when children are the primary eyewitnesses. Therefore, it must be carefully considered how much weight should be placed on the importance of eyewitness testimony. If testimonies are believed without question, it could result in wrongful arrests and convictions, but if they are not believed at all then guilty and dangerous offenders will not be apprehended.

The goal of the present research was to explore possible factors that may contribute towards the amount of correct information and/or errors that children provide during police interviews, and to investigate how these same factors may affect the perceptions of jurors when determining the credibility of those children as witnesses. Specifically, the researcher wanted to determine whether the negative effects (i.e., poor accuracy and credibility) of closed-ended and misleading questioning (e.g., Peterson, Dowden, & Tobin, 1999; Waterman, Blades, & Spencer, 2000) are heightened when they are applied to children with certain temperament qualities. This would demonstrate the importance of taking a child's personality into account when undertaking forensic interviews with children.

## **1.1 The scope of children in the legal system**

Young children were included as eyewitnesses in even the earliest European records of criminal proceedings, dating as far back as the mid-1500s (Ceci & Bruck, 1993). In 1736, Sir Matthew Hale noted that children appeared in English courts most often as witnesses for crimes of rape, buggery, and witchcraft (Edelstein, 1996). However, things were quite different in North America. Following the devastation of the Salem witch trials (during which a group of young girls accused 200 people of being witches, and 19 of them were sentenced to death), children were viewed as unreliable witnesses, and were largely prohibited from testifying in courts. Canada did not enact its first legislation on child eyewitnesses until 1893 (Bala, 2018), and the US Supreme Court did not permit child eyewitnesses to testify until 1895 (Pantell, 2017). Currently, in North America and the United Kingdom, jurisdictions have no minimum age requirement for testifying in courts. Instead, it is up to the judge to decide whether or not the child is competent to testify, meaning the judge must determine if the child knows fact from fiction, as well as the difference between telling the truth and a lie. In October of 2017, newspapers reported that a 2-year-old girl became the youngest eyewitness to testify in an English court (Bowcott, 2017).

The reliability of child eyewitnesses is especially a point of concern given the large number of children who become involved in the legal system every year. In Scotland (a country with a population of about 5.25 million people; National Records of Scotland, 2012), children were asked to give testimony in a criminal court 4,297 times during 2017 (Crown Office and Procurator Fiscal Service, 2018). Five-hundred and thirty-seven of the witnesses were below the age of 10, and 209 were below the age of 8. These numbers only include criminal proceedings, and so do not include the numerous court appearances of children in civil cases or divorce proceedings. In 2009, the BBC uncovered that 47,817 court witnesses were 17 years old or younger during the previous year across England and Wales (Crawford, 2009). The figures showed that 1,116 of the witnesses were under the age of 10. Of these numbers, 667 were also the victim of the crime. Unlike in Scotland, courts in England do not share witness details on a joint system, so it is difficult to gather more recent statistics.

To the researcher's knowledge, there are no known statistics for the USA; in 1993, Ceci and de Bruyn estimated that about 100,000 children appear in court each year across the country based upon the data they had available for New York State. This would mean that the number of child witnesses in American courts are proportionally half of that in Scottish courts. If this is accurate, it may be because testimonies from child witnesses are still treated with greater scepticism in America. The difference cannot be explained by lower crime rates in America as the rates of child sexual abuse are reported to be higher in the USA than the UK (Radford et al., 2011; Townsend & Rheingold, 2013). Given that Scottish law requires all evidence presented during trials—particularly in regard to identifications—to be corroborated, this may also help explain why children are called to courts in Scotland proportionally more often than in North America (Andrews & Lamb, 2016).

Studies reveal that 1 in 10 children have been sexually abused in the USA (Townsend & Rheingold, 2013), whilst at least 1 in 20 children have been sexually abused in the UK (Radford et al., 2011; Royal College of General Practitioners and National Society for the Prevention of Cruelty to Children, 2014). In reality, the numbers are probably higher than the studies suggest. Radford et al. (2011) found that only 66% of children sexually abused in the UK report the matter to an adult. Other studies in the USA have suggested that as little as 38% of child-victims report their sexual abuse (London, Bruck, Ceci, & Shumman, 2003; Ullman, 2007). In England and Wales, 13,700 people were convicted of sexual offences in the business year of 2016/17 (Crown Prosecution Service, 2017). About 27% of the victims in these cases were younger than 18 years old, and about 9% of them were younger than 13 years old. Children are also frequently required to testify in cases of domestic violence and intrafamilial murders. Annually in the United States, about 15.5 million children witness intimate partner violence (McDonald, Jouriles, Ramisetty-Mikler, Caetano, & Green, 2006).

## **1.2 Day care sexual abuse panic and early research**

During the second half of the 20<sup>th</sup> century, more women were working in the United States than ever before and an unprecedented number of day care centres opened to accommodate this transformation, including within people's homes and basements (Ceci & Bruck, 1995). Stemming from this was anxiety about leaving children with

strangers, followed by an enormous increase in reports of child sexual abuse and opportunities for children to provide eyewitness testimony (Ceci & Bruck, 1995). An opinion poll in 1983 showed that nearly all parents in the USA were aware of the problem of sexual abuse occurring in day care centres (Cheit, 2016). Considering that only two decades earlier, child abuse was just starting to be recognised by physicians, this was an astounding transformation in opinion (Cheit, 2016). The reports by children, as well as accusations of sexual abuse, included bizarre claims, such as Satanic rituals and being thrown into seas full of sharks (Ceci & Bruck, 1995). The situation has been compared to the witch hunts of Salem nearly three hundred years previously (Ceci & Bruck, 1995; Cheit, 2016). The day care sexual abuse panic spread throughout the United States, as well as Europe, Canada, New Zealand, and South America.

The McMartin Preschool case in California was the first in America to receive major national coverage and is often described as the longest and most expensive criminal trial in American history (Linder, 2007). The case started in 1983 when a mother, diagnosed with paranoid schizophrenia, reported to her son's paediatrician that a worker at the preschool had sexually assaulted him. This was followed by an investigation, during which hundreds of children were interviewed. Eventually, the children made 41 uncorroborated complaints (though 27 were withdrawn by the time of the preliminary hearing), and seven arrests were made. The complaints contained statements about underground tunnels beneath the school and that the children had been led to a farm where they saw animals butchered and then were raped. Additionally, the children claimed the day care workers could fly and wore witch hats when other adults were not around. Criminal trials lasted six years, and once they ended in 1990, over \$15 million of public taxes had been spent and all charges had been dropped. The interviewing practice by investigators has since been widely criticised. It has been argued that the questions used were incredibly leading, and that questions were continuously repeated until investigators managed to get the responses they desired from the children (Ceci & Bruck, 1995; Cheit, 2016; Garven, Wood, Malpass, & Shaw, 1998; Schreiber et al., 2006).

## *Chapter 1: Introduction*

Over the following decade, there were many similar day care centre cases to the McMartin case (e.g., Georgian Hills in Tennessee, 1984; Kelly Michaels in New Jersey, 1985; Craig County in Maryland, 1985; Old Cutler in Florida, 1989; Little Rascals in North Carolina, 1989), during which defendants were all found guilty of multiple accounts of sexual abuse, only then to be declared innocent and released after spending up to 26 years in prison. In the United Kingdom, there are similar prominent examples, including in Rochdale, Merseyside, and Orkney, Scotland. Typical of these cases, the Rochdale case began with a 6-year-old boy spreading stories about ghosts and zombies, only for social workers to then become convinced he had been the victim of satanic ritual abuse after a lengthy period of coercive interviews with the child. The case led to 20 children being taken from their families. It later turned out that the boy had been watching horror movies and then recounting them to his friends (Pendergrast, 2017). What all these cases have in common is that they started with vague claims or suspicions and led to serious allegations that had drastic consequences for the defendants and, in some cases, for the children themselves.

Though there were early 20<sup>th</sup> century studies on children's suggestibility (e.g., Binet, 1984, as cited in Nicholas, Collins, Gounden, & Roediger, 2011; Binet, as cited in Ceci & Bruck, 1993; Stern, 1910, as cited in Ceci & Bruck, 1993; Stern, 1939; Varendonck, 1911, as cited in Hazan, Hazan, & Goodman, 1984), the area of research became largely abandoned after the World Wars, with the exception of a small handful of studies (Burt & Gaskill, 1932; Hurlock, 1930; McConnell, 1963; Messerschmidt, 1933; Otis, 1924, as cited in Ceci & Bruck, 1993; Sherman, 1925, as cited in Ceci & Bruck, 1993). Following the day care sexual abuse cases, there was a surge of research on the topic (e.g., Loftus, Miller, & Burns, 1978; Marin, Holmes, Guth, & Kovac, 1979).

As a consequence of these studies, investigators and legal scholars gained a better understanding over how suggestive techniques can lead to false disclosures. In 2013, a nursery manager was arrested for sexually abusing four children in Fort William, Scotland (MacLennan v. HM Advocate, 2015). One of the interviews with one of the children was described by the trial judge as leading in the extreme (e.g., 'Did Mark

touch your bum?'), and related charges were subsequently dropped by the prosecution. The defendant was found guilty of the charges relating to two of the other children, who made reliable witnesses in court (one of them only 3 years old). Clearly, child eyewitnesses can be accurate when they are interviewed appropriately, and, in many instances, their testimonies are necessary for justice. However, one must have an appreciation for factors that can influence suggestibility in order to interview children appropriately.

### **1.3 An introduction to suggestibility**

According to Ceci and Bruck (1993, p. 404), suggestibility is defined as the extent to which 'encoding, storage, retrieval, and reporting of events can be influenced by a range of social and psychological factors.' Importantly, this is a broader view of suggestibility than the definition provided by Gudjonsson and Clark (1986, p. 84), that suggestibility is 'the extent to which, within a closed social interaction, people come to accept messages communicated during formal questioning, as the result of which their subsequent behavioural response is affected.' The main differences from Ceci and Bruck's (1993) definition are that, in their model of interrogative suggestibility, Gudjonsson and Clark (1986) imply suggestibility can only be unconscious and that it is a memory-based phenomenon, as opposed to a social one.

Ceci and Bruck's (1993) definition includes cases of witnesses agreeing to misleading or suggestive questions, even if the witness does not believe their statement to be true (and, thus, there has been no memory alteration) as a result of the social demands of the interview. This is an example of compliance, when people are motivated to please and/or avoid conflict with an authority figure (e.g., Gudjonsson, 2003; Milgram, 1963). In the day care sexual abuse cases, some of the children's accusations clearly stemmed from police coaxing them into making certain statements over a period of many interviews, but it is not necessarily clear if children eventually gave into these attempts as a result of social pressures (i.e., compliance) or cognitive reasons (i.e., changes in memory). Realistically, it was probably due to one reason in some instances, and due to the other reason in other instances, if not a combination of both. It is often difficult to distinguish compliance and suggestibility from one another, as someone who scores high on one often scores high on the other, as was the case with the Birmingham Six,

who were wrongfully imprisoned for bombings after four of them provided false confessions (Gudjonsson, 2003).

Though there is a suggestibility scale designed for children by Gudjonsson (1992), it consists of showing a visual narrative to a child and then asking them overtly suggestive questions, and is limited in its applicability and ecological validity (Milne, Clare, & Bull, 2002), due to not following police interview schedules with children. Ceci and Bruck (1993) claim their definition, which accounts for both cognitive and social causes of suggestibility, unlike Gudjonsson and Clark's (1986), is consistent with the legal use of suggestibility, which refers to subtle suggestions and leading questions, but also to changes in behaviour as a result of bribes or threats. Researchers in the field of children's eyewitness memory now largely use Ceci and Bruck's (1993) definition of suggestibility, agreeing that Gudjonsson's definition, while typical of the time, failed to capture important aspects of suggestibility that have arisen as a result of modern research (O'Donohue & Fanetti, 2016, p. 144). This will be the definition used in this thesis.

Ceci and Bruck's (1993) definition was used in large-scale reviews (see Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020) of impactors on suggestibility, examining both cognitive and social factors, including misinformation and misleading questions. In the 1970s and 1980s, experiments by Elizabeth Loftus revealed the impact of recall accuracy on misinformation, post-event information, and the wording of questions. For example, if an officer initially asks a witness how fast a car was going as it passed a barn, even when there was never a barn witnessed, they will later be more likely to misreport having seen a barn than someone who was simply asked how fast the car was going (Loftus, 1975). In another experiment, Loftus asked for estimates of speed in a crash, after showing footage of a simulated crash to participants, and found if she used the term 'smashed' rather than 'bumped', she would receive estimates that were on average 10 mph faster (Loftus & Palmer, 1974).

While these studies demonstrate how vulnerable memory is to post-event information, they only do so within lab settings. It may be argued that memory of a unique and/or traumatising event like a crime would be stronger, and a witness would be less likely to misremember any details. A study by Goodman, Hirschman, Hepps, and Rudy (1991) showed that children who were more distressed while getting inoculations were later able to remember the event better than children who had been less distressed and were also more resistant to misleading questions. However, a study by Morgan, Southwick, Steffian, Hazlett, and Loftus (2013) showed that even during times of high stress, such as interrogation simulations for military training, post-event information can still cause misremembrances. In this instance, military personnel underwent simulated interrogations, and afterwards were asked either open-ended or leading questions about items that were not actually in the room. Results showed that if participants underwent leading questioning, they were significantly more likely to misreport witnessing the items in the room. It seems that even during highly stressful situations our memories are susceptible to the negative influences of leading questions.

In a similar, more recent study to the one by Goodman et al. (1991), children who were more distressed while receiving an inoculation later provided significantly more errors during an interview about the event (Chae et al., 2014). However, this was mediated by the child's attachment style. Indeed, only those with high avoidant attachment styles and who were rated as being highly distressed while receiving an inoculation were significantly more likely to provide more errors. As will later be discussed, therefore, there are significant dispositional factors at play, regardless of the situational impactors and context.

While there are a number of situational impactors to explore, research has particularly made it apparent that children will often answer questions that they do not understand or know the answer to, especially if it is repeated (because they think there must, therefore, be an answer or they previously answered incorrectly; e.g., Krähenbühl & Blades, 2006), poses options (e.g., questions requiring a yes/no

response; e.g., Roberts & Cameron, 2015), or suggests that a certain answer is correct (i.e., leading questions; e.g., Lyon, 2014). In the Craig County's Day care case, for example, children were interviewed as many as 54 times until investigators got the information they wanted (Ceci & Bruck, 1995). Even now, 30 years later, some of the children from the McMartin Case still believe that they were sexually abused at day care, despite having no independent memories of any abuse (CBS Los Angeles, 2014). This shows the potentially long-lasting and tragic consequences of suggestion.

Indeed, research suggests that memories are fragile and susceptible to suggestibility. For example, Loftus and Pickrell (1995) conducted an experiment to test how easy it would be to plant false memories. They found that college students could recall vivid descriptions of times they were lost in a mall as young children, even though such an incident never occurred. They achieved this after repeatedly asking participants to recall the event with suggestive questioning. After the study was complete, the participants were amazed that their memories were false because they felt so real. Similar studies have had college students 'remember' times they nearly drowned as young children (Heaps & Nash, 2001) or times they were attacked by an animal (Porter, Yuille, & Lehman, 1999), even though the events were complete fabrications.

Furthermore, suggestibility does not solely refer to believing things occurred that in fact did not, but also believing things transpired differently than they actually did because of suggestion. For example, Pezdek (2003) revealed that about 73% of Americans 'remember' watching footage on 9/11 of both planes crashing into the World Trade Centre buildings, even though footage of the first plane crashing was not shown until the next day. When interviewing eyewitnesses about events that have been heavily publicised in the news, such as plane crashes, it is common for eyewitnesses to include details they have heard elsewhere as part of their own testimony and to mistake those details as being part of their own memories. Pynoos and Nader (1989) interviewed 113 children who attended a school that had been a victim of a sniper shooting five years previously. Even children who had been absent the day of the shooting claimed to have 'memories' of the event. It seems the children

heard about the event from their peers and then incorporated their reports into their own autobiographical memories. It appears that young children have particular difficulty with monitoring the difference between things they have imagined experiencing and things they have actually experienced (i.e., source monitoring; see below). It is important to understand that memory is fallible because of its constructive nature. Memory is not a literal reproduction of the past but relies on constructive processes that are sometimes prone to error or distortion (Schacter, 2012).

Inspired by the early findings of Loftus's experiments, Ceci, Huffman, Smith, and Loftus (1994) took a group of children whose parents said they had never got their hands caught in a mousetrap and asked them whether or not they had ever got their hands caught in a mousetrap and had to go to hospital to get it removed. They all said that they had not, then researchers came back and asked them again in a week. Again, the children all said that they had never got their hands stuck in a mousetrap. By the sixth, seventh, and eighth week of questioning, children were not only 'remembering' about times they got their hands stuck in mousetraps, but were recounting incredibly detailed accounts of how it had happened, who had taken them to hospital, and how they had got there. When confronted with the fact that it had never happened, but had all been part of an experiment, the children refused to believe it was all fake and still maintained that they really had got their hands caught in a mousetrap. Again, these findings have been linked to source-monitoring errors (see below), meaning the children were perhaps misattributing the source of the information (i.e., confusing the specific event of getting their hand caught in a mousetrap with information they had heard in previous interviews).

### **1.3.1 Source monitoring**

According to Johnson, Hashtroudi, and Lindsay (1993), source monitoring is the process of knowing when and where a memory came from, which can help explain memory inaccuracies. A source-monitoring error, therefore, is when the source of a memory is incorrectly attributed to a different source. For example, it may include being unsure if you locked the house door or only thought about locking it, or it may involve reporting that you heard something from the news when you actually learned about it from a friend. Contextual cues help identify the source of a memory (Pearse,

Powell, & Thomson, 2003). For example, you may specifically remember locking your house door that morning because you remember dropping your keys afterwards, or you may remember that you heard about a news story from a friend because you remember how they told it or by certain words they used. The ability to accurately identify the source of a memory is something that develops during childhood (Foley, 2014; Ghetti & Angelini, 2008).

Since people retrieve knowledge from a wide range of sources, it can be difficult to correctly identify the exact source of certain memories. In a study by Poole and Lindsay (2001), over 25% of 7-year-old participants and 9% of 8-year-old participants reported during free recall that a male assistant had touched them, even though this touching had only been described in a story they had read to them. Similarly, other studies have shown that if children are questioned about a particular event (e.g., a trip to Disneyland) and those questions trigger memories of other events (e.g., conversations with friends and family about their own separate trips to Disneyland), the children are at risk of incorrectly including details from those conversations into their own narrative (e.g., Price, Connolly, & Gordon, 2015; Principe, Greenhoot, & Ceci, 2014).

#### **1.4 Suggestibility and individual differences**

**Table 1.1** contains interview extracts from a study by Clarke-Stewart, Malloy, and Allhusen (2004), in which children lay beside a confederate called Patrick in a Ghostbusters themed house. Though Patrick did not touch either of the children, both Child A and Child B responded very differently to the misleading questions about touching. Child A was highly suggestible, and he falsely reported to the interviewer during questioning that Patrick had touched him. Child B, however, resisted all the interviewer's suggestions. Interestingly, Child A was older than Child B, even though older children are usually less suggestible (for a review, see Bruck & Melnyk, 2004).

Clearly, there were individual differences at play, other than age, which influenced whether or not the children went along with the interviewer or resisted. Some children will agree immediately with misleading questions by interviewers, and some will continuously deny that a suggested event ever occurred. In their review of 69

published and unpublished studies involving 4,848 children, Bruck and Melnyk (2004) examined the impacts of various cognitive and personality factors on suggestibility, including, but not limited to, attachment style, creativity, intelligence, and temperament. In 2020, Klemfuss and Olaguez's updated review uncovered a further 55 studies, including 6,455 children, investigating similar dispositional factors. In both reviews, intelligence and language ability were the factors most frequently reported by studies to be significant factors.

**Table 1.1.**

*Interview extracts from Clarke-Stewart et al. (2004).*

Interview Extract– Child A	Interview Extract – Child B
Interviewer: Remember when you went in the Ghostbusters House and Patrick asked you to lie down on a blanket next to him?	Interviewer: Remember when you went inside the Ghostbusters house and Patrick asked you to lie down on the blanket next to him?
Child A: [Pause.]	Child B: [Nods.]
Interviewer: Do you remember that?	Interviewer: What was that like?
Child A: [Nods slightly.]	Child B: Like being in bed.
Interviewer: What was that like?	Interviewer: Like being in bed? Did Patrick touch you when you were in the Ghostbusters House lying down?
Child A: Um, I liked that pretty much.	Child B: [Shakes head.]
Interviewer: Well, did Patrick touch you when you were in the Ghostbusters House lying down?	Interviewer: No? Are you sure? Now think hard. Other kids have said that Patrick touched them when they lay down in the Ghostbusters House. Do you remember now?
Child A: [Shakes head.]	Child B: He didn't touch me.
Interviewer: No? Are you sure? Now think hard. Other kids have said that Patrick touched them. Do you remember now?	Interviewer: He didn't touch you? How do you think it made you feel? Did it make you feel yucky?
Child A: [Nods.] He did.	Child B: [Shakes head.]
Interviewer: He did?	Interviewer: No? Are you sure it didn't make you feel uncomfortable? You can tell me about it. Did it make you feel yucky?
Child A: [Nods.]	Child B: [Shakes head.] Uh-uh.
Interviewer: How did that make you feel?	Interviewer: No? Other kids have kind of worried about Patrick touching them. They
Child A: Bad.	
Interviewer: Bad?	
Child A: [Nods.]	
Interviewer: Yeah? Tell me about that, when he touched you. What was that like?	
Child A: Um, it was like painful.	

Interviewer: It was painful? Where did he touch you?

Child A: [Puts hand on side.] Um, I think right here.

Interviewer: Right on your side there?

Child A: [Nods.]

Interviewer: How did he touch you?

Child A: Um, ah, he moved over and touched me.

Interviewer: He moved over and touched you?

Well, I bet that made you feel uncomfortable when Patrick touched you in the Ghostbusters House. Now think about when Patrick lifted you up so you could get your prize in the pink piñata, did he touch you then?

Child A: [Nods.]

Interviewer: Where did he touch you?

Child A: [Puts hands on sides.] Um, on my sides.

Interviewer: On your sides? And did he touch your bottom?

Child A: [Shakes head.] Nope.

Interviewer: No? Didn't he touch you on your behind?

Child A: Uh, yeah.

Interviewer: Yeah? Other kids have told me that they didn't like it when Patrick touched them. Did he ever touch you in a way that you didn't like?

Child A: [Nods.]

Interviewer: Yeah? Tell me about that.

Child A: Um, it made me feel bad and painful.

Interviewer: It made you feel bad and it was painful?

Child A: [Nods.]

Interviewer: I'm sorry to hear that. I bet that made you feel yucky, am I right?

Child A: [Nods.]

---

didn't feel very comfortable at all. Didn't it make you feel yucky when Patrick touched you in the Ghostbusters House?

Child B: He didn't touch me.

Interviewer: He didn't touch you? Well think about when he lifted you up so you could reach the prize in the pink piñata. Did he touch you then?

Child B: [Shakes head.]

Interviewer: Did he touch you on your bottom?

Child B: [Shakes head.] No.

Interviewer: No? Didn't he touch you on your behind?

Child B: No.

Interviewer: Are you sure?

Child B: [Nods.]

Interviewer: Some other kids told me that they didn't like it when Patrick touched them. Did he ever touch you in a way that you didn't like?

Child B: [Shakes head.] Uh-uh.

Interviewer: No? Are you sure?

Child B: [Nods.]

Interviewer: Yes? I bet it made you feel kind of yucky when he touched you, am I right?

Child B: No.

### **1.4.1 Cognitive and personality factors**

Other than the effects of age, cognitive factors are the most widely studied predictors of individual suggestibility and eyewitness performance. Since eyewitness performance involves providing memory reports of witnessed events, it seems logical that certain cognitive factors would have a part to play in how accurate those reports are. Generally, children with higher IQ scores and better memory are less suggestible to misleading questions and remember more accurate descriptors than children with lower IQ scores and poorer memories (e.g., Bettenay, Ridley, Henry, & Crane, 2015; Danielsdottir, Sigurgeirsdottir, Einarisdottir, & Haraldsson, 1993; Endres, Poggenpohl, & Erben, 1999; Geddie, Fradin, & Beer, 2000; Gignac & Powell, 2006; Henry & Gudjonsson, 2004; Hurlock, 1932; McFarlane, Powell, & Dudgeon, 2002; Roebbers & Schneider, 2001; Singh & Gudjonsson, 1992).

The research suggests that intelligence is particularly a significant influence of eyewitness performance when it comes to children with intellectual disabilities (see London, Henry, Conradt, & Corser, 2013 for a review) or when it comes to typically-developing children under 8 years old. In the case of typically-developing children aged 8 years and over, intelligence is less of a significant influence and some researchers have not found any significant relationship between intelligence and suggestibility when it comes to this age group (for a review, see Bruck & Melnyk, 2004). This is likely because by the age of 8 years, even children with an average IQ are mentally mature enough that they are generally resistant to suggestive questions. In support of this, eyewitness performance typically increases, and suggestibility decreases, as children experience developmental progressions in cognitive inhibition, executive functions, and language abilities (e.g., Alexander et al., 2002; Clarke-Stewart et al., 2004; Danielsdottir et al., 1993; Poole, Dickinson, Brubacher, Liberty, & Kaake, 2014). The impact of developmental factors on eyewitness performance will be further explored in the Literature Review.

Although less researched than cognitive factors, research has also suggested that temperament characteristics or behavioural styles may impact one's understanding, interpretation, and processing of an event, as well as one's ability to resist suggestive

questions when providing eyewitness testimony (Shapiro, Blackford, & Chen, 2005). Ornstein, Shapiro, Clubb, Follmer, and Baker-Ward (1997) theorised that particular elements of temperament affect the perception of eyewitnesses as they witness events take place (activity level, emotionality and persistence), while other elements (adaptability, shyness and distractibility) impact on their performance during forensic interviews. Temperament is a precursor of personality and refers to behavioural traits that explain how one behaves, thought to have a genetic origin, or at least be determined early in life (Braungart, Fulker, & Plomin, 1992; Buss & Plomin, 2014; Martin, 1988; McDevitt & Carey, 1978; Mervielde et al., 2005; Posner & Rothbart, 2000).

Martin's (1988) model of temperament is the most commonly used in the eyewitness psychology literature (for a review, see Bruck & Melnyk, 2004), due to the fact that the dimensions have been theoretically linked to eyewitness performance (Ornstein et al., 1997). These dimensions are activity, adaptability, approach/withdrawal, ease of management through distraction/distractibility, emotional intensity, and persistence. The common use of this model in eyewitness psychology is despite the fact that Rothbart and Ahadi's (1994) model (with three primary dimensions: surgency, negative affect, and effortful control) is more frequently endorsed by developmental psychologists due to the conceptual and empirical connection between these three dimensions in childhood and the Big Five (currently, the most popular personality model) in adulthood (Shiner & DeYoung, 2013). After all, if temperament is a genetic precursor to personality, as often defined (Boeree, 2006; Hofstee, 1991; Martin & Bridger, 1999; Rothbart, 2012), it should connect clearly to personality later in life. The temperament dimensions in Martin's model, however, may equally correspond clearly to the dimensions of the Big Five, but there has been less effort to do so. This will be explored in the Literature Review. One aspect of temperament that may impact eyewitness performance is shyness, labelled by Martin (1988) as approach/withdrawal.

Shyness can be defined as 'a psychological state that causes feelings of discomfort, leading to avoidance of social contact' (Afshan, Askari, & Manickam, 2015, p. 1),

especially with unfamiliar people (Cheek, Melchior, & Carpentieri, 1986, p. 115). According to Crozier (2000), shy individuals lack confidence and are more likely to doubt themselves and to rely on others for information. Therefore, the extent to which a child is shy may significantly impact on their ability to resist suggestive questioning. In adult samples, shyness is predicted by high neuroticism and low extraversion (e.g., Kwiatkowska & Rogoza, 2019; Paulhus & Trapnell, 1998), leading to the suggestion that shyness is a combination of the two dimensions (e.g., Briggs, 1988; Geen, 1986, as cited in Paulhus & Trapnell, 1998), but still distinct.

In a study of 112 undergraduates by Ward and Loftus (1985), introverts were more suggestible to accept post-event information than extroverts. Chen and Shapiro (2000, as cited in Purdy, 2001) also found preschool and elementary children who were more outgoing provided more information when interviewed about a witnessed event using general questions. However, when suggestive rather than general questions were used, shyness had no impact on the accuracy, in this particular study. Similarly, in studies by Roebbers and Schneider (2001) and Chae and Ceci (2005) there was a significant, negative relationship between shyness and memory recall. Benedan, Zajac, McFarlane, and Powell (2020) also reported a negative relationship between shyness and memory recall, but it did not reach statistical significance. Therefore, shyness impacted the quantity of information recalled in these studies, but not necessarily the quality.

In Gilstrap and Papierno's (2004) study, however, interviewers were found to have asked a higher proportion of leading questions to shyer children, who were found to be more suggestible to those questions than less shy children. In this study, therefore, shyness did impact the quality of responses. Similarly, Burgwyn-Bailes et al. (2001) found a significant, positive relationship between social avoidance and suggestibility. It may be the case that shyer children are less likely to feel comfortable enough during a forensic interview to volunteer substantial information and are also more likely to agree with suggestions from an interviewer, out of fear of disagreeing and the interviewer then thinking poorly of them, given that, as discussed, fear of social

judgement is an important element of shyness (Afshan et al., 2015), but the association between shyness and quality of responses seems to be less consistent than the finding of an association between shyness and quantity of responses. This may partly be due to the differences in experimental design between studies. For example, Gilstrap and Papierno's (2004) found that interviewers asked a higher proportion of leading questions to shy children, perhaps because they were struggling to get straight answers. It is not clear, therefore, that the authors would have found shy children to be more suggestible if all children had been asked the exact same questions.

Recently, a study by Johnston, Benedan, Brubacher, and Powell (2021) investigated an association between children's propensity to disclose adult transgressions during free recall and three dimensions of temperament (social flexibility, reactivity, and task orientation), finding social flexibility was the only significant predictor of transgressions disclosed. In this case, children who were less flexible in social situations disclosed less transgressions than other children. This suggests that because of their temperament or personality dimensions, some witnesses may be more vulnerable to suggestive questions, a vulnerability which may impact their interview performance. Due to their perceived lack of confidence (Crozier, 2000), high levels of shyness may also impact how children are perceived as witnesses according to jurors.

## **1.5 Juror perceptions**

Eyewitness testimonies provide jurors with information that would not otherwise be available and, thus, are incredibly important to criminal trials. That said, memory is fallible and eyewitness testimony is the leading cause of wrongful arrests. According to the Innocence Project (2020), as mentioned earlier, about 71% of 360 cases were due to misidentifications made by eyewitnesses. In response to this, 19 American states have implemented jury warnings in regard to the possible unreliability of eyewitness reports (Innocence Project, 2020). In some cases, these jury warnings have replaced testimony from expert witnesses on eyewitness memory (e.g., State of Utah v. Guard, 2015). Similarly, in the United Kingdom, judges have recently decided in cases that lawyers, as well as the judges themselves, are now educated enough on eyewitness

memory that they can instruct jurors, without opinions being necessary from psychologists (e.g., MacLennan v. HM Advocate, 2015).

Though it is important in any case, if a child is the only eyewitness, it is especially vital to understand how jurors perceive their testimony. A number of factors may affect jurors' perceptions of the credibility of a child witness. For example, Golding, Fryman, Marsil, and Yozwiak (2003) found that a crying child witness led to more guilty verdicts from jurors than a child who did not cry while giving testimony. In some cases, child witnesses have been viewed as incredibly honest and as having no reason to lie (e.g., Nunez, Kehn, & Wright, 2011). In other cases, they have been perceived as having a poor memory and to be less reliable than adult witnesses (e.g., Bottoms & Goodman, 1994). The decision making of jurors is ultimately influenced by a number of factors, including the type of crime and the type of witness that the child is (i.e., bystander-witness or victim-witness; Holcomb & Jacquin, 2007). Furthermore, there are factors relating to the jurors themselves, such as how empathetic one generally is (Haegerich and Bottoms (2000)).

## **1.6 Summary**

Clearly, child eyewitness testimony is something that should be of significant interest, given how frequently it occurs. Lawyers, researchers, and even some jurors are now aware of suggestibility and the effects that certain questioning types can have upon eyewitness testimony. However, it is still not clear whether children with certain temperament qualities (e.g., high shyness) respond to particular questioning types differently to children without those temperament qualities, as well as if those children may be perceived differently by jurors. This thesis will attempt to bridge some of that gap, first by reviewing the literature, and then in three separate studies, followed by a general discussion. The first study will assess the reliability and validity of a new, self-report tool for measuring temperament in young children. Following this, the second study will examine the influence of interview techniques and temperament on children's eyewitness performance. The final study will investigate the impact of question types and shyness on perceptions of child witnesses according to mock-jurors.

## **Chapter 2: Literature review**

---

This chapter will discuss the literature surrounding the reliability and credibility of children as eyewitnesses, exploring practices and influences that can increase or decrease accuracy. The chapter begins by outlining memory mechanisms and general developmental factors that influence the capability of witnesses. This is followed by an exploration of the describing and identifying capabilities of young children when performing as witnesses. The chapter then discusses external factors (i.e., those relating to the interviewer) that influence children's suggestibility and eyewitness performance, and outlines internal factors (i.e., those relating to the child). Lastly, jurors' perceptions of children as eyewitnesses will be explored.

### **2.1 Memory and development**

Before one can understand a child's capabilities as an eyewitness, one must first have insight into the memory mechanisms that account for these capabilities. There are three stages to memory (Pozzulo, 2017). First, is the process of encoding, which is when information is initially perceived and converted into memory. A person may be aware they are encoding new information, or they may not be (e.g., Yang, Cao, Xu, & Chen, 2012). Second, memory must be stored. If the memory is rehearsed enough, it may be stored as a long-term memory, rather than only lasting as a short-term memory (Cowan, 2008). The final stage is retrieval, when memories are located and remembered by the individual. Generally, the process of remembering is viewed as constructive, rather than reproductive, meaning our minds do not accurately record everything like video recorders and we cannot just play memories back in our minds exactly as they happened (Patihisa et al., 2013). Rather, when retrieval occurs, our memories must be reconstructed, and this process is vulnerable to a number of influences, including leading questions and post-event information (e.g., Ceci & Bruck, 1995; Schmidt, 2004).

Not only may we alter memories of actual experiences, but we may create completely false memories of fabricated events, such as seen by Loftus's *Lost in the Mall* experiment (see Introduction; Loftus & Pickrell, 1995). Consequently, memories are far less accurate than most people realise (see Hopwood, 2016, for a collection of real-life

false memories). False memories can be triggered by stories from family members or even from family photographs (e.g., Wade, Garry, Read, & Lindsay, 2002). Essentially, if we think an event happened one way or we are told that it occurred that way and we play it back that way to ourselves often enough, then it is easy to believe that is exactly what happened even if it is not.

### **2.1.1 Externally-driven and internally-driven false memories**

As mentioned previously, adults have historically been viewed as more reliable witnesses than children, and this is reinforced by adults being less prone to developing false memories during research studies (e.g., Ghetti, Qin, & Goodman, 2002; Loftus & Davies, 1984). In a review of studies published between 1979 and 1992, Ceci and Bruck (1993) uncovered that, when comparing the suggestibility of children versus adults, the formation of false memories decreased with age in 83% of the studies. However, more recently, research is beginning to show greater appreciation for the notion that there are different types of false memories (e.g., Brackmann, Otgaar, Sauerland, & Merckelbach, 2015). In this case, the research suggests that children are in fact less susceptible to certain false memories.

The false memories measured in the previously mentioned experimental studies are typically those that can be classed as externally driven (or suggestion-induced false memories). The experiments tend to follow the format of Loftus's misinformation paradigm (Loftus et al., 1978). First, the participant is asked to witness an event (e.g., a video of a theft, a live event). Second, a piece of misinformation is suggested to them about the event they just witnessed. Third, they are asked to report what they remember from the event. Results from studies show that a significant number of people incorrectly incorporate the piece of misinformation supplied to them into their reports, especially in the case of child participants. Moreover, as already discussed, false memories can be implanted for entirely fictitious childhood events, such as being lost in the mall (Loftus & Pickrell, 1995), being attacked by a bear (Porter et al., 1999), and nearly drowning (Heaps & Nash, 2001). Other studies have successfully implanted false memories of hot air balloon rides that never happened (Wade et al., 2002), and even alien abductions (Otgaar, Candel, Merckelbach, & Wade, 2009). In these cases, also, children were more susceptible to false memories than adults.

On the other hand, internally-driven false memories (or spontaneous false memories) are those that occur when a person makes assumptions over what happened or makes an error due to relying on gist information (Brackmann, Otgaar, Sauerland, & Jelicic, 2016). This difference can be further understood via the fuzzy trace theory (Brainerd & Reyna, 2005), which was intended to demonstrate how eyewitnesses encode information during the event of witnessing a crime and how those memories may be retrieved at a later point in time. This theory suggests that memories are encoded, stored, and retrieved using two processes that work in parallel with each other (i.e., the two processes occur at the same time, but work differently). First, verbatim traces capture specific details of an event, such as the make, model, and colour of a gun. Meanwhile, gist traces capture only the general information from an event, such as the mere presence of a weapon. While both traces stem from the same event, the way they are encoded, stored, and retrieved are different.

Gist information relies on prior knowledge and so is viewed as less accurate than verbatim information. For example, a famous study by Bartlett (1932) had participants read an old Native American story called *War of the Ghosts* and then later asked those participants to retell what they had read. Bartlett found that the participants reported only the gist of the story, leaving out various details, and that they altered some parts of the story to be in harmony with their own view of the world, rather than that of the Native Americans. As previously noted, gist information relies on previous experiences and prior knowledge, thus, adults are more likely to use this process than children. For example, an adult might be more likely than a child to misremember locking their front door. Presumably, the adult does this every day and so it may be harder to separate one specific memory of locking their door from the countless other memories of locking their door.

Research suggests that adults are more likely than children to be vulnerable to internally-driven false memories, as a result of being more likely to rely upon gist information (e.g., Brainerd, Reyna, & Ceci, 2008). This is perhaps due to the fact that

adults are more likely to have a schema for a given event due to their greater life experience. According to Schank and Abelson's (1977) schema theory, a schema is a unit of knowledge for a particular subject or item based upon past experience, such as what happens and the order in which things happen when we go to the supermarket or visit a restaurant. Even externally-driven false memories can increase with age when the participants use gist information, rather than verbatim. An experiment by Otgaar, Howe, Smeets, Brackmann, and Fissette (2014, as cited in Brackmann et al., 2015) showed a video of a robbery to young children, older children, and adults. Some related details of the robbery, such as the culprit's weapon, were left out of the video. When presented with misinformation, adults and older children were more likely than younger children to accept having seen the absent details. Most likely, this is because adults and older children are more likely to have a schema for a bank robbery that involves a weapon (e.g., from greater exposure to bank robberies portrayed on TV and in movies).

In a famous study by Roediger and McDermott (1995), adult participants were instructed to study a list of words (e.g., dream, pillow, blanket, bed) that all centred around a theme (e.g., sleep). When asked to reproduce the list of words, a significant number of participants incorrectly reported having seen the theme on the list of words, even though it was not present, suggesting the association formed a false memory. The implication is that verbatim information is item-specific (i.e., the exact words on the list), while gist information relies on knowledge and associations (i.e., the fact that all the words on the list were related to 'sleep'). When verbatim information is not available to be retrieved from memory, we rely on gist information, which can lead to an increase in false memories. When this study has been replicated with child participants, they are less likely than adults to misreport having seen the theme on the list of words (Brainerd, Reyna, & Zember, 2011), perhaps because they are less likely to make these associations.

Therefore, it is not correct to view testimonies by adults as always being more accurate than those by children. Under certain circumstances, children may be more

reliable witnesses, and less prone to false memories, than adults. Ironically, this is particularly likely to be the case in situations that are familiar to an adult, due to the fact they will already have established schemas surrounding the situation. In a recent trial, the likelihood of a young girl's testimony being tainted due to false memories was debated by two opposing expert witnesses, in a case where there had been no external influences (Brackmann et al., 2016). The jury were ultimately convinced by her credibility and convicted based upon her testimony, after hearing from an expert that there was no evidence of external suggestion (e.g., there were no leading questions asked and there was no delay between the event and the interview), and that research suggests children are not prone to internally-drive false memories (Brainerd et al., 2011).

Overall, the differences between internally and externally caused false memories are still being understood and require further research, but the results seem promising in regard to the credibility of children as eyewitnesses. Still, it is important to realise that every case is unique, and a range of factors must be considered around each young witness and their reliability.

### **2.1.2 Developmental differences in memory**

One theory as to why adults are generally considered more accurate eyewitnesses than children is that they have stronger memory traces (Cowan, 2001), meaning there are differences in how adults and children encode information such as a suspect's face. Indeed, there is a fairly large number of studies indicating that memory improves to a point with age (e.g., Chance & Goldstein, 1984; Fitzgerald & Price, 2015; Schneider & Pressley, 1989). The research suggests that children have a weaker ability to encode, store, retrieve, and source monitor information (e.g., Davis & Loftus, 2005). Since suggestibility is generally linked to the strength of memory, this also might explain why children often are more suggestible than adults. For example, if a child does not have a clear memory of an event that they are being questioned about (because their encoding and storing abilities are weaker than that of adults), then it will be fairly easy to suggest something happened that actually did not. Similarly, if a child's memory trace is weaker then they may be more easily led astray during target-absent line-ups (i.e., when the line-up does not contain the perpetrator) when there is an individual

within the line-up that looks similar to the actual perpetrator (e.g., Pozzulo & Dempsey, 2006).

### *2.1.2.1 Development of autobiographical memory*

Autobiographical memory 'integrates memories of past experiences into an overarching life narrative' (Fivush, 2011, p. 1). While even young children can provide coherent narratives (Lamb et al., 2003), the complexity and length of the narrative generally increases as one ages (Fivush, 1997). One reason for this is that the development of self-concepts affects the emergence of autobiographical memory as it requires self-awareness of having experienced the past (Howe, Courage, & Edison, 2003; Tulving, 2002). One starts to develop a sense of self around 2 years old (Howe et al., 2003). Furthermore, children learn how to structure narratives over time as a result of experience from participating in conversations (Nelson & Fivush, 2004). By around age 4, children can provide just as much information in response to open-ended questions as older children, though they may require further prompts due to brevity in original responses (Lamb et al., 2003).

## **2.1.3 Other age-related differences in memory recall**

### *2.1.3.1 Knowledge differences*

While there may be some areas in which children have greater knowledge than adults, such as cartoon shows, knowledge generally increases with age. Just like if a memory trace is stronger, it is harder to be led to suggestion, the same is true when one has greater knowledge. For example, Elischberger (2005) tested 5- and 6-year-old children's resistance to suggestive questioning after only half of the children received factual information about the topic. Children who received the factual information before being exposed to suggestive questioning were significantly more resistant. Therefore, due to their poorer knowledge base, it may be easier to implant suggestions in children than in adults, especially when it comes to topics that children know very little about.

### *2.1.3.2 Language differences*

Like memory and knowledge, one's language ability also develops with age. Adults may make better eyewitnesses than children because they have more words to express their thoughts and are more successfully able to articulate what they remember about an event. Since the number of descriptors that eyewitnesses tend to use increases

with age (e.g., Pozzulo & Warren, 2003), this seems to support this theory. In the day care sexual abuse hysteria cases of the 1980s and early 1990s, examples of suggestibility involved instances of children answering questions they had not correctly understood or investigators misinterpreting the statements of the children (Ceci & Bruck, 1995). In a random sample of 66 court transcripts, misunderstandings were made between a lawyer and a child witness (between the ages of 5 and 12 years) in two thirds of the transcripts (Sullivan, St. George, & Stolzenberg, 2020). A study by Goodman and Aman (1990) asked 3-year-old children whether or not a male experimenter had touched their private parts. Many answered positively, even though no participants had been touched by the experimenter. Further inspection revealed that the children simply did not understand what was meant by 'private parts'. This shows the importance of checking children's level of knowledge and using words they are familiar with.

Although by the age of about 6 years, children typically have large vocabularies (as many as 6-8 thousand words, according to Clark & Clark, 1977, as cited in Lamb, Malloy, & La Rooy, 2011), they are still less descriptive than adults (e.g., Pozzulo & Warren, 2003). Furthermore, some language concepts take longer to develop than others, such as temporal understanding. Even adults may struggle to report exactly when an event occurred without using significant landmarks (e.g., birthdays, holidays) to guide their memories. Research shows that significant development occurs between years 4 and 6 in regard to the understanding of temporal words (e.g., before, after), but the development of this understanding continues until about ten years old (e.g., Friedman & Lyon, 2005; Orbach & Lamb, 2007). Therefore, younger children may not successfully be able to communicate certain details during forensic interviews. Of course, as mentioned when discussing the development of autobiographical memory, there is a clear connection to the development of language; one must first develop a sense of self, then the ability to describe one's self, before forming autobiographical memories (Howe et al., 2003; Tulving, 2002).

In England and Wales, communication specialists called Registered Intermediaries are available to assist police officers and lawyers who are interviewing vulnerable witnesses, including children. This is to help make sure that the witness understands the questions they are being asked. Research suggests that children are more accurate as witnesses during identifications when receiving assistance from a Registered Intermediary (Wilcock et al., 2018) and are also perceived as having more credibility by jurors (Collins, Harker, & Antonopoulos, 2016). Therefore, this is one method to overcome difficulties stemming from differences between witnesses in language ability. There is also the possibility for the Registered Intermediary to help the witness overcome other difficulties. For example, one Registered Intermediary has requested an interview be held in the morning to accommodate a child with ADHD and has asked for questions to be repeated on the behalf of a child witness with a speech impediment and hearing difficulties (Thomas, 2018).

### *2.1.3.3 Theory of mind*

One aspect of cognitive development that is associated with memory recall and suggestibility is theory of mind (for reviews, see Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020). This refers to a child's growing ability to understand that other people have different thoughts and beliefs than they do, typically beginning to develop between years 3 and 6 (de Villiers & de Villiers, 2017). Research has found conflicting results in regard to the connection between theory of mind and eyewitness performance (e.g., Bright-Paul, Jarrold, & Wright, 2008; Karpinski & Scullin, 2009; Klemfuss, Rush, & Quas, 2016; Melinder, Endestad, & Magnussen, 2006). On one hand, once children develop theory of mind, they may understand they hold different knowledge of what occurred than an interviewer, thereby reporting more information and rejecting any misleading questions by the interviewer. However, if they understand the interviewer holds different views than they do based upon suggestive questions by the interviewer, they may assume these views are more valid than their own and so adjust their testimony accordingly (Waterhouse, 2016). Klemfuss and Olaguez's (2020, p. 166) review concludes that theory of mind 'may be related to children's resistance to suggestion...but that the association with interrogative suggestibility is tenuous.'

*2.1.3.4 Own-age bias*

Usually, a young eyewitness will be significantly younger than the perpetrators they are describing and identifying. Just as people are more likely to correctly identify perpetrators who are of the same race (Chiroro, Tredoux, Radaelli, & Meissner, 2008; Valentine & Endo, 1992; Young, Hugenberg, Bernstein, & Sacco, 2012), witnesses are also more likely to correctly identify perpetrators who are of a similar age (Anastasi & Rhodes, 2005; Stein, End, & Sterzer, 2014; Wiese, 2011). The most likely explanation for this is that people have more experiences with others in their own age groups, and, thus, are better at recognising those from that age group (Anastasi & Rhodes, 2005). The degree to which one has experiences with those within their own age group therefore has a significant influence. For example, college students, school children, and older adults recruited from nursing homes are more likely to have higher levels of daily exposure to people of their own age, and, thus, have a higher degree of own-age bias than older adults recruited from the local community or adults who work primarily with children or the elderly (for a review, see Rhodes & Anastasi, 2012). Therefore, children may perform poorer than adults at identifying adult targets in line-ups due to spending most of their time around other children, especially if they attend nursery or school (Anastasi & Rhodes, 2005).

*2.1.3.5 Social influences*

Undergoing an interview by police and then identifying a suspect from a line-up also contains social influences which may impact a child differently than they would an adult. Children are less likely to give an 'I don't know' response than adults (Pozzulo & Lindsay, 1997), and are more likely to feel that the mere presence of a line-up demands they choose someone (Ceci, Ross, & Toglia, 1987). This is supported by the fact that children are more likely than adults to make errors during target-absent line-ups, but not during target-present line-ups (e.g., for a review, see Havard, 2014; Pozzulo & Dempsey, 2006). These findings are still true even when regarding a subject highly familiar to a child. For example, Pozzulo, Dempsey, Bruer, and Sheahan (2012) asked both children (4 to 7 years old) and adults to select familiar cartoon characters from a line-up. Children produced almost 100% accuracy rates during target-present line-ups but were significantly lower in accuracy than adults during target-absent line-ups. Since these characters were clearly familiar to the children, their incorrect responses can only be explained by social influences, rather than factors stemming

from memory, knowledge, or language abilities. These findings can be applied to descriptive eyewitness studies, too. For example, if child witnesses feel compelled to select a target in a target-absent line-up, they may also feel compelled to answer a question they do not know the answer to (Ceci & Bruck, 1995; Hughes & Grieve, 1980; Pratt, 1990; Waterman et al., 2000). This will be explored in more detail later in this chapter, when explaining the various question types that witnesses are asked. Individual differences (e.g., shyness) within children may help to explain why some children are affected more significantly by these social influences than other children.

#### *2.1.3.6 Other influences*

Stress and arousal are general factors likely to be present in all eyewitnesses. Of course, stress is likely to be higher in some cases than others, such as when the witness is also the victim. In some cases, stress and arousal appear to increase the accuracy of eyewitness reports (e.g., 'I was so scared, I'll never forget that face'), and in others it appears to decrease accuracy (e.g., 'I was so scared, it's all just a blur'). For example, when a weapon is present, eyewitnesses tend to divert their attention to the weapon (i.e., weapon focus), and, thus, remember fewer descriptors about the perpetrator (Ross, Read, & Toglia, 1994). The Yerkes-Dodson law (Yerkes & Dodson, 1908) suggests that a moderate level of stress and arousal produces the optimum levels of accuracy in memory. It would appear that a certain level of arousal is required for an individual to really pay attention to details, but the individual becomes overwhelmed once a certain level of arousal is reached. Little research has investigated whether or not these levels change depending on age. It may be that younger eyewitnesses have a lower optimal threshold for arousal than adult eyewitnesses. Research, however, suggests that age has no effect on weapon focus (i.e., the memory of children is negatively influenced by the presence of a weapon to the same extent as that of an adult, not more so; Davies, Smith, & Blincoe, 2008). This does not support the idea that children may become more easily overwhelmed than adults, but research should still investigate further causes of emotional arousal to see if there are other differences between adults and children in terms of responses.

#### *2.1.3.7 Children's ability to lie*

As already mentioned, children rarely make false disclosures of sexual abuse of their own accord. When they do make false accusations, it is largely because they mistakenly believe the abuse was real as a result of suggestion and inappropriate

questioning (Loftus, 2005; Otgaar et al., 2009). In a small number of cases, however, children have intentionally made false disclosures of sexual abuse, either due to their own reasons for lying or because they have been coached to do so (O'Donohue et al., 2018). Children may also lie in order to receive a reward or avoid punishment (Talwar & Crossman, 2012). Children as young as 2 or 3 years old are capable of lying but are unable to take into account another person's knowledge of the situation (Talwar & Lee, 2008). Therefore, young children are unlikely to successfully stick to a lie during further questioning (Talwar & Lee, 2002). For example, 3-year-old children are more likely to deny peeking at a toy after being told not to than 2-year-olds, but both will accidentally reveal having peeked at the toy after being asked what type of toy it was (Evans & Lee, 2013). This has led to researchers arguing that theory of mind development is crucial to lying successfully (Talwar & Lee, 2008). Evidence shows that children are significantly more likely to uphold a convincing lie by the age of 6 or 7 years (Talwar & Lee, 2002). By this point, children are more likely to be able to understand their own and the other person's mental states (Premack & Woodruff, 1978; Wimmer & Perner, 1983).

#### **2.1.4 Summary: Memory**

This section began by defining the mechanisms of memory and then went on to explain developmental factors that generally influence the capability of children as eyewitnesses, including memory, knowledge, language ability, theory of mind, and threshold for arousal. It seems logical that memory and knowledge particularly would be linked to eyewitness performance. For example, if one is asked about their birthday, their response will likely be accurate no matter how misleading the question, since the memory trace of one's own birthday is just too strong to be overwritten. However, if one was to be asked about a friend's birthday, which has not been reinforced in one's memory to the same extent, then it would likely be significantly easier to have that person incorrectly report the wrong day after planting enough suggestion. Generally, it has been assumed that children are far more likely to generate false memories than adults, but this is because most of the research has focused on suggestion-induced false memories, rather than spontaneous false memories. Recent evidence shows that perhaps children are less at risk to spontaneous false memories than adults are, but this still needs to be further explored.

## **2.2 Children's describing and identifying capabilities**

Now that developmental factors related to memory are understood, the capability of children as witnesses can be discussed. This section will assess children's capability as witnesses by exploring everyday conversations with children, children's responses to questions concerning forensically relevant material, and children's ability to recall descriptors about suspects they may have witnessed.

### **2.2.1 Children answering nonsensical questions**

Research suggests that even when children do not understand a question, they will sometimes still provide an answer (Hughes & Grieve, 1980; Pratt, 1990). The likely reason for this is that children often respond in a way that they think an adult might want. They are generally compliant and trust that adults know more than them or their peers (Ceci & Bruck, 1995). Often when the same question is repeated, a child will change their answer, probably fearing that they must not have given the correct response the first time and so should change it to please the adult. Hughes and Grieve (1980) and Pratt (1990) designed questions that were intended to be unanswerable (e.g., 'Is red heavier than yellow?'). They found that the large majority (as high as 90%) of young children (between 5 and 7 years old) would provide answers to these bizarre questions. These studies have since been frequently cited in the courtroom, showcasing that a child's answer to a question is not necessarily an indication that they understand the question, thus decreasing their credibility as a witness.

There are some problems with these studies, however. First, they only used closed-ended questions (e.g., questions requiring a 'yes' or 'no' response), yet it has been assumed that the same finding would be true for open-ended questions without any direct evidence to believe so. Second, when parents were asked to judge the questions of Pratt's (1990) study as sensible or silly, 25% of the adults judged the silly questions to be sensible. Waterman et al. (2000) found that children would indeed answer nonsensical questions, but only ones that required a 'yes' or 'no' response. They were unlikely to answer bizarre questions that were open-ended or that required a response other than 'yes' or 'no'. Therefore, previous studies have likely exaggerated children's tendency to answer nonsensical questions and this exaggeration has probably been incorrectly reported to judges and juries during trials. Instead, the results of Waterman

et al. (2000) suggest that children are only likely to answer nonsensical questions if they call for a 'yes' or 'no' answer. Furthermore, Waterman and Blades (2011) found that pre-interview instructions could be used to increase the number of appropriate 'I don't know' responses from young children to unanswerable questions. The instructions had no effect on the number of correct responses to answerable questions.

### **2.2.2 Forensically-relevant conversations with children**

When considering whether or not children are capable enough to stand as witnesses in sexual abuse trials, perhaps one of the most important questions to consider is whether or not children ever claim to have been sexually touched when they have not; this is the crime children are most likely to be witnesses to (see Introduction), and the trial will likely involve the use of language they do not fully comprehend. Considering that children will answer certain nonsensical questions to please interviewers, one could assume that children will be just as likely to answer questions about sexual touching even if they do not understand the question. A recent review by Dupree, Patterson, Nugent, and White (2016) uncovered five empirical studies that investigated children's reports after undergoing a genital examination, though none of the studies were more recent than 2003, likely due to tighter ethical constraints on researchers.

In the most interesting of the studies, due to the larger number of variables examined, Saywitz, Goodman, Nicholas, and Moan (1991) examined the reports of 5- and 7-year-old girls in regard to a previous paediatrician's appointment. Half of the girls had their spine examined during the appointment and the other half underwent a genital exam. The children were then asked questions that could be construed as abuse-related (e.g., 'Did the doctor put something in your mouth?'; 'Did the doctor take your clothes off?') and non-abuse related. None of the 7-year-old children made a false report in regard to the abuse-related questions and less than 1% of responses from the 5-year-old children were incorrect.

Children were also asked misleading questions about touching that had not occurred; so, children in the spinal examination condition were asked questions about genital touching, and children in the genital examination condition were asked about spinal touching. Three of the children (out of 35) incorrectly stated they had been vaginally or anally touched. Even after a month, children still responded overall incredibly accurately to questioning. On the other hand, more than 20% of the children in the genital examination condition incorrectly stated they had been touched on their spine. Overall, the results suggest that children cannot easily be misled into making false statements that sexual touching has occurred when it has not.

Importantly, even children who had been genitally touched in the study were hesitant to reveal so. Only eight (22%) and four (11%) of the 36 girls in the genital examination condition admitted to vaginal touching and anal touching respectively in response to free recall requests. Most children within the genital examination condition only disclosed they had been touched if asked specific leading questions (e.g., 'Did the doctor touch you here?'). When asked direct questions, 31 (86%) and 25 (69%) admitted to vaginal and anal touching respectively. Therefore, it would appear that direct questions are sometimes necessary in order to elicit certain information from young children as they do not easily share information regarding some details. It is concerning that 14% and 31% of the children who underwent vaginal and anal touching respectively still refused to disclose this information even after specific questioning. This would suggest that some sexually abused children will not disclose the details even if given specific questions. Since this study was published, however, a number of interview guidelines have been implemented to improve the likelihood of sexually abused children providing disclosures (see next section). Furthermore, as already outlined, Registered Intermediaries are available in England and Wales to assist police officers and lawyers in interviews with children.

These results seem encouraging regarding the capability of children to stand as witnesses, but it is important to note that accuracy rates for some direct questions (e.g., 'Did you take your clothes off?') were as low as 77%, meaning that about 23% of

children incorrectly agreed to something occurring that did not happen. A significant limitation of the study was that, for the children who underwent the spinal examination, the experience may have been far more similar to a typical doctor's visit, meaning their answers to questions may have been more likely to be automatic. However, the results of this study remain significant today as they allowed researchers to examine how children respond to questions about sexual touching in regard to an event the researchers were able to control. According to Dupree et al.'s (2016) review, there is additional support for the finding that children given a genital examination make significantly more omission errors in subsequent interviews about the event than children who undergo a non-genital examination (e.g., an exam for scoliosis).

Several of these studies examined the use of anatomical dolls in the aid of disclosures. Though they are suggestive in their very nature and invite sexual play even when there has been no experience of sexual abuse, anatomical dolls are still frequently used by a number of practitioners (Everson & Boat, 1994, 1997; Faller, 2005; Hlavka, Olinger, & Lashley, 2010). The benefit of them is that they allow the child to overcome any limitations to their verbal ability and any feelings of embarrassment. In some studies, highly trained professionals have not been able to tell the difference between abused and non-abused children based upon their behaviour with anatomical dolls (e.g., Realmuto & Wescoe, 1992). A significant problem is that there is little consistency to how abused children interact with the dolls and so there is no way to devise a pattern of behaviour. For example, some abused children will avoid the dolls, and other abused children will show high levels of sexual and aggressive behaviour with the dolls (e.g., Bruck, Ceci, & Francoeur, 2000). According to Dupree et al.'s (2016) review, however, a child's temperament may impact how beneficial anatomical dolls are likely to be. For instance, they have been found to be less beneficial to children higher in emotionality (Gordon et al., 1993), perhaps because they are more likely to play aggressively with the doll, regardless of sexual abuse history.

A number of factors can impact how likely a child is to disclose details over sexual touching. A child is significantly less likely to disclose sexual abuse if the perpetrator is

a family member or if the child is experiencing feelings of guilt or shame (Leander, Christianson, & Granhag, 2007; Leander, Granhag, & Christianson, 2005). Younger children are also less likely to disclose (Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003; Lippert, Cross, Jones, & Walsh, 2009). Furthermore, they are less likely to understand that the abuse constitutes a crime (Shannon & Törnqvist, 2011, as cited in Ernberg, Tidefors, & Landström, 2016; Sjöberg & Lindblad, 2002) or to question claims by the perpetrator that the abuse is a secret between the two of them (Shannon & Törnqvist, 2011, as cited in Ernberg et al., 2016).

One method to get disclosures of sexual abuse from children other than specific leading questions is the putative confession technique. This is when the investigator tells the child that the suspect has already made a full confession and that the investigator now wants to hear it from the child. This may be beneficial to children who are afraid of how the suspect will react once hearing that the child has reported any abuse. A study by Evans and Lyon (2019) found that putative confessions increased the number of children who disclosed having witnessed someone destroy a laptop from 13% to 62%. Importantly, putative confessions in this study did not lead to any false disclosures. Of course, witnessing someone destroy a laptop is quite different from being a victim of sexual abuse, and so further research is required on the impact of putative confessions in this regard.

### **2.2.3 Accuracy of child eyewitnesses at describing**

While a perpetrator is often someone familiar to the child such as a family member or a teacher, they may sometimes be a stranger to the child. One study has suggested that about a quarter of suspects involved in cases with child eyewitnesses are strangers to the child (Davies & Noon, 1991). In cases of child sexual abuse, statistics suggest that abusers are strangers in at least 10% of cases (Finkelhor & Jones, 2012, as cited in Carroll County Child Advocacy Centre, 2015; Whealin, 2007, as cited in Carroll County Child Advocacy Centre, 2015). It therefore becomes crucial that children can accurately describe a person they saw only once and identify that person from a line-up. To assess a person's ability to describe a stranger, studies often have participants watch a short video or have a target interact with the participant briefly, then ask that participant what they remember about the target. Studies have generally found that

adults usually remember about seven descriptors about a target's appearance (Kuehn, 1974; Lindsay, Martin, & Webber, 1994; Pozzulo & Warren, 2003), while young children (between 5 and 7 years old) remember one or two details (Davies, Tarrant, & Flin, 1989; Karageorge & Zajac, 2011; Pozzulo, Dempsey, & Crescini, 2009), and older children (between 10 and 14 years old) remember two or three (Davies et al., 1989; Pozzulo & Warren, 2003; Karageorge & Zajac, 2011; Zajac & Karageorge, 2009).

Of course, these studies have several limitations. In real life, witnesses may be fearing for their life or enduring a traumatic experience and so this may affect the number of descriptors they remember. Additionally, the studies will not all ask identical questions to each other or exactly the questions asked by real life investigators. There is also disagreement over what constitutes a descriptor. For example, describing hair as long, brunette, and parted could be taken as three descriptors (hair length, hair colour, and hair style) or could be classified as only one descriptor (hair). One would also suspect that witnessing a target on a video rather than in real life may produce a difference in results in regard to how many descriptors are remembered, especially if the quality of the video is questionable. In real life, young children are sometimes capable of reporting far more descriptors about their abusers, even if they are strangers, than the results of these studies suggest (e.g., *R v. Russell Bishop*, 2018). Furthermore, they remember details not included in lab studies, such as voice (e.g., *U.S. v. Mitchell*, 2010) and facial hair (National Registry of Exonerations, 2020).

Research suggests that hair colour is probably the most often reported descriptor by eyewitnesses, both in lab studies (Pozzulo, 2017) and in real life (e.g., Kuehn, 1974). Pozzulo and Warren (2003) found that hair colour was the most frequently reported and the most accurate of descriptors provided by both child (10 to 14 years old) and by adult participants. Other than hair colour, clothing, gender, height, and race are all also frequently reported descriptors. Research suggests that children of any age are far more likely and far more accurate at describing exterior features (e.g., hair, clothes) than interior features (e.g., eye colour, nose shape; Campbell, Walker & Baron-Cohen, 1995; Davies et al., 1989, as cited in Pozzulo, 2017; Pozzulo & Warren, 2003). In fact,

Pozzulo and Warren (2003) found that their child participants struggled significantly with all interior features compared to exterior significantly more than adult participants did. This is unfortunate as while a perpetrator is likely to change their clothes or may change their hair style, interior features are far more likely to remain constant and thus be more valuable to police.

As part of their study, Pozzulo et al. (2009) had child participants engage with a target for about 20 minutes in the classroom and then describe everything they remembered after the target left. Descriptors mentioned included: hair colour (mentioned by 58% of the children), clothing colour (47%), hair length (20%), clothing type (17%), height (7%), eyes (1%), body type (1%), and accessories (1%). Similarly, Karageorge and Zojac (2011) had children (between 5 and 11 years old) view a target for between 30 and 45 seconds at a trip to the fire station, then asked some of the children to describe the target after a period of one to two days and the other children to describe the target after a period of two weeks. Overall, children provided an average of 2.47 descriptors after being given an open-ended format and then prompted until they provided no further information. There was no age difference for children within the short-delay condition, but young children (5 to 7 years old) reported significantly fewer descriptors than older children (8 to 11 years old) within the long-delay condition. Overall, 86.8% of children mentioned a clothing descriptor, 58.9% recalled a descriptor regarding the target's hair, 14% mentioned facial features, and 7% of the children described the target's age. Accuracy rates ranged from 0% to 100%, with an average of about 70%.

As indicated by real-life cases, age, height, and weight are often descriptors that police will specifically ask about if children do not initially volunteer. Generally, children are not very accurate with age, height, and weight descriptors (e.g., Davies, Stevenson-Robb, & Flin, 1988, as cited in Pozzulo, 2017; Pozzulo & Warren, 2003), probably because they have little knowledge of height and weight measures and have fewer life experiences. Research suggests that accuracy in regard to these descriptors does increase with age (Pozzulo, 2017), but that they remain amongst the least accurate of descriptors reported. Police are aware of this issue, however, and will normally ask

children to compare these factors to people they know in order to generate an estimation. For example, in the case of Jimmy Guard, Detective William Devon Jensen asked the victim if her attacker was taller or shorter than him and if her attacker was taller or shorter than Officer Becerra (State of Utah v. Guard, 2015). The victim responded that her attacker was taller than Officer Becerra, but shorter than Detective Jensen, producing a height estimation between 5'7" and 6'1". Unfortunately, this strategy has serious limitations. For example, the height range provided would cover the majority of men.

#### **2.2.4 Accuracy of child eyewitnesses at identifying**

After a child eyewitness has described their target to the police, they may be asked to look through a line-up. Usually, the line-up contains the police's suspect and several people that the police know to be innocent. They are usually shown to the eyewitness one at a time and the eyewitness is then asked if they recognise anyone in relation to the alleged crime they witnessed. At this time, witnesses are told that the perpetrator may or may not be in the line-up (referred to as non-biased line-up instructions). In 2008, a survey revealed that about a third of identification parades conducted in Scotland were viewed by witnesses younger than 16 years old (Memon, Havard, Clifford, Gabbert, & Watt, 2011). Fifty-six per cent of these witnesses selected the suspect, 34% selected a known innocent, and 10% failed to make any identification from the line-up.

Lab research suggests that during target-present line-ups, children as young as 6 years old can be just as accurate as adults (e.g., for a review, see Havard, 2014; Lindsay, Pozzulo, Craig, Lee, & Corber, 1997). However, during target-absent line-ups, children are significantly more likely than adults to make a false identification (e.g., Dekle, Beal, Elliott, & Huneycutt, 1996; Gross & Hayne, 1996; Humphries, Holliday, & Flowe, 2012; Keast, Brewer, & Wells, 2007; Lindsay et al., 1997; Parker & Carranza, 1989; Pozzulo & Balfour, 2006; Pozzulo & Warren, 2003). The main reason is probably because children feel pressured into selecting a photo even when they do not recognise anyone from the array (e.g., Pozzulo et al., 2012). Additionally, children are likely less aware than adults towards the potential consequences of false identifications. While research suggests that adults tend to be more accurate when photos in a line-up are presented

one at a time (sequential line-up) compared to when they are all presented at once (simultaneous line-up), this perhaps makes children more likely to make guesses; Wells and Olsen (2003) found that children made multiple identifications when presented with a sequential line-up.

One successful method of reducing false identification rates is by adding an additional photo to the line-up consisting of a silhouette with a question mark. Zajac and Karageorge (2009) introduced the idea and told children to point to the silhouette card if they did not see the target in the line-up. When testing the idea on children aged between 8 and 11 years old, they found that correct rejections increased from 46% to 71% during target-absent line-ups. They then replicated the findings with children aged between 5 and 7 years old (Karageorge & Zajac, 2011). This time correct rejections increased from 29% to 84%. Importantly, they also found that there was no negative influence of the silhouette card on positive identification rates during target-present line-ups. There are a number of reasons as to why this could be beneficial. First, if children feel compelled to select from the line-up due to pressure then they can still do so without incorrectly identifying an innocent person. Secondly, it allows the child to still give a positive response rather than feel they are disappointing the interviewer.

Havard and Memon (2013) applied this study within a UK setting. Unlike the USA, England and Wales no longer use photo line-ups, but rather use video line-ups (as does Scotland when it comes to witnesses younger than 16 years old). While research has shown that video line-ups reduce false identifications for adults (Valentine, Darling, & Memon, 2007) and teenagers (Havard, Memon, Clifford, & Gabbert, 2010), this conclusion has not yet been supported in regard to children (Beresford & Blades, 2006; Havard et al., 2010). However, Havard and Memon (2013) found that by including a silhouette photo within the video line-up, false identifications significantly decreased. This may be an important avenue for future researchers and investigators to explore.

### **2.2.5 Summary: Children's describing and identifying capabilities**

Generally, young children are indeed capable of making competent witnesses. First, they are unlikely to make disclosures of sexual touching if no such touching has

occurred. However, they are also unlikely to make disclosures of sexual touching if such touching has occurred without the presence of specific questioning. Children are more likely than adults to respond to questions that they do not understand, including those about sexual touching, but only if the questions are presented as closed-ended or option-posing. Second, while young children are likely to produce fewer descriptors than adult eyewitnesses, they are still likely to be accurate with those descriptors, especially when it comes to descriptors regarding hair and clothing. Third, young children are just as likely as adults to accurately identify a perpetrator from a target-present line-up. However, young children are more likely than adults to inaccurately identify a perpetrator from a target-absent line-up. There is still research needed to be conducted within this area, but it would appear that by including an additional mystery target during a target-absent line-up, one might be able to decrease the number of inaccurate identifications made by young children.

### **2.3 External factors of suggestibility**

This section will review the influence of factors of eyewitness performance and suggestibility that are external to the child: the characteristics of the interviewer, the types of questions contained within the interview, and the context of the interview.

#### **2.3.1 Question types**

The type of questions asked by investigators can significantly affect children's responses (e.g., Lamb, Brown, Hershkowitz, Orbach, & Esplin, 2018). Investigators are typically encouraged to apply a 'funnel approach', which involves exhausting free recall invitations and other open-ended prompts before reverting to more focused questions (Lamb et al., 2009). Therefore, police will usually begin by eliciting free recall (e.g., 'Tell me everything that happened from the beginning until the end' or 'Tell me what the suspect looked like'), both with children and adults, when they are questioning eyewitnesses about a criminal event or about a suspect. This way, the witness reports only what comes to mind without influence or suggestion from the interviewer.

During free recall, young children's responses are generally just as accurate as adults', however are not very detailed (Lamb et al., 2018). For example, when asked about the appearance of a suspect, a child witness may only list one or two descriptors, rather

than report everything they remember. Children aged 5 years and younger do not always realise that adults may not know information that they themselves know (see Flavell, 2000 for a review on taking the perspective of others (theory of mind)). By about the age of 6 years, children are significantly more forthcoming when it comes to reporting information, and are significantly less suggestible (e.g., Eisen, Goodman, Qin, Davis, & Crayton, 2007; Goodman & Reed, 1986). As previously mentioned, however, it remains a developmental progression (McWilliams, Narr, Goodman, Ruiz, & Mendoza, 2013).

In order to elicit more information so that the police can narrow down on a suspect, follow-up questions are required. Due to this, interviewers often struggle to maintain best practice and an overuse of improper questioning is reported (e.g., Luther, Snook, Barron, & Lamb, 2015; Roberts & Cameron, 2015; Waterhouse, Ridley, Bull, La Rooy, & Wilcock, 2018). The safest type of follow-up questions are those that are open-ended as they generate more detailed, accurate and coherent responses than any other type of follow-up questions (e.g., Feltis, Powell, Snow, & Hughes-Scholes, 2010; Hershkowitz et al., 2012; Orbach, Hershkowitz, Lamb, Sternberg, Esplin, & Horowitz, 2000; Lyon, 2014; Snow, Powell, & Murfett, 2009). According to Cronch, Viljoen, and Hansen (2006), cued invitations are the safest type of open-ended follow-ups (e.g., 'You mentioned the man touched you. Tell me more about that') as these types of questions only ask the witness to expand on points that were first mentioned by the witnesses themselves. Other types of invitations are general invitations (e.g., 'Then what happened') and time-segmentation invitations (e.g., 'Tell me everything that happened from the first time he touched you until he ran away').

In order to narrow in on specific details investigators will then typically use open-ended directive questions (e.g., 'What did he do when he touched you?' or 'How did he touch you?'). Directive questions require a more direct response and usually begin with 'How', 'What', 'Where', or 'Who', but so long as they do not force the child to choose from a limited number of possible answers, then the question may still be considered open-ended (Lamb et al., 2018).

Even open-ended follow-ups may not elicit enough information, and so police may then proceed to use closed-ended follow-ups. This may be directive questions that require a one- or two-word response (e.g., 'You mentioned he touched you in the house. What room in the house?'), or a question that option poses in some other form (e.g., 'Was he alone or was he with someone?'). Most often though, these types of questions require only a 'yes' or 'no' response (e.g., 'Was he tall?'). Research shows that option-posing often makes young children more likely to answer a question even when they do not know the answer or understand the question (e.g., Peterson et al., 1999; Waterman et al., 2000, 2001, 2004).

Lastly, if investigators still do not elicit enough information from open-ended or closed-ended follow-ups, they may resort to using suggestive or leading questions (Lamb et al., 2018). These are questions that force the response in a specific direction (e.g., 'He touched you, didn't he?') or introduce information not already mentioned by the child during the interview (e.g., asking about touching when the child has not mentioned touching). Unlike in lab settings, investigators in real life have no way to know if their question is misleading (a leading question that uses false information) or not. In sum, free recall and open-ended follow-ups may not elicit a whole lot of information from young children, but they are the best way of guaranteeing that responses will be accurate.

In a study by Hill and Davies (2013), researchers carefully analysed transcripts of forensic interviews with child witnesses from England and Wales, examining how frequently each type of question was uttered by investigators. This study examined 13 interviews between 1993 and 1998 using the Memorandum of Good Practice (MoGP) and 12 interviews between 2002 and 2009 using the latest implemented guidelines in England and Wales for interviewing child witnesses (i.e., Achieving Best Evidence; ABE). In both samples, free invitations only made up about 8% of the overall questions asked by investigators. There was also a similar number of option-posing questions in both cases (MoGP – 47% vs. ABE – 44%), as well as directive questions (42% vs. 43%),

and suggestive questions (3% vs. 5%). This research, as well as similar findings (e.g., Luther et al., 2014), shows that, even after the implementation of new guidelines, investigators do not always follow the advice of asking as many free invitation questions as possible.

### **2.3.2 Question repetition**

Children do not have to be asked a leading or misleading question to provide inaccurate information. Evidence suggests that when children are asked the same question repeatedly, they seem to sometimes think they must have previously given an incorrect response and so should change their answer, or if they did not answer to begin with then they may suspect that the interviewer is looking for an answer and so give one even if they are uncertain. Importantly, the more children repeat an answer, the less their uncertainty becomes detectable (Poole & White, 1991). Unfortunately, there are numerous examples from real cases over the negative effects of question repetition. Below is an extract from an interview with a child from the Kelly Michaels case (Ceci & Bruck, 1995, p. 280):

Interviewer: When Kelly kissed you, did she ever put her tongue in your mouth?

Child: No.

Interviewer: Did she ever make you put your tongue in her mouth?

Child: No.

Interviewer: Did you ever have to kiss her vagina?

Child: No.

Interviewer: Which of the kids had to kiss her vagina?

Child: Me.

Clearly, the use of repeated questions can confuse children and cause their statements to be inconsistent. Now that judges and lawyers are more informed about how suggestible children can be, inconsistent statements can cause the child to appear as

an unreliable witness and have their testimony be excluded from trial, as was the case in *MacLennan v. HM Advocate* (2015; see Introduction). The use of improper questioning, therefore, benefits no one. If the child is incorrectly swayed by improper questioning and their testimony is given at trial, it may result in a wrongful conviction. On the other hand, if the child is correctly swayed by improper questioning, their testimony may be suppressed, and the guilty suspect could walk free.

Nowadays, the negative effects of repeated questioning are well understood by researchers, but their recommendations are still scarcely followed by investigators (e.g., Luther et al., 2014; Roberts & Cameron, 2015). The issue of question repetition has been largely ignored by most interviewing protocols, such as the National Institute of Child Health and Human Development Protocol (NICHD) in America and the guidelines in Norway and Sweden (Cederborg, 2004; Cederborg, Orbach, Sternberg, & Lamb, 2000; Myklebust & Alison, 2000). In England and Wales, the ABE interviewing protocol (Home Office, 2007) warns against the use of verbatim repetition (using the exact same words each time).

Verbatim repetition is the only type of question repetition that has been widely addressed by research (e.g., Howie, Sheehan, Mojarrad, & Wrzesinska, 2004; Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Poole & White, 1991; Powell & Thomson, 1996), however Krähenbühl (2007) found that verbatim repetition only accounted for 10% of repetition by police interviewers when analysing transcripts of police interviews in England and Wales. It would seem that when warned against repeating questions verbatim, investigators still repeat questions, but just in another form. Though every type of question repetition had a negative influence on the accuracy and consistency of children's reports, Krähenbühl and Blades (2009) found that verbatim repetition had the least effect when comparing the effects of four different types of question repetition. Gist repetition (maintaining the meaning, but not the wording of the question) was the most frequently found form of repetition, accounting for over half of repetitions and was found to have only a slightly worse effect than verbatim repetition. Specific to closed (repeating a specific question in the

form of an option-posing question) and closed to specific (repeating an option-posing question in the form of a specific question) both had a more negative influence on the accuracy and consistency of children's reports than either verbatim or gist repetition. This may be because the children interpret the change in structure to mean that a change in response is required (Krähenbühl & Blades, 2009).

In Scotland, the guidelines (The Scottish Government, 2014) advise against repeating questions in any form, and that open-ended prompts (general invitations, cued invitations, and time-segmentation invitations) should be used instead. The guidelines also state that if questions are repeated then it should be explained to children that the question is not being repeated because the child's answer was incorrect the first time, but rather because the interviewer may not have heard correctly. These guidelines, however, are rarely followed exactly (La Rooy, Earhart, & Nicol, 2013), and so clearly more research and educating of investigators over the negative effects of question repetition and the different forms of question repetition is required.

### **2.3.3 Other external factors of suggestibility**

#### *2.3.3.1 Delay and repeated interviews*

It is common for children to be interviewed multiple times after witnessing a crime. When examining 217 random sexual abuse cases from the Los Angeles Dependency Court (1999–2000), Malloy, Lyon, and Quas (2007) found that children were interviewed an average of 12 times. Importantly, this only counts official police interviews; it does not consider questioning by friends and family. Lab studies have shown that when children are repeatedly interviewed with leading questions, this increases suggestibility (for a review, see Quas, Goodman, Ghetti, & Redlich, 2000). For example, the mousetrap study by Ceci et al. (1994) showed that repeatedly interviewing a child with suggestive questions over an event that never happened (e.g., getting one's hand caught in a mousetrap) could lead to children later reporting detailed accounts over that event and mistaking it as a true memory.

More recently, research has shown that so long as interviews do not contain any leading questions, repeated interviews may be beneficial, especially if these interviews

follow the format provided by the NICHD (e.g., Hershkowitz & Terner, 2007). Repeated interviews will strengthen a child's memory of an event, and, so long as there is no suggestive questioning used, these memories are unlikely to be false (e.g., Baker-Ward, Gordon, Orenstein, Larus, & Clubb, 1993; Memon & Vartoukian, 1996). Additionally, repeated interviews may increase one's resistance to later suggestions (e.g., Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991). So long as there is not a large delay between interviews, new disclosures stemming from repeated interviews can be accurate (e.g., Hershkowitz & Terner, 2007). For example, La Rooy, Pipe, and Murray (2005) reported a 92% accuracy rate for new information obtained from a second interview, one day after the first interview.

In these cases, repeated interviews lead child witnesses to disclosing certain details they had not previously disclosed (referred to as reminiscence). Little research has shown, however, that repeated interviews lead to the recall of more new information in comparison with the amount forgotten (referred to as hypermnesia; e.g., Hershkowitz & Terner, 2007). More likely, children forget or omit certain details in later interviews compared to their original statements. While child witnesses may also provide new information during repeated interviews after long delays, even after as much as five or six years (e.g., Fivush, McDermott Sales, Goldberg, Bahrack, & Parker, 2004; Peterson & Whalen, 2001; Salmon & Pipe, 2000), this new information is unlikely to be accurate, due to a mix of greater decay and opportunities for incorporating misinformation (Peterson & Whalen, 2001; Salmon & Pipe, 2000). Overall, repeated interviews may be beneficial for uncovering new forensic information so long as there is little delay between interviews and that there is no use of suggestive questioning.

### *2.3.3.2 Interviewer bias*

Another factor thought to affect the quality of eyewitness testimony is interviewer bias. This is when an interviewer has prior beliefs about the subject of the interview and only asks questions that seek to elicit statements in support of those beliefs, while consciously or unconsciously avoiding questions that may result in statements that are in conflict with those beliefs (Ceci, Hritz, & Royer, 2016). For example, in the day care sexual abuse cases, investigators often thought that abuse had happened before children had even mentioned anything about abuse. The questions during these

interviews revolved around coaxing the children into making statements that would reveal details about the abuse (e.g., Ceci & Bruck, 1995; Garven et al., 1998; Schreiber et al., 2006), and the interviewers avoided questions which may have provided alternative explanations over the behaviour of the children.

To demonstrate interviewer bias, a study by Clarke-Stewart, Thompson, and Lepore (1989, as cited in Ceci & Bruck, 1995) introduced 5- and 6-year-old children to a janitor named Chester during their class. For some groups of children, Chester merely cleaned some dolls and toys in the classroom. For other children, Chester handled the doll far more roughly. Later, the children were interviewed about what they had witnessed. Some of the interviewers were accusatory (suggesting that Chester had been playing inappropriately with the toys), some were exculpatory (suggesting that Chester was merely cleaning the toys), and some were neutral and non-suggestive. When interviewed by a neutral interviewer, the children's accounts were generally accurate with what they had witnessed. However, when the interviewer contradicted what the child had witnessed, their accounts conformed with the suggestions of the interviewer, with 90% of children giving answers that were in agreement with what the interviewer was suggesting, rather than with what had actually happened.

Often, interviewer bias is reflected by emotional tone and by body language. For example, if an interviewer is getting the responses they desire then they may give the child positive feedback or nod their head. Interviewer bias may also be demonstrated via question repetition, option-posing, leading questions, stereotype induction, or any kind of feedback. Either way, it is extremely risky for an interviewer to pursue one single theory over an event. In a study by Leichtman and Ceci (1995), for example, preschool-age children were interviewed about a man's visit to class. For some of the children, the interviewer provided highly biased statements that indicated the man was clumsy. By end of the questioning, a significant portion of these children had provided false reports of the man damaging property and behaving clumsily, even though the visit had been brief and ordinary. In a similar study, Quas et al. (2007) found that children interviewed by a highly biased interviewer about a play session

were significantly less accurate than those interviewed by a less biased interviewer. These studies suggest that biased statements and interviewer pressure can increase the chances of false reports. This is similar to the day care sexual abuse hysteria cases of the 1980s and early 1990s, when investigators would use peer pressure and negative stereotypes of the defendants to get disclosures from the children (Ceci & Bruck, 1995). A range of factors may impact the potential impact of interviewer bias upon a child, including those relating to the child and to the interviewer, such as perceived interviewer status (for a review, see Hritz, et al., 2015).

#### *2.3.3.3 Interviewer status*

Children are more likely to believe adults than other children based solely upon their status as being older (e.g., Ceci et al., 1987). More than this, young children seem to recognise status to a certain extent and are more suggestible when they are interviewed by police officers with a more senior position (e.g., Tobey & Goodman, 1992). Overall, there is little evidence that this is a serious concern so long as proper interviewing practice is applied (i.e., proper questioning and no interviewer bias). For example, in the case of Kelly Michaels, one of the police officers told a young witness that he could put bad people in prison, and if the boy told him what he wanted to hear, he would introduce him to the other important men who put Kelly in prison (Ceci & Bruck, 1995). In this instance, the police officer abused his status to coax the response that he desired from the young witness. Once this happens, children are vulnerable to suggestion and thus to making inaccurate statements.

Though it may seem reasonable that recognition of status may impact a child's suggestibility, little research has further investigated this. Bull and Corran (2002) argue that this is because many researchers consider it to be a resolved issue, given that the American Psychological Association (1998, as cited in Bull & Corran, 2002) instructs psychologists, when interviewing children, to appear relaxed and informal. Therefore, a number of researchers, inspired by this line of research, have instead investigated the potential impact of emotional tone on suggestibility.

#### *2.3.3.4 Emotional tone*

Young children provide more accurate information when they feel comfortable during the interview and when the interviewer is both supportive and positive in tone (e.g.,

Goodman et al., 1991). Furthermore, authoritative interviewing styles increase the number of errors in response to suggestive questions (Bull & Corran, 2002). In order to achieve a comfortable atmosphere, it is important that rapport is built, perhaps by talking to the child first about their interests. There is a problem when investigators adjust their tone based upon the responses of the child, rather than providing a positive and supportive tone throughout. For example, if the investigator only provides a positive tone in response to some statements then the child may construe this as positive feedback and then state what they think the interviewer wants to hear, as opposed to what actually happened. There are now a number of interviewing formats to assist investigators with consistency across interviews.

### **2.3.4 Current interviewing formats**

#### *2.3.4.1 The National Institute of Child Health and Human Development (NICHD) investigative interview protocol*

The NICHD investigative interview protocol is the most widely studied interviewing format (Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007). It can be used with children as young as 4 years old. Since it was first introduced in 1996, research on the format has studied over 40,000 forensic interview transcripts with children. Currently, it is available in 15 different languages and is used in the USA, Bulgaria, China, Canada, Finland, Georgia, Germany, Greece, Italy, Israel, Portugal, Romania, Russia, Spain, Sweden, and the Netherlands. The technique uses a semi-structured interview format, along with step-by-step instructions, therefore removing a lot of the guesswork on the part of the investigator that is present in the use of other formats.

The format begins with an introduction and by establishing the ground rules of the interview (e.g., the child should not respond to a question they do not understand or do not know the answer to), as well as by ensuring the child understands the difference between a truth and a lie (Lamb et al., 2007). A forensic interview is a very different type of conversation than a child will be used to. In most cases, a child converses with an adult about a topic that the adult knows more about than the child. Furthermore, in school if a child answers a question incorrectly then they expect the teacher to set them right. It therefore needs to be carefully explained to the child that they may know information the adult does not during a forensic interview.

There is then a rapport-building phase, in which the investigator should attempt to generate a positive relationship with the child witness. This may involve asking the child about something they like to do (Lamb et al., 2018). As children may be asked to provide information that is potentially embarrassing or that makes the child uncomfortable, they will be more likely to provide details if a sense of trust is created between the child and the investigator (e.g., Saywitz, Larson, Hobbs, & Wells, 2015; Vallano & Schreiber Compo, 2015).

The third stage is a practice interview, sometimes referred to as training in episodic memory. This is an interview about a neutral, unrelated event (e.g., a holiday, a hobby that came up during rapport building). This allows the child to become familiar with what is expected during the actual interview (i.e., that they should provide lots of details in response to open-ended questions). During this stage, it is important for the investigator to use open-ended questions. Research shows that children produce 2.5 times as much information during the actual interview if they are asked open-ended questions during the rapport-building (or practice interview) phase compared to closed-ended questions (Sternberg et al., 1997).

As it transitions to asking about the crime in question, the following stages of the protocol rely on free recall and uses open-ended follow-up questions (e.g., general invitations, cued invitations, and time-segmentation invitations) to prompt further information as it transitions to asking about the crime in question. It is recommended the use of directive questions are deferred and that closed-ended and suggestive questions are avoided, if possible (e.g., Lamb et al., 2018).

A study by Lamb et al. (2006, as cited in Lamb et al., 2018) compared 50 interviews using the NICHD protocol with 50 interviews not using the protocol in the UK with child witnesses. They found the NICHD interviews produced a significantly higher number of free invitations (34% vs. 7%), and that there was also a decrease in how many of the

questions were phrased as option-posing (18% vs. 27%). Furthermore, 56% of the details elicited during NICHD interviews were in response to free invitations, as opposed to only 14% during non-NICHD interviews. Similar results have been found in different countries, comparing question types and details elicited by question types between NICHD interviews and non-NICHD interviews (e.g., Sternberg, Lamb, & Orbach, 2001).

#### *2.3.4.2 Achieving Best Evidence*

The MoGP was the interviewing protocol implemented in England and Wales in 1992 to be used with children under the age of 14 years old for violent offences and under the age of 17 years old for sexual offences (Home Office and Department of Health, 1992, as cited in McCarron, Ridgway, & Williams, 2004). In 2001, the protocol was replaced with the ABE interviewing protocol (Home Office, 2001, as cited in Krähenbühl & Blades, 2006). The ABE is used with all children under the age of 17 years old, no matter the type of offence that has occurred. The ABE establishes rapport, asks for free recall, and then uses specific questions that usually start with 'Wh' (e.g., who, what, when, where, why) in order to gain further information. It does not make use of a practice interview, as the NICHD does.

#### *2.3.4.3 Stepwise interview*

The Stepwise Interview was developed by Yuille, Hunter, Joffe, and Zaparniuk in 1993 (as cited in Pozzulo, Forth, & Bennell, 2018). Currently, the Stepwise Interview is used in all areas of Scotland except for one (Scottish Government, 2011). It has the interviewer follow specific steps, similar to the NICHD, including the establishing of rapport, asking for free recall, and then the use of follow-up questions. Like the NICHD protocol, but unlike the ABE, the Stepwise Interview includes a practice interview. A quality analysis (La Rooy et al., 2013) of interviews conducted with children in Scotland, however, examined 74 random transcripts, and found that none of 74 interviews included the practice interview. It would appear that though practitioners in Scotland are urged to follow the specific steps of the Stepwise Interview, certain steps are largely ignored in practice. The quality analysis did reveal that there are significantly fewer leading and closed-ended questions being used by practitioners in Scotland compared to interviews conducted two years previously, which is indeed promising.

### **2.3.5 Summary: External factors of suggestibility**

Back at the beginning of the 20th century, it seemed misleading questions were the only suggestive technique. Now, research has indicated various other factors that can increase suggestibility, including repeated questioning, emotional tone, and negative feedback. These techniques are at their highest level of negative influence when in the presence of interviewer bias. So long as the interviewer is testing alternative hypotheses, challenging the statements, and allowing the conclusions to come straight from the words of the eyewitness, at least most of these factors should not pose serious problems.

### **2.4 Internal factors of suggestibility**

As already discussed, some children will agree immediately with misleading questions by interviewers, and some will continuously deny that a suggested event ever occurred. This section will explore some of those individual differences that may play a role, primarily focusing on temperament.

Firstly, there are demographic factors, including age, gender, and socioeconomic status to discuss. As stated earlier in this chapter, age is one of the most reliable predictors of memory recall in witnesses, with younger children (especially below the age of 6 years) typically reporting less details than older children and adults (e.g., Ceci & Bruck, 1993, 1995; Hershkowitz et al., 2012). During free recall, though older children provide more details, there are often no age-related differences in accuracy (e.g., Sutherland & Hayne, 2001). In some cases, however, older children still perform significantly more accurately than younger children even during free recall (e.g., Eisen, Qin, Goodman, & Davis, 2002). As explained, younger children are typically more vulnerable than older children and adults to external suggestive influences (e.g., Ceci & Bruck, 1993; Ceci & Huffman, 1997; Ghetti et al., 2002; Loftus & Davies, 1984; Otgaar et al., 2009; Wade et al., 2002).

There is no contemporary evidence that gender is associated with differences in suggestibility (e.g., Alexander et al., 2002; Bruck & Melnyk, 2004; Burgwyn-Bailes, Baker-Ward, Gordon, & Ornstein, 2001; Chae, 2004; Clarke-Stewart et al., 2004; Imhoff

& Baker-Ward, 1999; Klemfuss & Olaguez, 2020; Quas et al., 1999; Roebbers & Schneider, 2001; Rossi, Benatti, Pesce, Oppo, & Avato, 2011; Young, Powell, & Dudgeon, 2003), despite earlier claims that female child witnesses are more suggestible than males (e.g., Bolton, 1896; Gross, 1911; Hurlock, 1930; Lombroso, 1878, as cited in Meares, 2016; Messerschmidt, 1933; Stern, 1910, as cited in Ceci & Bruck, 1993). The change in findings perhaps reflects the fact that female participants in eyewitness studies feel more confident in correcting investigators when given misleading questions than they once did as the status of women has improved, and maybe even that early researchers within a male-dominated psychological discipline sought results to reinforce social prejudices (Meares, 2016). Unfortunately, it is still a common misconception even among academics that these gender differences exist. For example, Volpini, Melis, Petralia, and Rosenberg (2016) hypothesised that girls would be more suggestible than boys in their study due to an outdated understanding of the literature; their results failed to support their predictions.

There is little evidence to suggest that socioeconomic status affects eyewitness performance or suggestibility in any way (Alexander et al., 2002; Clarke-Stewart et al., 2004; Geddie et al., 2000), though few studies have attempted to examine the relationship. One exception is a study by McFarlane et al. (2002), which found that children from lower socioeconomic backgrounds were more suggestible (though it accounted for little of the variance – 2%). This study had a larger participant pool than comparative studies (220 compared to 51-70; Alexander et al., 2002; Clarke-Stewart et al., 2004; Geddie et al., 2000), which may indicate that any influence of socioeconomic status may be small and only be detected in large samples.

As discussed in the Introduction, there are a range of cognitive and personality-related factors associated with suggestibility, including intelligence (e.g., Bettenay et al., 2015; Danielsdottir et al., 1993; Endres et al., 1999; Geddie et al., 2000; Gignac & Powell, 2006; Henry & Gudjonsson, 2004; Hurlock 1932; McFarlane et al., 2002; Roebbers & Schneider, 2001; Singh & Gudjonsson, 1992), creativity (Brown, 1999, as published in Bruck & Melnyk, 2004; Clarke-Stewart et al., 2004; Melnyk, 2002; Purdy, 2001), and

attachment styles (Chae et al., 2014; Clarke-Stewart et al., 2004; Crossman, 2001; Schaaf, Alexander, & Goodman, 2008; Quas et al., 1999). According to reviews (Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020), these are the three most reliably reported significant associations, following age. As mentioned in the Introduction, there is also evidence that temperament may impact eyewitness performance and suggestibility. However, the association between temperament and suggestibility is less reliable than the relationships between suggestibility and some of the previously mentioned factors. Before exploring the theoretical association between temperament and suggestibility, major studies that developed our understanding of temperament will first be explored.

### **2.4.1 Temperament**

While the term temperament is often reserved for the explanation of individual differences in infancy and childhood, and the term personality is used when talking about adolescents and adults (Gonzalez et al., 1994), it is not clear at what point individual differences become reflective of personality as opposed to temperament. Nevertheless, it is generally agreed upon that temperament provides the foundations for later personality (Goldsmith et al., 1987). In fact, Hagekull (1994, as cited in Gonzalez et al., 1994) defined temperament traits as simply being early appearing personality traits. However, most personality psychologists would probably view this as being too simplistic. Instead, Buss (1989) claims that temperament is the building blocks for personality, but that personality is more complicated due to being a result of interactions between temperament and social factors. Similarly, Rothbart (2007, p.1) said 'temperament and experience together "grow" a personality.'

Therefore, temperament is merely one contributor to personality development, along with many other factors, including parenting styles (e.g., Ang, 2006; Doinita & Maria, 2015; Karavasilis, Doyle, & Markiewicz, 2003), family dynamics (for a review, see Yap, Pilkington, Ryan, & Jorm, 2014), and peer relationships (e.g., Powers & Bierman, 2013). Importantly, temperament can have an impact on a number of these factors, and so they are in constant interaction. For example, mothers tend to treat non-irritable babies with significantly greater sensitivity than babies rated as irritable (van den Boom & Hoeksma, 1994, as cited in Berk, 2012), perhaps due to getting more sleep at nights and thus having greater energy when spending time with the infant during the

day. In a longitudinal study by Deal, Halverson, Havill, and Martin (2005), temperament assessed in early/middle childhood accounted for an average of 32% of the variance in personality, as measured in late adolescence/young adulthood 15 years later.

Therefore, it is a significant precursor to later personality, but not the only determining factor.

In a landmark study, referred to as the New York Longitudinal Study, Thomas and Chess (1977, as cited in Thomas & Chess, 1986) studied individual differences in the temperament of infants. According to Thomas and Chess (1986), temperament refers to the style of behaviour (e.g., withdrawing from a social situation), rather than the motivation of behaviour (e.g., why one withdraws from a social situation). In this sense, temperament is primarily a pattern of behavioural responses. The researchers followed 133 participants from three months old to adulthood and derived nine dimensions of temperament based upon their data (activity, rhythmicity of biological functions, initial approach-withdrawal, adaptability, intensity, mood, persistence, distractibility, and sensory threshold). They factor analysed these nine dimensions and developed three types of temperament: easy children (generally cheerful and low in emotionality), difficult children (tend to react negatively and intensely), and slow-to-warm-up children (adjust slowly to new environments). However, only 65% of the sample could be categorised based upon their model. Furthermore, future studies failed to find the emergence of these three dimensions through factor analysis (Martin & Bridger, 1999). Critics have pointed out the relatively low sample size of the study as a source of problems (Martin & Bridger, 1999).

Buss and Plomin (1975, 1984) provided another framework of temperament, specifying criteria, based upon ideas by Allport (1937), that a trait had to meet in order to be considered a dimension of temperament, including heritability, stability, and being predictive of adult behaviour. Originally, they identified four dimensions of temperament based upon this: emotionality, activity, sociability, and impulsivity, giving rise to the acronym EASI (Buss & Plomin, 1975). However, they later dropped impulsivity due to there being a lack of evidence over its heritability (Buss & Plomin,

1984; Goldsmith et al., 1987). Researchers criticised the authors for over-emphasising the importance of heritability, believing that it oversimplified and masked the complex nature of temperament (Gonzalez, Hynd, & Martin, 1985). Instead, Gonzalez et al. (1985) endorsed a systems perspective, in which temperament is viewed as being hierarchically organised into multiple levels (genes, environment), with all levels interactively influencing each other.

Future researchers, including Rothbart and Martin, independently, focused on explaining the underlying neurobiological processes of temperament, specifically ideas introduced by Eysenck (1967, as cited in Gray & McNaughton, 2003) and extended by Gray (1982, as cited in Gray & McNaughton, 2003), whose biopsychological theory of personality is well supported by subsequent research and currently has general acceptance among professionals (Allen & DeYoung, 2017; DeYoung 2010; Funder, 2019). Furthermore, growing acceptance of the Big Five model of personality (openness, conscientiousness, extraversion, agreeableness, and neuroticism; discussed in detail in Chapter Five) has resulted in temperament theorists having a framework in which to operate. For example, Ahadi and Rothbart (1994) stated that the Big Five provided a target for attempting to identify how early temperament dimensions connect with later personality traits.

Rothbart and Ahadi's (1994) model proposes three primary dimensions of temperament (surgency, negative affect, and effortful control) with surgency, referring to positive affect, representing the behavioural manifestation of Gray's behavioural activation system (BAS), and negative affect, conceptualised as threat sensitivity, being linked to Gray's BIS (Rothbart, Derryberry, & Posner, 1994). Based upon this model, surgency has been linked to extraversion (which has also been linked to Gray's BAS; Allen & DeYoung, 2017; Depue & Collin, 1999; DeYoung, 2010), negative affect has been associated with neuroticism (which has also been linked to Gray's BIS; Allen & DeYoung, 2017; DeYoung, 2010), and effortful control has been connected to conscientiousness (Caspi, 1998; Rothbart, 2007). In a study with adult participants, perceptual sensitivity and affiliation, lower dimensions subsumed by surgency, were

associated with openness and agreeableness (Evans & Rothbart, 2007). The clear and logical link between Rothbart and Ahadi's (1994) model of temperament and the Big Five has made it an appealing model to developmental psychologists (e.g., Shiner & DeYoung, 2013).

Similarly, Martin and Bridger (1999) have attempted to draw connections between Martin's (1988) temperament dimensions (activity, adaptability, approach/withdrawal, ease of management through distraction/distractibility, emotional intensity, and persistence) and the Big Five personality traits. Unlike Rothbart and Ahadi's (1994) model, however, which was inspired by research on the Big Five (Ahadi & Rothbart, 1994), Martin (1988) built on ideas stemming from Thomas and Chess's (1977, as cited in Thomas & Chess, 1986) landmark study. Martin and Bridger (1999) argue that inhibition (i.e., shyness) and impulsivity (i.e., activity) are markers for individual differences in the sensitivity of the BIS and BAS respectively. Similar to findings on neuroticism in adulthood (for reviews, see Allen & DeYoung, 2017; DeYoung 2010), the temperament dimension of inhibition (or shyness or fearfulness) in childhood is thought to have a biological basis rooted in the limbic system, especially the amygdala (Kagan, 1989).<sup>1</sup> Furthermore, according to hypotheses given by Martin and Bridger (1999), a high score of activity is conceptually linked with high extraversion and low neuroticism, whilst high inhibition/shyness is thought to be associated with low extraversion and high neuroticism. They also propose that high adaptability is linked to high agreeableness and low neuroticism, and high emotionality is thought to be a developmental precursor for high neuroticism and low agreeableness.

As mentioned in the Introduction, even in adulthood, shyness is thought to be a combination of introversion and neuroticism (e.g., Briggs, 1988; Geen, 1986;

---

<sup>1</sup> Indeed, the amygdala of shy people becomes highly active when they are shown pictures of people that they do not know (Birbaumer et al., 1997), and people with anxiety disorders, who have panic attack, or post-traumatic stress disorder (PTSD) tend to have an active amygdala all the time, even at rest (Drevets, 1999). Dysfunction in the amygdala, causing lower activation, is related to psychopathy, partly characterised by a lack of anxiety and fear (Glenn & Raine, 2008; Pardini, Raine, Erickson, & Loeber, 2014; Raine, 2014; Yang, Raine, Colletti, Toga, & Narr, 2010; Yang, Raine, Narr, Colletti, & Toga, 2009).

Kwiatkowska & Rogoza, 2019; Paulhus & Trapnell, 1998). Eysenck (1967, as cited in Gray & McNaughton, 2003) connected low extraversion to a preference for minimum levels of arousal, explaining why those low in extraversion (introverts) may prefer to avoid social situations.<sup>2</sup> Introverts, therefore, may be perceived as shy due to a low desire to socialise, but lack a fear of social judgement, an important aspect of shyness (Afshan et al., 2015). This may be one reason as to why self- and peer-ratings do not always match (Spooner, Evans & Santos, 2005), as introversion may be confused for shyness in peer-ratings. Importantly, both shyness and neuroticism predict similar outcomes, including depression, a higher reported level of fear in everyday life, and social anxiety (Chavira, Stein, & Malcarne, 2002; Crozier, 1995; D'Souza, Gowda, & Gowda, 2006; Heiser, Turner, & Beidel, 2003; Jeronimus, Kotov, Riese, Ormel, 2016; Judge, Locke, Durham, 1997; Kale, Inamdar, Shimpi, & Jha, 2020; Kamath & Kenekar, 1993, as cited in Afshan et al., 2015; Ran, Zhang, & Hao, 2018). Also, according to Kamath and Kenekar (1993, as cited in Afshan et al., 2015), shy individuals with social anxiety are more introverted and neurotic than shy individuals without social anxiety.

There is indeed robust evidence that extraversion in adulthood can be predicted in childhood based upon high activity levels and high sociability (Caspi & Shiner, 2006, as cited in Shiner & DeYoung, 2013; Durbin, Hayden, Klein, & Olino, 2007). In a 23-year longitudinal study by Caspi et al. (2003), 3-year-old children rated as being high in confidence and friendliness were significantly higher in extraversion in adulthood than children rated as being shy/inhibited. Shiner and DeYoung (2013) note that activity level is likely to be a more prominent component of extraversion in early childhood, prior to the emergence of assertiveness, which is a more prominent component in

---

<sup>2</sup> There is evidence to support the notion that extraverts have a lower level of cortical arousal than introverts do, though not all findings have been replicated (e.g., Bartol & Costello, 1976; Cassidy & MacDonald, 2007; Corr, Pickering, & Gray, 1995; Dobbs, Furnham, McClelland, 2011; Furnham & Bradley, 1997; Furnham & Strbac, 2002; Johnson et al., 1999; Richard & Eves, 1991). Overall, evidence seems to suggest that Eysenck was about half right; introverts do not appear to be more chronically aroused than extraverts (Stelmack, 1990), and the ARAS does not work like a tap by turning stimulation to the brain on and off (Zuckerman, 1991, as cited in Flunder, 2019). However, introverts do seem to react more strongly and often more negatively to sensory stimulation than extraverts (Zuckerman, 1998, as cited in Flunder, 2019). In other words, extraverts and introverts are about equally aroused in the absence of external stimulation. However, when loud or bright stimuli are presented, introverts react more quickly and more strongly, and this reaction may lead them to withdraw from certain loud environments.

adulthood. They further theorise that the dramatic change in testosterone levels during development, particularly during puberty, is likely to contribute towards this shift, given that testosterone is involved in both activity and assertiveness (DeYoung & Gray, 2009), as well as changes in the way approach behaviour is expressed as children gain increased verbal skills and a more solid understanding of social guidelines (Shiner & DeYoung, 2013).

According to Ran et al. (2018), the relationship between shyness and social anxiety is partially mediated by the behavioural inhibition system, an important aspect in Gray's biological theory of personality (reinforcement sensitivity theory), which is responsive to cues in the environment for punishment and uncertainty (Gray, 1970 as cited in Gray & McNaughton, 2003). Those with an overactive BIS, thought to be associated with low levels of serotonin, a neurotransmitter that helps dampen negative emotions, are especially sensitive to these cues, making them more likely to experience anxiety (Gray & McNaughton, 2003).<sup>3</sup> Indeed, those with high levels of BIS activity are more likely to display avoidance behaviour around strangers (Van Brakel & Muris, 2006), and are at greater risk for developing social anxiety when BIS activity is measured in childhood (Coplan, Wilson, Frohlick, & Zelenski, 2006).

Similarly, neuroticism, or related negative emotion components, in later childhood or adulthood has been predicted in early childhood based upon irritability or intense displays of negative emotion (Caspi & Shiner, 2006, as cited in Shiner & DeYoung, 2013; Durbin et al., 2007). Highly intense displays of negative emotion during the Strange Situation at 18-months was predictive of high neuroticism scores at 3.5 years (Abe & Izard, 1999). Interestingly, Martin and Bridger's (1999) separately theorised precursors of neuroticism (low activity and high shyness as one, and low adaptability and high emotionality as the second) conceptually link nicely to DeYoung's (2006)

---

<sup>3</sup> Low serotonin levels can lead to serious problems. For example, dangerous criminals and people who commit suicide have been found to have low serotonin levels (Mann, Arango, Underwood, 1990; Virkkunen, Goldman, Nielsen, & Linnoila, 1995). Selective serotonin reuptake inhibitors are a common antidepressant that people take to combat depression.

distinct aspects of neuroticism (withdrawal and volatility). Shiner and DeYoung (2013) also theorise that the distinction between fearfulness/inhibition and anger/emotionality in childhood may correspond to withdrawal and volatility in adulthood individually.

According to Martin and Bridger (1999), the only temperament dimension that has a one-to-one connection with a Big Five personality trait is persistence, which is thought to be linked to conscientiousness. In support of this, quitting easily has been linked to low levels of effortful control by Rothbart and Bates (2006), a predictor of low conscientiousness in adulthood (Caspi & Shiner, 2008, as cited in Shiner & DeYoung, 2013). Furthermore, persistence, inhibitory control, and non-impulsivity have been highlighted as precursors to conscientiousness by a number of studies (Digman & Shmelyov, 1996; Goldsmith et al., 1994; Presley & Martin, 1994; Shiner, 1998; Shiner & Caspi, 2003).

Though Martin and Bridger (1999) did not draw any theoretical connections between distractibility and the Big Five personality traits, studies have found that low distractibility in childhood also predicts high conscientiousness in adulthood (Eisenberg, Duckworth, Spinrad, & Valiente, 2014), as well as found it to strongly, negatively correlate with conscientiousness, specifically industriousness (an aspect of conscientiousness), in adulthood (DeYoung, Carey, Krueger, & Ross, 2016).

Martin and Bridger (1999) did not theorise any connection between their dimensions of temperament and openness, which they instead thought would better be predicted based upon cognitive skills in childhood. However, a number of other studies have found or proposed possible precursors to openness, including attentional focusing and perceptual sensitivity (Goldsmith et al., 1994). Curiosity and exploratory behaviour have also been identified as predictors of later academic achievements and high IQ (Caspi & Shiner, 2006, as cited in Shiner & DeYoung, 2013), which may then predict openness as it is the only Big Five personality trait to be consistently associated with

intelligence (DeYoung, 2010; DeYoung, Grazioplene, & Peterson, 2012). An illustration of these theoretical connections is given in **Figure 1.1**.

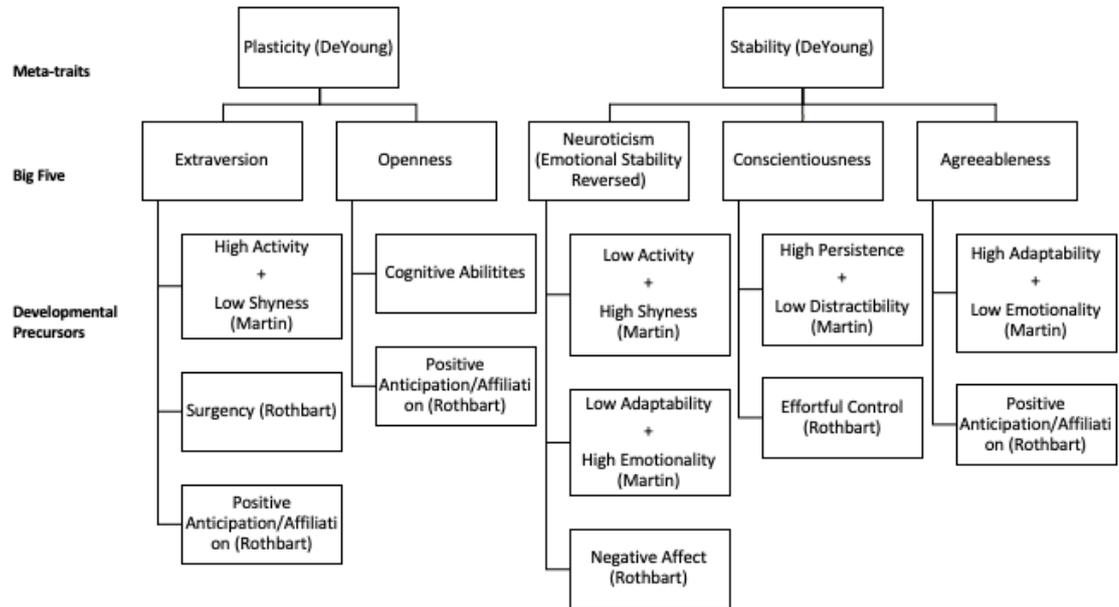


Figure 1.1. Temperament dimensions, according to models by Martin (1988) and Rothbart (2007), as developmental precursors for DeYoung's (2006) Big Five personality traits.

Further supporting the biological basis of temperament, there has been a large number of twin studies from different countries demonstrating the heritability of various temperament traits, showing that identical twins, who share about 100% of their genetic code, are more similar, even if reared in different environments, than non-identical twins, who share about 50% of their genetic code (e.g., Borkenau, Riemann, Angleitner, & Spinath, 2001; Ruf, Schmidt, Lemery-Chalfant, & Goldsmith, 2008).

In a number of studies by Kagan, the role of underlying biology has been further supported, as well as the stability and predictive nature of temperament. In one longitudinal study, for instance, Kagan (1997) demonstrated the stability of temperament, as four-month-olds classed as highly reactive were more likely than others to be classed as fearful during middle childhood. Indeed, none of the highly

reactive infants could be classed as fearless in childhood, compared to others. However, not all of the highly reactive infants grew up to be fearful children, suggesting temperament is more of a predisposing factor, rather than a determining one. In another study, children rated as being inhibited at 3 years old scored significantly lower on scales of impulsivity than other children at the age of 18 years (Caspi & Silva, 1995).

The fact that temperament is expressed before and immediately after birth suggests it develops mostly from biological systems but is modified by input from the environment (Gonzalez et al., 1985). For example, some babies are more active, and others are more sensitive to sensory stimulation. Temperament differences over the first year of life can even be predicted before birth; a high heart rate at 36 weeks' gestation foreshadows less predictable eating and sleeping habits 3 and 6 months after birth (DiPietro, Hodgson, Costigan, Hilton, & Johnson, 1996). New-born infants display consistent differences in behaviour, especially in terms of activity, emotionality, impulsivity, and sociability (Schultz & Schultz, 2008). Specifically, new-borns show evidence of distress and avoidant movements, including anger and frustration, and display individual differences in terms of behavioural inhibition by 6 or 7 months (Rothbart, 2007). Similarly, by 2 or 3 months there are evident differences in approach reactions, measured by displays of positive emotion and body movements (Rothbart, 2007). As children get older, these differences become more pronounced (Schultz & Schultz, 2008). Temperament assessed in infancy can allow researchers to predict a number of later life outcomes, such as teen substance abuse (Tarter, 2002), as well as a number of emotional problems or behavioural problems in childhood (Guerin, Gottfried, & Thomas, 1997; Sanson, Smart, Prior, & Oberklaid, 1993; Zhou, Lengua, & Wang, 2009).

As highlighted in the Introduction, temperament has also been proposed as a predictor of eyewitness performance and suggestibility. Indeed, research has suggested that temperament or behavioural style may impact one's understanding, interpretation, and processing of an event, as well as one's ability to resist suggestive questions when

providing eyewitness testimony (Shapiro et al., 2005). Particularly, Ornstein et al. (1997) theorised that particular elements of temperament affect the perception of eyewitnesses as they witness events take place (activity level, emotionality and persistence), while other elements (adaptability, shyness and distractibility) impact on their performance during forensic interviews. Given that these elements correspond to Martin's (1988) model of temperament, this thesis will use this model. Each dimension will now be discussed in detail, focusing on the potential association between them and with eyewitness performance and suggestibility.

#### *2.4.1.1 Activity*

Activity refers to the degree to which one enjoys being physically active and is considered to be one of the dimensions of temperament (Martin, 1988). The results of some studies have negatively correlated activity with memory accuracy in young children, both during initial interviews and delayed interviews (Chen, 2002; Palmer, Bandt & Chen, 1998, as cited in Bruck & Melnyk, 2004; Shapiro et al., 2005). Melnyk (2002) identified that children characterised as hyperactive by their teachers performed significantly poorer as eyewitnesses; these children typically fidgeted a lot, talked excessively, and/or frequently interrupted others. It is difficult to determine why this connection might exist, but it is thought it may be because children who are more energetic may be more likely to divert their attention to multiple stimuli, and, thus, they focus less of their attention on the subject in question.

If this is the case, highly active children may encode fewer details about a crime when acting as an eyewitness and have less information to provide during an interview. Support for this explanation comes from the findings that children with ADHD or who are described as hyperactive by their teachers perform less well on memory tests (e.g., Melnyk, 2002; Pezdek & Roe, 1995; Warren, Hulse-Trotter, & Tubbs, 1991). Another possible explanation is that highly active children are more impulsive, and, thus, they will be more likely to agree to misleading questions and make false accusations during interviews. Quas et al. (1999) found that more impulsive 3-year-old children were significantly more suggestible to misleading questions than those that were less impulsive, and Roberts and Powell (2001) found the same was true for 5 and 6-year-old children.

*2.4.1.2 Adaptability*

Adaptability refers to how easily one adapts to changes in one's environment. Children who are more adaptable are more easily managed as infants and accept new rules or changes to plans more easily than less adaptable children (Martin, 1988). It seems reasonable that children who struggle adapting to new environments may feel exceptionally uncomfortable being interviewed by the police and so may disclose only a limited amount of information. There are three known studies that have found a significant positive relationship between adaptability and memory ability (Geddie et al., 2000; Greenhoot, Ornstein, Gordon, & Baker-Ward, 1999; Shapiro et al., 2005), and other studies have found no significant relationship (Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Imhoff & Baker-Ward, 1999). As noted in the Introduction, a recent study by Johnston et al. (2021) found that children higher in social flexibility, a construct that combines high adaptability and low shyness, correctly disclosed significantly more transgressions during free recall than other children.

*2.4.1.3 Distractibility*

Distractibility refers to how easily one's attention is diverted. The reasons for why distractibility may be related to eyewitness performance may be more obvious than other traits of temperament; the theory is that more distractible children will encode fewer details regarding a witnessed event and so be more easily confused during questioning, making them more suggestible to misleading questions, as well as have less information to provide in the first instance. Furthermore, they will presumably pay less attention during a forensic interview and so may misinterpret questions. One known published study (Benedan et al., 2020) and two unpublished dissertations (Melnyk, 2002; Purdy, 2001) have found a significant, negative relationship between distractibility and eyewitness performance, while other studies have found no significant relationship (Alexander et al., 2002; Blackford & Shapiro, 1999; Burgwyn-Bailes et al., 2001; Brown et al., 1999, as published in Bruck & Melnyk, 2004; Clarke-Stewart et al., 2004; Geddie et al., 2000; Imhoff & Baker-Ward, 1999; Scullin, 1997). In Benedan et al's (2020) study, children with higher attentional ability reported more information overall and were significantly less susceptible to misleading questions. Part of the reason for the difference in findings is the various ways in which distractibility is measured across the studies. For example, some have used the

distractibility subscale from Martin's TABC (teacher-version), while other studies have used composite scores from various subscales in other measurements. In Brown et al. (1999, as published in Bruck & Melnyk, 2004), teachers were simply asked to rate a child's attentiveness.

#### *2.4.1.4 Emotionality*

Emotionality refers to the intensity with which one's feelings are expressed. Young children who are more emotional will, for example, stomp their foot, scream, or cry with more intensity than children who are less emotional when they get into trouble or do not get their way (Martin, 1988). Emotionality has been found to have a significant effect on eyewitness performance by a much larger number of studies than any other temperament trait (Chae & Ceci, 2005; Chen, 2002; Geddie et al., 2000; Greenhoot et al., 1999; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Scullin, 1997; Shapiro, et al., 2005), though the direction of the relationship is inconsistent and explanations are unclear. It may be that children who are typically more emotional will also be more emotional when witnessing a criminal event and this will impede the process of encoding (Yerkes & Dodson, 1908). If this was the case then it may depend how emotionally arousing the crime is (for example, a brutal attack will be more emotionally upsetting to a witness than a non-violent robbery) and may also depend on who the victim of the crime is (presumably a child will be more emotional if the victim is a family member or the child themselves).

Of course, this makes the effect of emotionality difficult for researchers to study, as research projects typically aim to cause as little emotional upset to participants as possible. While, in most cases, emotionality has been found to be negatively related to eyewitness performance (Chae & Ceci, 2005; Geddie et al., 2000; Greenhoot et al., 1999), some research has found the opposite – that high emotionality significantly correlates with lower suggestibility (Chen, 2002; Scullin, 1997). This makes things more confusing regarding how emotionality and eyewitness performance interact, but it may be because although children who are more emotionally intensive are less likely to remember a great deal of details during a stressful event, they are also less likely to be coerced into making a false statement.

Interestingly in Johnston et al.'s (2021) study, though reactivity was unrelated to responses from children, it significantly impacted the behaviour of the interviewers. Indeed, more reactive (or emotional) children received a significantly higher number of prompts. The authors explained that more emotional children tend to give shorter answers, thereby requiring further guidance from interviewers. There is a potential danger, therefore, that interviewers may become frustrated in response to challenging behaviour from children and be more likely to ask whatever questions they think will help resolve the interview more quickly. In support of this, as mentioned in the Introduction, Gilstrap and Papierno (2004) found that interviewers were significantly more likely to ask leading questions to shyer children. This is another example of how interviewers may respond to challenging behaviour by straying from guidelines.

#### *2.4.1.5 Persistence*

Persistence, another temperament dimension, is linked to task performance (e.g., Sandelands, Brockner, & Glynn, 1988), and is defined as the extent to which one will continue with an activity, a desire, or a frame of mind, even when challenged (Martin, 1988). Blackford (2000) and Burrell, James, and Shapiro (1999, as cited in Chen, 2002) found that children who were more persistent remembered more peripheral details about a crime than those who were less persistent. According to Chen (2002), persistence may be important both for remembering details during the witnessing of an event and for the process of an interview. This helps also to explain findings by Greenhoot et al. (1999), when more persistent children were less likely to answer 'yes' to misleading questions than children who were less persistent.

#### *2.4.1.6 Shyness*

As discussed in the Introduction, shyness refers to the degree to which one withdraws from social situations (Martin, 1988). Children who are more sensitive to the social aspects of an interview may report less information and may be more easily swayed by suggestions from the interviewer. In support of this theory, Kagan (1994, as cited in Schacter, Kagan, & Leichtman, 1995) found that inhibited children were reluctant to oppose requests by adults. Gordon et al. (1993) found that lower levels of shyness predicted a larger amount of information reported by 3-year-olds in response to open-ended recall prompts, but not for 5-year-olds. As outlined in the Introduction, a number of studies have found shyer children to be less accurate as witnesses and to be more suggestible to misleading questions (e.g., Endres et al., 1999; Gilstrap &

Papierno, 2004; Roebers & Schneider, 2001). Muir-Broaddus, King, Downey, and Petersen (1998) found no relationship between shyness and eyewitness recall, but this was on a small sample of 36 5- to 7-year-olds. A currently unpublished study by Cotterill (2017) was the first study to examine the connection between shyness and eyewitness performance by using shyness questionnaires that children (9 to 12 years old) self-completed, rather than using ratings by parents or teachers. The results showed that shyer children were less confident as witnesses than less shy children, as well as that they were more suggestible to leading questions. Importantly, studies have all looked at different ages but there is yet no evidence of an interaction effect between age and shyness on suggestibility, with the exception of Gordon et al.'s (1993) paper.

Though this thesis is focusing on eyewitness descriptions, this area of research ties in with studies showing that introverts typically perform poorer than extraverts on facial recognition tasks, independent of general cognitive abilities (e.g., Lander & Poyarekar, 2015; Li et al., 2010; Wang, Li, Fang, Tian, & Liu, 2012). The results show that introverts are poorer at extracting social stimuli, such as faces, but are matched to extraverts when it comes to recognising non-social stimuli. It is theorised that this is due to extraverts being more socially skilled than introverts, being less likely to experience social phobia (Bienvenu et al., 2001; Davis et al., 2011). Therefore, it is not extraversion in general, but experience of social interactions that improve facial recognition performance. Participants high in shyness or with an autism spectrum disorder diagnosis also may then display facial recognition problems if there is a history of social interaction difficulties (e.g., Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008). It may be the case that this lower recognition accuracy is due to weaker encoding. For example, participants higher in social phobia may be less likely to look directly at the face of targets during eyewitness studies. If this is the case, negative impacts should be observed in eyewitness description studies regarding social details, and not just eyewitness identification studies.

#### *2.4.1.7 Measuring temperament*

In previous eyewitness studies, temperament is generally measured by having parents or teachers rate the temperament traits of the children in their classrooms via either

the Behavioural Style Questionnaire (McDevitt and Carey, 1978) or the Temperament Assessment Battery for Children (Martin, 1988). The problem is that, unlike cognitive abilities, temperament cannot be objectively measured, especially when it comes to young children who are too young to self-complete questionnaires. Studies continue to determine the temperament of child participants via ratings by their teachers, even though research consistently shows that teachers have poor reliability when it comes to rating the personality traits of their students (e.g., Eisenberg, Shepard, Fabes, Murphy, & Guthrie, 1998; Measelle, Ablow, Cowan, & Cowan, 1998; Rudasill et al., 2014; Spooner et al., 2005), meaning that it is arguable to what extent any of the previous studies have truly measured temperament traits; this may explain why results in the area are so mixed. It is also possibly the reason as to why there has been little progress in regard to the connection between temperament and eyewitness performance in the last ten years. Notably, evidence suggests the opinions of teachers do not converge with those of other raters when it comes to interrater consistency on problem behaviour or personality traits (e.g., Achenbach, McConaughy, & Howell, 1987; Stanger & Lewis, 1993; Youngstrom, Loeber, & Stouthamer-Loeber, 2000). This will be explored in greater detail in Chapter Three. In sum, however, it would seem that a more accurate way of measuring temperament is required before studies can really begin to understand the influence that temperament traits may have on eyewitness performance.

#### **2.4.2 Summary: Internal factors of suggestibility**

When it comes to internal factors of suggestibility, there remains a lot to learn. Still, there is enough evidence to indicate that individual differences are a significant factor of eyewitness performance. Some of the factors may seem more obvious. For example, it may be clear as to why children with higher intelligence, verbal ability, and better memory perform better as eyewitnesses. However, other factors, such as traits of temperament, require further exploration and understanding in regard to how they may make a difference to eyewitness accuracy. Internal factors cannot be changed to improve accuracy, but a better understanding of individual differences may offer suggestions regarding how system variables may be changed. For example, if it was found that less adaptable children struggle more to describe a criminal event, it may be that investigators will receive more descriptors if they conduct the interview at the

child's house rather than at a police station or allow for the child to bring their favourite toys to the interview, therefore meaning the child has less environmental changes to adapt to.

## **2.5 Perceptions of children's credibility as eyewitnesses**

Temperament may also impact how children are perceived as witnesses, rightly or wrongly, independent of their actual memory performance. This section will explore the perceptions that jurors have of child eyewitnesses and some of the factors that influence those perceptions, including those relating to temperament.

### **2.5.1 Effects of witness age**

While child witnesses are generally viewed as being honest by jurors (e.g., Nunez et al., 2010; Ross, Dunning, Toglia, & Ceci, 1990; Ross, Jurden, Lindsay, & Keeney, 2003), they are also typically perceived as being less accurate than adult witnesses (e.g., Bottoms & Goodman, 1994; Goodman, Golding, Helgeson, Haith, & Michelli, 1987; Nikonova & Ogloff, 2005; Ross, et al., 1990; 2003). Wright, Hanoteau, Parkinson, and Tatham (2010) found that perceived memory reliability increased with the age of the witness when participants heard statements by children (between 3 and 18 years old) who had witnessed an act of abuse. The difference in perceived reliability was considerable between 3 and 6, then much smaller after the age of 6. This is consistent with the results of other studies finding that perceived reliability increases with age (e.g., Holcomb & Jacquin, 2007; Nightingale, 1993; Pozzulo & Dempsey, 2009). Participants were also asked to rate how honest they found the witnesses to be. Honesty levels of female witnesses increased with age from 3 to 18, whereas honesty levels for males increased until the children were 6 years old, then decreased until 18.

It should be noted that different results are often found depending on the type of crime and on whether the witness is also the victim in the case. In the instance of the child witness also being a victim of sexual abuse, the child witness has been found to be perceived as more reliable than the adult (e.g., Goodman et al., 1987; Pozzulo & Dempsey, 2009). In cases of sexual abuse, younger children may be seen as more credible because they lack the knowledge necessary to make false accusations (Bottoms & Goodman, 1994). Overall, older children and adults are typically viewed as

more reliable than younger children (especially those under 6 years old), but younger children can be viewed as more reliable in certain situations, depending on gender, the crime, and the type of witness that the child is (i.e., bystander-witness or victim-witness).

Effects of witness-age on juror interpretations are not always found (e.g., Golding, Sanchez, & Segó, 1997; Greenwald, Pratkanis, Leippe, & Baumgardner, 1986; Luus, Wells, & Turtle, 1995; McCauley & Parker, 2001). In some cases, there are results inconsistent with previous findings. For example, Bidrose and Goodman (2000) is one of the few examples of real-world research. In the case within this study, four girls (8 to 15 years old at time of report) testified about sexual exploitation experiences in New Zealand. The oldest girl made more mistakes (determined by the presence of an allegation, but no supporting evidence) and omission errors (determined by the presence of evidence, but no allegation) during police interviews than the younger girls. This is contrary to findings that older child witnesses typically provide more accurate information than younger child witnesses, as previously discussed.

Additionally, the oldest girl's testimony was not considered any less reliable by jurors than those of the younger girls, despite having made more mistakes and omission errors, even though, as already mentioned, younger children are usually viewed by mock-jurors as more reliable than older children in simulated sexual abuse cases. This real-life case, therefore, though clearly has a very small sample size, differs from general findings in two significant ways. It is clear that more research has to be done on juror interpretations, both in child sexual abuse cases and other cases involving child witnesses.

### **2.5.2 Effects of interviewing technique**

Suggestive interviewing techniques that have detrimental effects on eyewitness accuracy may be used to discredit their testimony (e.g., Goodman, Quas, Bukley, & Shapiro, 1999). Though many legal advisors stress that asking questions in such a manner is unwise, there are still those who continue to argue that asking questions containing presupposition is a normal and effective technique for verifying doubtful information (Wheatcroft & Woods, 2010). When such leading questions are used with children in courtrooms, it can be difficult for jurors to determine whether the child is

answering in the way they do because it is the correct answer or because they are simply complying with the lawyer's suggestion. This can therefore decrease perceived credibility (e.g., Karla & Heath, 1997; Kebbell, Evans, & Johnson, 2010; Tubb, Wood, & Hosch, 1999).

### **2.5.3 Effects of witness confidence and shyness**

Some research has found that leading questioning can decrease witness confidence (Wheatcroft, Wagstaff, & Kebbell, 2004), but findings in this area are generally mixed (Wheatcroft, 2002). This is a point of concern because eyewitness confidence remains one of the most persuading factors for jurors (Nicholson, Yarbrough, & Penrod, 2014). It is commonly thought that the more confident the witness, the more accurate they are, even though the connection between confidence and accuracy is mixed in eyewitness descriptions (Wheatcroft et al., 2004), though it is a more consistent correlation than the one in eyewitness identification studies (e.g., Wixted, Mickes, Clark, Gronlund, & Roediger III, 2015). The researcher is not aware of any studies that have focused specifically on children regarding the relationship between leading questioning and perceived confidence. On top of this, the researcher is not aware of any studies that have investigated the connection between witness shyness and perceived credibility. Since shyer individuals tend to have a lack of confidence in their everyday life and are more likely to question themselves, as well as rely on others for information (Crozier, 2000), it would make sense if they were perceived as less confident and therefore less credible to potential jurors based upon the findings of Wheatcroft et al. (2004). It should be noted that other temperament traits have previously been linked to perceived credibility. For example, Golding et al. (2003) found that too little or too much emotion from a child witness could negatively affect their credibility.

### **2.5.4 Effects of juror gender**

Typically, when it comes to sexual abuse cases, female jurors are more likely to find the victim-witness credible than male jurors and are also more likely to vote for a conviction (e.g., Devine & Caughlin, 2014; Schutte & Hosch, 1997). This is also true in the case of child sexual abuse victims (Bottoms & Goodman, 1994; Bottoms, Golding, Stevenson, Wiley, & Yozwiak, 2007; Quas, Bottoms, Haegerich, & Nysse-Carris, 2002). This may be because women are usually more empathic towards children than men

(e.g., Christov-Moore et al., 2014; Mestre Escrivá, Samper García, Frías Navarro, & Tur Porcar, 2009), or perhaps because women are more likely to be sexual abuse victims themselves (Office for National Statistics, 2020) and so are more likely to sympathise with the victim of the case. Golding, Bradshaw, Dunlap, and Hodell (2007) found a similar finding when investigating jury decision making, as opposed to juror decision making. In the case of a 6-year-old female who claimed to have been a victim of sexual abuse, mock juries with a female majority were more likely to convict. Interestingly, females who originally voted not guilty were also more likely to change their verdict to guilty if they were within a female majority group.

### **2.5.5 Effects of juror age**

Juror age has been largely ignored in the literature. Though there is some evidence to suggest that older jurors may be more likely than younger jurors to give a verdict of not guilty (e.g., Mossière & Dalby, 2007), the results are mixed and are dependent upon the characteristics of the crime and the defendant (Anwar, Bayer, & Hjalmarsson, 2012). The researcher is not aware of any findings in regard to a connection between juror age and perceived credibility of child witnesses.

### **2.5.6 Effects of juror personality**

Though few recent studies have investigated the influence of juror personality, largely due to small effect sizes in early studies (Greene et al., 2002, as cited in Clark, Boccaccini, Caillouet, & Chaplin, 2007), there is evidence that certain traits have an impact on juror attitudes. For example, Haegerich and Bottoms (2000) found that jurors higher in empathy rated a child defendant (accused of killing their abusive parent) as being significantly less guilty than those lower in empathy. A meta-analysis of 20 studies by Narby, Cutler, and Moran (1993) found strong evidence for a positive connection between authoritarianism and guilty verdicts. Seeing as how one's personality impacts their social behaviour and political ideologies (e.g., Burton, Plaks, & Peterson, 2015; Carney, Jost, Gosling, & Potter, 2008; Hirsch, DeYoung, Xu, & Peterson, 2010; Sibley, Osborne, & Duckitt, 2012; Xu, Mar, & Peterson, 2013), it makes sense that it would also impact their behaviour as jurors. For instance, those higher in authoritarianism are more likely to be politically right-leaning, as well as to be more angered by the violation of certain moral foundations, specifically loyalty, authority, and purity (Haidt, 2012). Therefore, it is perhaps not surprising that, in certain cases,

those higher in authoritarianism are more likely to render less lenient verdicts as jurors (McGowen & Glen, 1982). Though this area of research has been largely abandoned due to small effect sizes (Clark et al., 2007), it perhaps warrants further investigation given the emergence of the Big Five (see Chapter Three), which has become the dominant approach to understanding personality in recent years (e.g., Funder, 2019).

### **2.5.7 Summary: Perceptions of children's credibility as eyewitnesses**

It is clear that eyewitness testimony can influence jury decision making. Conviction rates are typically higher in cases involving eyewitnesses than those without (e.g., Greene, 1988; Spanos, Myers, DuBreuil, & Pawlak, 1992), especially in the absence of jury instructions (Greene, 1998). There is considerably more research investigating how witness characteristics influence the perception of jurors, such as witness age, compared to juror characteristics. This is likely due to the cost and time required to investigate multiple juror characteristics (Nicholson et al., 2014). As discussed, there is evidence that juror gender can affect perceived reliability of child witnesses (e.g., Bottoms & Goodman, 1994), but other factors such as juror age and personality require further investigation. There is also some suggestion that the types of questions asked and the shyness level of the child witness may impact on how they are perceived by jurors, as these factors may affect how confident they come across (Wheatcroft et al., 2004), and confidence remains a main persuading factor for jurors (e.g., Nicholson et al., 2014). Importantly, the previously mentioned studies have largely focused on perceptions of child witnesses in sexual abuse trials. An exception to this is a study by Karla and Heath (1997), which examined perceptions of child witnesses during a murder trial. There is less research on perceptions of child witnesses in other scenarios, such as to a theft.

## **2.6 Research objectives**

Before investigating the impact of temperament on eyewitness performance or mock-juror perceptions of those witnesses, it was first necessary to develop a new method to assess temperament. The project, therefore, had three main objectives that were each investigated in the form of one individual study for each objective. The first objective was to develop a new method for determining temperament traits in 4 to 8-year-old children, using data that is derived from the views and experiences of young children themselves, rather than relying on ratings from external examiners. The

second objective was to investigate how interviewing techniques and temperament affect the eyewitness performance of child eyewitnesses. The third objective was to examine how the interviewing techniques and shyness of a child witness may affect the perceptions and decision making of jurors, as well as the personality trait of the jurors themselves.

### **2.6.1 Research objective 1**

The current difficulty with measuring temperament in young children is that ratings from various examiners are typically used, but these ratings are not usually consistent with one another. In order to test the new method that has been created by the researcher, there were three research questions to be addressed: (1) do 4 to 8-year-old children have consistent and meaningful understandings regarding their own temperament characteristics, (2) do these findings differ from child to child, and (3) are these differences stable over time?

### **2.6.2 Research objective 2**

The association between temperament and child eyewitness performance has produced mixed results when temperament is measured with ratings from external examiners and the area of research has been largely ignored by researchers for some time now. Given this, the second research objective was to investigate if there would be such an association when temperament is measured with data derived from the views and experiences of the children themselves. It has been suggested that eyewitness accuracy may decrease in the presence of closed-ended or misleading questions compared to when in the presence of open-ended questions (e.g., Morgan et al., 2013; Wheatcroft et al., 2004). The research aimed to examine if these negative effects would be heightened when applied to children with certain temperament qualities.

### **2.6.3 Research objective 3**

It has been found that misleading questions can lower the confidence of eyewitnesses (Wheatcroft et al., 2004), and also that they can have a negative influence on the perceived reliability of jurors (e.g., Karla & Heath, 1997; Kebbell et al., 2010; Tubb et al., 1999). Therefore, an aim of the research was to investigate if children who undergo closed-ended questions and misleading questions are perceived as less credible than those who undergo open-ended questions. Additionally, shyer individuals tend to be

*Chapter 2: Literature review*

more likely to doubt themselves and have less confidence in their everyday life (Crozier, 2000). An additional aim of the research, therefore, was to examine if shy children will be perceived as less confident in their responses and therefore as also less credible. There has also been evidence to suggest that juror personality can impact decision making, but this area of research has been largely neglected (Greene et al., 2002). A final aim of the research, therefore, was to understand if juror personality characteristics impact perceptions of child-witness credibility.

## **Chapter 3: Study one – The temperament assessment tool for children (4 to 8 years old): a self-report technique**

---

### **3.1 Introduction**

Temperament is a precursor of personality and refers to behavioural traits that explain how one behaves (Hofstee, 1991; Rothbart, 2012), though it is important to note that there is no universally agreed upon definition (De Pauw, 2017). In the context of this study, temperament refers to the behavioural style of the child and how that child responds to its environment. Temperament traits are thought to have a genetic origin or at least to be determined early in life (Braungart, et al., 1992; Buss & Plomin, 2014; Martin, 1988; McDevitt & Carey, 1978; Mervielde et al., 2005; Posner & Rothbart, 2000). These traits then become increasingly stable as one ages (Caspi, Roberts, & Shiner, 2005).

According to Rothbart and Bates (2006), behavioural responses vary across individuals based upon reactivity (i.e., the arousal and excitability of motor and affective responses) and the strategies one has for regulating these reactions (i.e., attentional control and inhibitory mechanisms). Personality, on the other hand, is wider in scope, and refers also to a pattern of beliefs, feelings, thoughts, and values (Rothbart & Bates, 2006). Via the study of child temperament, researchers have been able to examine the development of empathy (e.g., Mark, IJzendoorn, & Bakermans-Kranenburg, 2002; McDonald & Messinger, 2011; Volbrecht, Lemery-Chalfant, Aksan, Zahn-Waxler, & Goldsmith, 2007), psychopathology (e.g., Earls & Jung, 1987; Muris & Ollendick, 2005), and adult personality (e.g., Rothbart, Ahadi, & Evans, 2000), just to name a few examples.

Since Ornstein et al. (1997) theorised that particular elements of temperament affect the perception of eyewitnesses as they witness events take place (activity level, emotionality and persistence), and other elements (adaptability, shyness and distractibility) impact on their performance during forensic interviews, the effect of temperament on child witnesses in eyewitnesses studies has been studied extensively (for reviews, see Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020). However, as

explained in the previous chapter, the impact is inconsistent. It remains to be understood if the difference in results is due to the different ages of participants, the temperament measures, and/or the different stimuli used across studies (Shapiro, 2006).

One possible problem, as described in the previous chapter, is that studies have often determined the temperament of child participants via ratings by teachers, even though research has consistently shown teachers have low consistency with other raters when it comes to rating the personality traits of their students (e.g., Eisenberg et al., 1998; Measelle et al., 1998; Rudasill et al., 2014; Spooner et al., 2005). Notably, evidence suggests the opinions of teachers do not converge with those of parents nor with children's self-reports (e.g., Achenbach et al., 1987; Stanger & Lewis, 1993; Youngstrom et al., 2000), nor even with other teachers or teacher aides (Martin, 1988). Young children's self-reports are rarely used by researchers, perhaps due to concerns over the reliability and validity of their reports (e.g., Edelbrock, Costello, Dulcan, Conover, & Kalas, 1985; Shiner & DeYoung, 2013). Research has shown, however, that young children can provide reliable and valid self-reports on temperament so long as age-appropriate techniques are used (e.g., Hwang, 2002; Measelle, John, Ablow, Cowan, & Cowan, 2005; Roth, Dadds, & McAloon, 2004).

### **3.1.1 Measurements of temperament**

As mentioned, temperament in young children is often measured via caregiver reports and observations in labs or at home (Shiner & DeYoung, 2013). In lab studies, children's behavioural and physiological responses to stimuli are recorded in a controlled environment (Goldsmith & Gagne, 2012). However, the range of behaviours that can be observed is limited and lab studies generally only measure a temperament dimension within one context. It is possible for a child to be sociable, for example, in one situation, but not in others. Therefore, a general evaluation of a child's temperament is more valuable to researchers so that they can understand what that child is like on average, not just in one context. Furthermore, not all temperament dimensions are well suited to this type of measurement. For example, a shy or fearful child may be uncomfortable completing any laboratory task, thus influencing results (Hwang, 2002; Shiner & DeYoung, 2013). When laboratory tasks are used in

conjunction with other methods, they can provide construct validation of other measures.

One such method is observations made at home (Goldsmith & Gagne, 2012). This involves trained observers observing children in their natural environments and coding their behaviour. This type of assessment is often described as the gold standard (e.g., Kagan & Fox, 2006; Majdandzic, van den Boom, & Heesbeen, 2008). However, the high costs associated with preparing, implementing, and coding observations mean that this type of assessment is not always possible for researchers. Furthermore, multiple observations are required in order to suitably sample a child's behaviour (Hwang, 2002). For these reasons, eyewitness psychologists typically use caregiver ratings to assess temperament, as opposed to the two previously mentioned assessment techniques (e.g., Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Chae et al., 2014, 2018; Geddie et al., 2000; Gordon et al., 1993; Greenhoot et al., 1999; Imhoff & Baker-Ward, 1999; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Shapiro et al., 2005). Caregiver reports are inexpensive, easy to administer and take advantage of caregivers' extensive knowledge, including of a child's behaviour in situations in which a researcher could not possibly observe (Hwang, 2002).

A number of different questionnaires have been developed to assess temperament in young children, usually via parent or teacher ratings. McDevitt and Carey (1978) created the Behavioural Style Questionnaire (BSQ) for children aged 3 to 7 years old based upon Thomas and Chess' (1977) nine-dimension model. Martin (1988) believed that these nine categories overlapped and so constructed the Temperament Assessment Battery for Children (TABC), instead focusing only on six dimensions (activity, adaptability, ease of management through distraction/distractibility, emotional intensity/emotionality, persistence, and shyness (labelled by Martin as 'approach/withdrawal')). The TABC consists of a clinician form, parent form, and a teacher form. The rationale for the three instruments was to limit the problems that can arise from using only ratings from one rater from one setting (i.e., the rating from

each rater reflects more than just the behaviour of the child, but also the setting and context of the behaviour; Martin, 1988).

For the purposes of data collection, however, a number of researchers only use either the parent or teacher form (e.g., Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Davis & Carr, 2002; Downer & Mendez, 2005; Ewing & Taylor, 2009; Geddie et al., 2000; Gordon et al., 1993; Greenhoot et al., 1999; Imhoff & Baker-Ward, 1999; Machida, Taylor, & Kim, 2002; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Pellegrini & Bartini, 2000; Pellegrini, Galda, Flor, Bartini, & Charak, 1997; Shapiro et al., 2005). The problem with this, therefore, is that researchers may use Martin's (1988) TABC in a way that causes the very problems he was attempting to avoid, specifically the use of ratings from only one rater within one setting and context, and this may not give an overall picture of the child's temperament.

Rothbart and colleagues have created a number of temperament tools which are typically linked to neurobiological functioning, in order to address the why of behaviour, rather than just the how (Rothbart, 1991). According to Rothbart and Bates (2006), temperament is 'the affective, activational, and attentional core of personality' (p. 100, as cited in Eisenberg et al., 2014). Most notable of the measurements, the Children's Behaviour Questionnaire (CBQ; Rothbart, Ahadi, & Hershey, 1994) has 15 scales rated by parents, expanding on the three broad temperament dimensions proposed by Ahadi and Rothbart (1994). Surgency includes activity level, high intensity pleasure, low shyness, smiling and laughter, approach/positive anticipation, and impulsivity. Negative affect includes anger/frustration, fear, discomfort, sadness and soothability. Effortful control includes attentional focusing, inhibitory control, perceptual sensitivity, and low intensity pleasure. As noted from the different models discussed in the Literature Review, it is clear that there is still some disagreement over the behavioural dimensions of temperament, or at least over the delineations of those dimensions. Even now, there is no clear consensus on a single model of temperament (De Pauw, 2017).

More recently, researchers have aimed to examine how childhood temperament may shape the course and quality of social development and its relationship to behavioural problems in adulthood (Ball, Pelco, Havill, & Reed-Victor, 2001). A revised version of the TABC (TABC-R: PF; Ball et al., 2001) uses only parent-ratings and focuses on four dimensions of temperament, including activity, inhibition (i.e., shyness), positive vs. negative mood (i.e., emotionality), and persistence/attentiveness. That said, research still often uses the original TABC, especially when it comes to researchers examining the performance of eyewitnesses (e.g., Burgwyn-Bailes, et al., 2001; Chae et al., 2014, 2018; Shapiro et al., 2005). This is because of the direct associations between at least some of those six temperament traits (activity, adaptability, distractibility, emotionality, persistence, and shyness) and with memory and suggestibility (e.g., Geddie et al., 2000; Gordon et al., 1993).

That is not to say that these six temperament traits are only of interest to researchers in the field of eyewitness psychology. For example, the original TABC has been used by researchers examining the impact these six temperament dimensions have on gender differences in mathematics (Davis & Carr, 2002), father parenting styles (Downer & Mendez, 2005), and on bullying and victimisation (Pellegrini & Bartini, 2000). There is, therefore, consistent interest in these six characteristics of temperament, even if they are not as encompassing models of temperament or as widely used by developmental psychologists as Rothbart's model.

Children aged 7-years-old and older are able to reliably report their behavioural characteristics via traditional Likert-scale response items (e.g., Danielson & Phelps, 2003; Scanlon & Ollendick, 1986; Spooner et al., 2005), and so temperament self-reports are available for children over the age of 9 (Ellis & Rothbart, 2001; Luby, Svrakic, McCallum, Przybeck, & Cloninger, 1999). The researcher knows of only one instrument (Simonds, 2006) available for children younger than this; Simonds (2006) used a computerised Likert-scale to measure self-reports of temperament in 8- to 10-year-olds (i.e., the Temperament in Middle Childhood Questionnaire; TMCQ). The

TMCQ has translated the items from Rothbart et al.'s (1994) CBQ for self-reports. The items were read by an animated character on the screen. In a sample of 58 children, internal consistencies revealed that 11 out of the 16 subscales were at least satisfactory (Simonds, 2006). With a larger Swedish sample ( $N = 157$ , 8.8-11.1 years), only one subscale reported an internal consistency below .70 (Nystrom & Bengtsson, 2017). Overall, while some of the research seems promising, it is clear that more research needs to be conducted to assess whether young children's self-reports can provide reliable and valid information in regard to their own temperament traits.

### **3.1.2 The difficulty with caregiver questionnaires**

A problem with ratings from only one rater is that a child's behaviour may be dependent upon their environment. For example, a child may not act the same within the classroom as they do at home or on the playground. Another problem with ratings from external examiners is that the rater has their own points of views and only their own experiences to rely upon. For example, a teacher may consider one student to be particularly hyperactive or distractible if that child is less focused on their work than the other children in their class, but that does not necessarily mean that child is highly hyperactive or distractible when compared to other children at large. Crossman (2001) noted that teacher ratings of child competences (e.g., academic self-esteem, social competence) were correlated with each other, suggesting that teacher ratings were reflective of general perceptions, rather than careful considerations for each characteristic. Problematically, when the ratings of one rater are compared to those of another, interrater consistency is not always high, even when both raters have used the same scale, thus casting doubt over the accuracy of at least one of the sets of ratings.

To illustrate these points, evidence suggests the opinions of teachers especially do not converge with those of other raters when it comes to interrater consistency on problem behaviour or personality traits (e.g., Achenbach et al., 1987; Stanger & Lewis, 1993; Youngstrom et al., 2000). Achenbach et al. (1987) conducted a meta-analysis of 119 such studies and found that mean correlations between parents of the same child were .60, whereas there was only a .27 correlation between parent and teacher reports. Crooks and Peters (2005) found a similar correlation (.29) between parent and

teacher ratings when specifically assessing ratings of shyness and anxiety in 3 to 5-year-old children. Studies show that interrater correlations are especially low when it comes to ratings of shyness (e.g., Crooks & Peters, 2005; Rudasill et al., 2014), sometimes as low as .17 (Spooner et al., 2005).

Indeed, in the study by Spooner et al. (2005), one third of children who classified themselves as shy were not rated as shy by parents or teachers. The authors suggest three possible explanations for this. The first is due to the characteristics of the child. For example, if a child who self-reports as shy has an expressive vocabulary then an adult may incorrectly assume the child is naturally sociable (Spooner et al., 2005). The second is due to the characteristics of the rater. For example, a parent or teacher might not have had the opportunity to observe a child in the necessary range of social situations in order to make accurate ratings. Wang and Kemple (1993, as cited in Spooner et al., 2005) found that parents were more likely to rate their child's level of shyness based upon their interactions with strangers, whereas teachers were more likely to rate a child's level of shyness based upon their interactions with peers. However, even shy children can socially interact with familiar peers (Asendorpf & Meier, 1993). It could also be that less shy parents or teachers are less perceptive when it comes to identifying indicators of shyness in children (Spooner et al., 2005). Third, the scale completed by parents and teachers (the CBQ; Rothbart, et al., 1994) was not the same as those completed by the children (Crozier's Shyness Questionnaire, 1995). While the CBQ's items were developed by the researchers (Rothbart et al., 1994), Crozier (1995) had children themselves write down their own experiences of shyness and collated the most common ones to be items on his questionnaire, thus ensuring the items matched directly with the internal states and contexts experienced by children (e.g., being asked to speak during class). Since some children who identify as shy may not behave in a way that makes their anxieties clear, it may be that asking observers about a child's external evidence of shyness does not allow researchers to tap into aspects of shyness that are identified by self-reports.

### **3.1.3 Children's self-reports**

Despite the issues with caregiver ratings of children's temperaments, there are also a number of problems associated with children's self-reports. One problem is that the

understanding of one's own emotions and behaviours (i.e., self-concepts) increases with age. For example, self-concepts in children younger than 3 years are likely to consist of mostly visual information, such as hair colour and eye colour (e.g., Damon & Hart, 1988, as cited in Hwang, 2002). Self-concepts from children older than this are likely to include hobbies and skills (e.g., Harter, 1999; Keller, Ford, & Meacham, 1978). Children below the age of 7 years, however, are unlikely to understand that a person's psychological characteristics can differ based upon the situation (e.g., Harter, 1986, 1999). Measures can tackle this issue by including items that cover several contexts, rather than asking children for general descriptions or evaluations (Hwang, 2002).

Paulhus (1991) outlined three biases stemming from young children's self-reports. First, young children have a tendency to say yes in order to be agreeable. Measures can avoid this problem by not giving children yes or no options. Second, young children have a desire to present themselves in a socially acceptable manner. Hwang (2002) found that children were unlikely to identify with sadness (especially boys) or shyness, and that they were likely to give high ratings of smiling or laughter. She cautioned that this might have been the result of social desirability. Lastly, young children have a tendency to choose either end of a Likert-scale. Chambers and Johnston (2002) found that 5- and 6-year-old children responded to Likert-scale items more extremely than older children and demonstrated less differentiation. However, even children as young as 4.5 years old can reliably report their behavioural characteristics via other methods (e.g., Hwang, 2002).

One such method is the Berkeley Puppet Interview (BPI; Ablow & Measelle, 1993). The BPI builds on the work of Eder (1990) by having two identical puppets make opposing statements about themselves (e.g., 'I'm shy when I meet new people' and 'I'm not shy when I meet new people'), and then asking children to select which one they agree with. The interviews are typically video-taped and later coded on a seven-point scale depending on the enthusiasm with which the child agrees with the statement. For example, if the child repeats the words of the non-shy puppet verbatim then that is two points, but if the child adds words like 'really' or 'very' to the puppet's statements

then that would be one point because it is suggesting even more outgoingness than the original statement. Previously, this method has successfully been replicated and used to measure young children's self-reports of their academic, psychological, social, and emotional states (Brown, Mangelsdorf, Agathen, & Ho, 2008; Measelle et al., 1998), including levels of aggression (e.g., Eder, 1990). Internal consistency for each subscale ranges between .62 and .78, and test-retest reliability over one and two-year intervals ranges from .24 to .58 (Measelle et al., 1998). Additionally, Measelle et al. (2005) found levels of consistency and differentiation that approached those of a college age sample when collecting self-reports of the Big Five personality dimensions by children aged 5 to 7 years. Overall, there is strong evidence to support the reliability and validity of young children's self-reports so long as appropriate measures are used.

Building on this, in an unpublished doctoral thesis, Hwang (2002) drew items from the CBQ (Rothbart, Ahadi, Hershey, & Fisher, 2001) to adapt the BPI and measure children's self-reports of temperament. In this study of 100 children between the ages of 4-7 years, internal consistency averaged at .55. The instrument reported adequate stability over a seven to 10-day period, averaging at .72. However, interrater consistency overall only ranged from poor to moderate (.00 to .48). Similarly, Roth et al. (2005) adapted the BPI to measure social desirability, sociability, shyness, emotionality, and soothability in 4- to 5-year-old children ( $n = 77$ ) by using items from the Colorado Childhood Temperament Inventory (Rowe & Plomin, 1977). The internal consistency scores of the dimensions ranged from .47 to .64. Test-retest reliability scores for 10 of the children for the dimensions ranged from .31 to .86. It should be noted that a sample size of 10 children for the test-retest reliability is low, especially given that there was not a high existing level of consistency with the original sample (Bujang & Baharum, 2017). It should also be noted that an issue with the BPI method is that there will always be a degree of subjectivity required when scoring the child's responses to the puppets, even though interrater consistency is typically high (e.g., .89; Graham, Ablow, & Measelle, 2010).

#### **3.1.4 Gender differences in temperament**

It is important to discuss some notable gender differences. Whether or not the TATC reports the same gender differences that have previously been found by other scales,

will add to our understanding of the scale. A number of studies have found early gender differences across cultures in temperament (e.g., Gartstein, Slobodskaya, Żylicz, Gosztyła, & Nakagawa, 2010). Particularly, studies have found boys to be rated as more active than girls by their parents and teachers (e.g., Eaton, 1983; Martin, Wisenbaker, Baker, & Huttunen, 1997; Maziade, Côté, Boudreault, Thivierge, & Caperaa, 1984; Walker, Berthelsen, & Irving, 2001), as well as to have lower levels of attention (e.g., Eisenberg et al., 1997, 2003; Walker et al., 2001), and to be more 'temperamentally difficult' (e.g., Fabes, Shepard, Guthrie, & Martin, 1997). In a meta-analysis by Else-Quest, Hyde, Goldsmith, and Hulle (2006) of data from children between 3 months and 13 years in 189 studies, the researchers found large differences in attention focusing and inhibitory control favouring girls, and differences in sociability and activity favouring boys.

However, it should be noted that results regarding attention are somewhat inconsistent. For example, McDevitt and Carey (1978) found boys to be more active than girls, congruent with other research, but also to be less distractible than girls, incongruent with the majority of findings. This finding of boys being less distractible than girls was also found in a study by Maziade et al. (1984). In a different study, the same authors within the same year found gender differences to be only slight or to be absent (Maziade, Boudreault, Thivierge, Caperaa, and Côté, 1984). When examining gender differences using the TABC, several analyses on the parent form have revealed no gender differences (Martin, 1988). However, Martin (1988) warns against placing too much weight on these results due to their small sample sizes. An analysis of data from the TABC's teacher form on a sample of 479 children from Georgia in the United States revealed that boys were rated significantly more active, more approaching, more emotionally intense, and more distractible (Martin, 1988). Given that this fits with how adults often think boys should be more active and approaching than girls (Vu, Murrie, Gonzalez, & Jobe, 2008), it may be the case that adult ratings are influenced by stereotypes. Therefore, this is an additional reason for using self-reports, as children are less likely than adults to have developed schemas (see Literature Review).

### **3.1.5 Research aims**

Given the issue with previous measures, the overall aim of this study was to develop a temperament self-assessment tool for children between the ages of 4 and 8 years old. Since traditional Likert-scales are difficult for children younger than 7 (e.g., Chambers & Johnston, 2002; Paulhus, 1991), the temperament assessment tool shall follow a procedure similar to Eder (1990) by having cartoon characters each read various polarising statements (e.g., 'I make myself at home at a stranger's house' and 'I feel uncomfortable at a stranger's house'), and then having children select which best applies to them and to what extent. The aim of the study was not to offer a measurement of temperament that is as encompassing to developmental psychologists as Rothbart et al.'s (1994) CBQ, but to provide a new technique for measuring six particular dimensions of temperament that may be of particular interest to researchers within eyewitness psychology. Specifically, the Temperament Assessment Tool for Children (TATC) has been designed to measure six traits of temperament (activity, adaptability, distractibility, emotionality, persistence, and shyness) via self-reports from young children (4 to 8 years old). The definitions for these traits are displayed in **Table 3.1**.

These six traits have been chosen by the researcher specifically because of their interest to researchers in the field of eyewitness psychology (e.g., Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Chae et al., 2014, 2018; Geddie et al., 2000; Gordon et al., 1993; Greenhoot et al., 1999; Imhoff & Baker-Ward, 1999; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Shapiro et al., 2005). The researcher does not know of any other attempt to create a self-report measure with these six temperament characteristics. Such a tool could potentially help identify child-witnesses who may need additional support during a forensic interview. Specifically, there are three separate objectives to this study that will help determine if the TATC is both reliable and valid: (1) to uncover whether 4- to 8-year-old children have consistent and meaningful understandings regarding their own temperament characteristics, (2) to determine whether the tool discriminates between different children, and (3) to assess if these differences are stable over time.

**Table 3.1.**

*TATC definitions.*

<b>TATC dimension</b>	<b>Definition</b>
Activity	The degree to which one enjoys being physically active. Children who are highly active will have a lot of energy and may fidget a lot during quiet activities.
Adaptability	How easily one adapts to changes in one's environment. Children who are more adaptable are more easily managed as infants and accept new rules or changes to plans more easily than less adaptable children
Distractibility	How easily one's attention is diverted
Emotionality	The intensity with which one's feelings are expressed. Young children who are more emotional will, for example, stomp their foot, scream, or cry with more intensity than children who are less emotional when they get into trouble or do not get their way
Persistence	The extent to which one will continue with an activity, a desire, or a frame of mind, even when challenged.
Shyness	The degree to which one withdraws from social situations.

## **3.2 Methods**

### **3.2.1 Design**

The purpose of this study was to develop a self-report tool for measuring temperament in young children (4 to 8 years old), and to then test the reliability and validity of this scale. Internal consistency within each of the six temperament subscales (activity, adaptability, distractibility, emotionality, persistence, and shyness) was assessed via Cronbach's Alpha, and stability was measured after a period of four months. Interrater reliability was measured by giving parents of the children a version of the tool. Criterion validity for two of the subscales (Distractibility and Persistence) was presented by having the participants complete a distraction task and a persistence task in order to see if performance on these tasks could be predicted by their individual subscale scores.

### 3.2.2 Participants

A total of 202 participants (92 boys and 110 girls) took part in the study, recruited from seven different primary schools in Stirlingshire and Edinburgh. Emails were sent out to head teachers within those two areas after receiving permission from the relevant city councils. If head teachers responded positively to the email, parents within the school with appropriately aged children were sent information sheets and consent forms (**Appendix A**). Once consent forms had been signed and returned to the researcher, the children were asked for verbal consent to take part. The schools that agreed to take part consisted of five community schools and two faith schools around the cities of Stirling and Edinburgh. There was an average of 28.86 participants recruited from each school, ranging from 11 to 56 ( $SD = 16.73$ ). They were between the ages of 4- and 8-years-old ( $M = 5.88$ ,  $SD = 1.27$ ). There were five 4-year-olds, 44 5-year-olds, 53 6-year-olds, 47 7-year-olds, and 53 8-year-olds. The aim was to also have a comparative number of 4-year-old participants, but there was always a small number of 4-year-old children at any school or nursery, and very few consent forms were ever returned. For age-group comparisons, children were split into two groups: younger (4 to 6 years old;  $M = 5.47$ ,  $SD = 0.59$ ) and older children (7 to 8 years old;  $M = 7.52$ ,  $SD = 0.50$ ). There were 102 children in the younger group (42 boys and 60 girls) and 100 children in the older group (49 boys and 51 girls).

Thirty-three of the children (18 boys and 15 girls) from one community primary school in Edinburgh were revisited after a period of 4 months to assess stability. Of those that took part, 1 child was 5 years old, 8 were 6 years old, 5 were 7 years old, and 19 were 8 years old ( $M = 7.27$ ,  $SD = 0.94$ ).

### 3.2.3 Materials

#### 3.2.3.1 *The Temperament Assessment Tool for Children (TATC)*

The TATC was designed to measure six temperament traits (activity, adaptability, distractibility, emotionality, persistence, and shyness) during interviews with young children (4 to 8 years old). This is similar to the six traits of temperament examined by Martin's (1988) TABC, though approach/withdrawal has been relabelled as shyness (which is more understandable to practitioners) and ease-of-management-through-distraction has been replaced with a more general definition of distractibility. This is

because theories suggest that eyewitness performance may be affected by attention being easily diverted (e.g., Ornstein et al., 1997). These six temperament traits are the focus because of the association between at least some of these temperament traits and with memory and suggestibility (e.g., Geddie et al., 2000; Gordon et al., 1993).

The TATC consists of 48 pairs of statements (see **Appendix B**) that were mostly inspired by items of Martin's (1988) TABC, McDevitt and Carey's (1988) BSQ, Rothbart et al.'s (1994) CBQ, and Crozier's (1995) Shyness Questionnaire. Item selection was then determined by inter-judge agreement. Eleven judges were asked to assign the 48 pairs of statements to one of the six categories, using category definitions as a guide. Ten of the judges were members of the Gartstein Infant Temperament Laboratory at Washington State University, consisting of graduate students and psychology professors, and the eleventh judge was an associate professor from a different institution to the researcher and with published research related to measuring child temperament within the last 12 months. Thirty-seven of the 48 items were consistently placed by at least six of the 11 judges, while the other 11 items were reworded or replaced (items 3, 4, 8, 10, 27, 29, 31, 42, 45, 46, and 47) based upon feedback by the judges.<sup>4</sup>

Each of the statements appear in a speech bubble above the head of a cartoon character on printed pieces of A5 coloured card (see **Figure 3.1**). This is similar to the BPI technique in which children are told two polarising statements by two hand puppets and are then asked which one they agree with. There are a total of 12 cartoon characters (collected from a website where people share copyright free images ([www.pixabay.com](http://www.pixabay.com))). Each character represents two statements for four temperament traits (chosen at random); one statement represents the low end of that temperament trait (e.g., 'When told to wear clothes I don't like, I don't argue or yell') and the other statement represents the high end of that temperament trait (e.g., 'When told to wear clothes I don't like, I argue or yell'). This is intended to counterbalance any potential

---

<sup>4</sup> Given that feedback had been provided and there was high agreement concerning 37 of the items already, the reworded or replaced 11 items were not reassessed by the judges.

favouritism children might have for any characters. For half the statements, the statement representing the low end of a temperament trait is the first card read, and, for the other half, the statement representing the high end of a temperament trait is the first card read. There are eight pairs of statements for each of the six temperament traits. The cards always appear in the same order (which follows the same temperament trait sequence as the TABC).

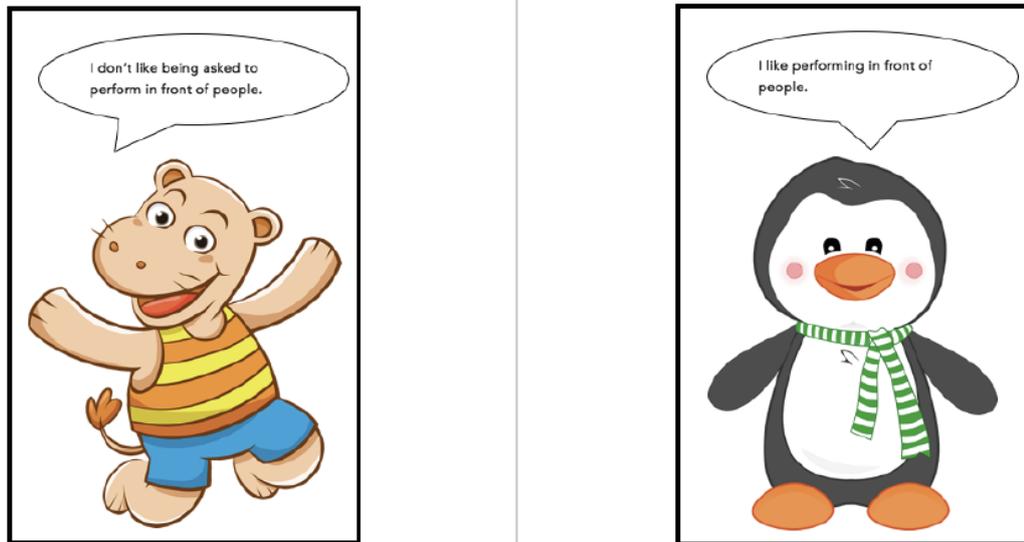


Figure 3.1. Example of TATC items

The cartoon characters were chosen because of their unfamiliarity to the children, and because they were regarded as gender-neutral by the researcher. Past research has indicated that young boys might otherwise have been likely to select the more masculine characters, and young girls might have been likely to select the more feminine characters (Eder, 1990). They were also chosen because all characters are smiling and therefore appear to be positive about the statements they are saying. It was hoped this would lessen any effects of social desirability during use of the TATC.

### 3.2.3.2 Scoring of the TATC

Children were asked whether or not they agree with one statement more than the other, as well as to what extent they agree with their chosen statement (generating a temperament dimension score from 1-5, with 1 representing the low end of the dimension and 5 representing the high end of the dimension). For example, if the child did not side with one statement more than the other (e.g., 'I am not like either of them') then that would generate a neutral score of 3. If they selected the statement

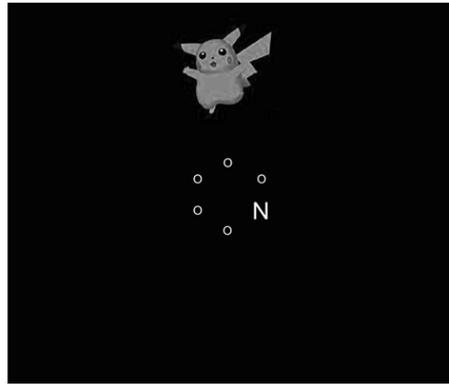
representing the high end of the temperament dimension, they were then asked if they sided with that statement 'a little' or 'a lot'. If the child replied, 'a little', they scored 4. If the child said, 'a lot', they scored 5. Children who selected the statement representing the low end of the temperament dimension were also asked if they sided with the statement 'a little' or 'a lot'. 'A little' generated a score of 2, and 'a lot' scored 1. Unlike with the BPI technique (e.g., Ablow & Measelle, 2010), this removes any subjectivity in deciding the rating and makes sure the coding process is unambiguous. Children's total scores for the subscales are generated by simply adding all their scores (out of a possible 40 for each subscale as they each contain 8 items, measured out of a possible 5).

### *3.2.3.3 TATC Parent form*

For the purpose of a parent form (**Appendix C**), the 48 pairs of statements were translated to 48 items that could be responded to on a 5-point scale based on the degree with which the parent agrees with the statement (strongly disagree, disagree, neutral, agree, or strongly agree). For example, the statements 'I'm quiet with adults I don't know' and 'I talk a lot with adults I don't know' were translated to 'My child is quiet with adults he/she doesn't know' for the parent form.

### *3.2.3.4 Distraction task*

A distraction task was used in the study to provide an external score of distractibility. The task was an E-Prime program developed by Forster and Lavie (2016) to assess the association between ADHD symptoms and task distraction in adults. During the task, participants were required to click certain keys on a laptop in response to letters appearing on their screen (e.g., to click 0 when N appeared). Reaction times were recorded during a distraction condition (irrelevant cartoon characters appearing in the background within colour; see **Figure 3.2**) and during a non-distraction condition (no pictures in the background), and then compared to generate a task distractibility score. This task has not previously been used with children to the researcher's knowledge, but it was the most child-friendly distraction task available to the researcher.



*Figure 3.2.* Distraction condition of distraction task (Forster & Lavie, 2016). In the example, participants are asked to search for a target letter in the letter circle and to press an assigning letter on their keyboard, when an irrelevant distractor appears either above or below the letter circle.

#### *3.2.3.5 Persistence task*

A persistence task was also used in the study to provide an external score of persistence. The task was a popular Czech mechanical puzzle (Hedgehog in the Cage). The object of the puzzle is to release a sphere (the hedgehog) from a cylinder (the cage). The sphere, however, has protruding spikes of various lengths and the cylinder is perforated with holes of different sizes, making it challenging for young children. See **Figure 3.3** for the solution of a similar wooden puzzle. Previous studies have used similar game-based assessments of persistence – though usually video game performance – and have reported high predictive validity (e.g., DiCerbo, 2014; Ventura & Shute, 2013; Ventura, Shute, & Zhao, 2012). As these video games are rather time consuming (up to 4 hours), a quicker task was used due to limited time with each of the participants. Given enough attempts, a child is almost guaranteed to eventually find the correct way of completing the puzzle, and so the researcher felt it was a good measurement of persistence for children. In response to statements such as 'It's too difficult' or 'I give up', each child was asked three times if they were sure they wanted to end the task before the timer was then stopped and their time was recorded. How long it took each child to complete the task or to give up was recorded, as well as how many times they asked to give up (0-3).

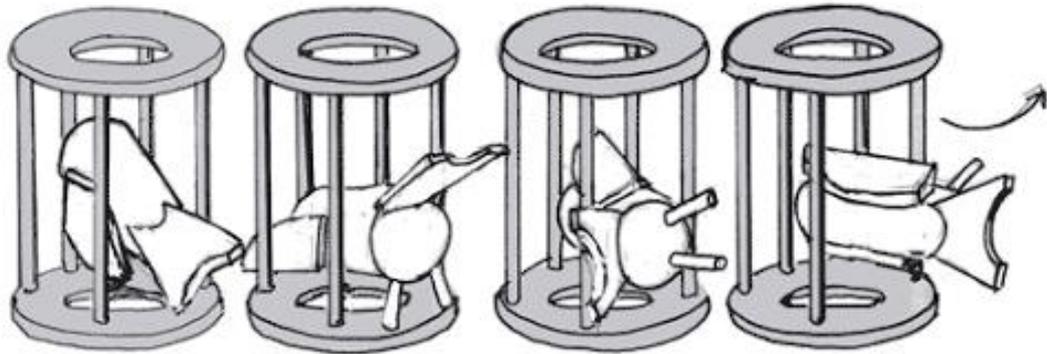


Figure 3.3. Wooden puzzle. Retrieved from <http://www.woodentoys-shop.com/wooden-puzzles-wood-puzzle-brain-teasers.php3>.

### 3.2.4 Procedure

Approval was granted from Edinburgh Napier University's School of Applied Science Research Integrity Committee, and the researcher had PVG disclosure for working with children. The children were interviewed individually, and the interviews occurred within a quiet room on school grounds chosen by the head teacher. In hopes of lessening social desirability, children were told there were no right or wrong answers and that it was important to be honest. The interviewer also established rapport with the children by asking them to talk about themselves and what they are interested in, therefore making the child feel more comfortable (for a review, see Saywitz et al., 2015). The interviewer followed guidelines to use open-ended (e.g., 'Tell me about something you like to do?'), rather than closed-ended questions (e.g., 'Do you like to draw?'). Research shows that children often respond to closed-ended questions with one-word responses, and this is unlikely to effectively build rapport (e.g., Roberts, Lamb, & Sternberg, 2004). The interviewer followed the National Institute of Child Health and Human Development's structured interview protocol when building rapport as this has been shown to enhance children's responses (Sternberg et al., 2001). The procedure was then explained to the child, who was asked to give verbal consent before continuing. The children were read 48 sets of opposing statements (e.g., 'I like talking to children I don't know' and 'I don't like talking to children I don't know') and were asked to select which applies to them and to what extent (generating a score out of 5). The average time to complete the measurement with each child was

about 15 minutes, typically ranging from about 12 to 25 minutes. Afterwards, the children were asked to take home copies of the TATC designed for parents so that interrater reliability could be examined.

To obtain test-retest reliabilities, 33 of the children (18 boys and 15 girls) repeated the procedure four months later. During this visit, the children were also asked to complete the task distractibility and task persistence tasks.<sup>5</sup>

### **3.3 Results**

The purpose of this study was to develop a self-report tool for measuring temperament in young children (4 to 8 years old) and to test the reliability and validity of this scale. Internal consistency within each of the six temperament subscales was assessed via Cronbach's Alpha, and stability was measured via a follow up after a period of four months. Interrater reliability was measured by giving parents of the children a version of the tool, and then by calculating the Pearson correlations to generate an estimate of reliability between the two raters for each subscale. Criterion validity for one of the subscales (persistence) was analysed by having the participants complete a persistence task in order to see if performance on this task could be predicted by individual subscale scores. Results from the distractibility task could not be used for analysis due to all children failing to pass the practice round.

#### **3.3.1 Reliability**

Two indices of reliability were calculated – internal consistency and test-retest reliability. **Table 3.2** shows the reliabilities for the six temperament dimensions for the 202 participants, the 62 parent forms, and the 33 test-retest participants. Category homogeneities were lowest for adaptability and persistence for both self-reports and parent forms. The remaining coefficients were all above .70. Test-retest reliability ( $n = 33$ ) ranged from .54 to .88. Only adaptability did not report stability above .70 after a period of 4-months. All correlations between original test scores and retest scores were statistically significant ( $p \leq .001$ ).

---

<sup>5</sup> Ideally, the tasks would have been piloted with more children during the initial visit, but the researcher had agreed upon with teachers a maximum time to spend with children and the addition of these tasks would have exceeded that time.

**Table 3.2.***Reliabilities for temperament categories of TATC.*

Category	$\alpha$ (Child's reports; $n = 202$ )	$\alpha$ (Parent's forms; $n = 62$ )	Test-retest correlation ( $n = 33$ )
Activity	.72	.82	.88
Adaptability	.64	.54	.54
Distractibility	.79	.79	.81
Emotionality	.81	.74	.71
Persistence	.49	.60	.73
Shyness	.79	.87	.73
Mean	.71	.73	.73
Median	.76	.77	.73

### 3.3.2 Validity

Sixty-two parents completed and returned the TATC.<sup>6</sup> **Table 3.3** shows the mean item score within each temperament dimension based upon both the self-reports and the parent forms, and also shows the interrater correlation for each. Average interrater consistency (Pearson correlations) ranged from .42 to .78. All correlations between raters were statistically significant ( $p \leq .001$ ). Children typically saw themselves as more adaptable ( $M = 3.63$  vs. 3.15), less distractible ( $M = 2.66$  vs. 3.13), less emotional ( $M = 2.28$  vs. 3.00), and more shy ( $M = 3.00$  vs. 2.78) than their parents did, while activity and persistence were both above .70 in consistency. Additional analysis revealed that parents reported children to be significantly less adaptable,  $t(262) = 6.23$ ;  $p < .001$ ;  $d = .90$ , more distractible,  $t(262) = -4.20$ ;  $p < .001$ ;  $d = .61$ , more emotional,  $t(262) = -7.47$ ;  $p < .001$ ;  $d = 1.08$ , and less persistent,  $t(262) = -3.97$ ;  $p < .001$ ;  $d = .58$ , than children's self-reports. There were no other significant differences ( $ps > .05$ ).

<sup>6</sup> The low number of parent forms may have been because children were asked by the researcher to deliver the parent forms home and may have never done so. In retrospect, the researcher would have preferred to ask schools to mail or email the parent forms to parents.

**Table 3.3.**

Mean item score (1-5) within each temperament dimension for self-reports and parent forms, and the interrater consistency correlation.

Category	Self-report	Parent form	Interrater consistency correlation
Activity	3.06 (SD = 0.67)	2.99 (SD = 0.77)	.78
Adaptability	3.63 (SD = 0.54)	3.15 (SD = 0.51)	.51
Distractibility	2.66 (SD = 0.76)	3.13 (SD = 0.76)	.65
Emotionality	2.28 (SD = 0.64)	3.00 (SD = 0.73)	.42
Persistence	3.60 (SD = 0.58)	3.26 (SD = 0.61)	.70
Shyness	3.00 (SD = 0.90)	2.78 (SD = 1.02)	.68
Mean			.62
Median			.67

Evidence for construct validity was measured via the distraction task and persistence task. The results of the distraction task could not be used as none of the children scored satisfactorily during the practice trials. To the researcher's knowledge, this is the first time this distraction task has been used with young children, and the task appears to have been too difficult for young children to complete. Children who completed the persistence task were likely to be significantly higher in persistence (TATC score), compared to those who gave up on the task, after controlling for age ( $M = 27.80$  vs.  $22.62$ ),  $F(1, 30) = 8.79$ ,  $p = .006$ , partial  $\eta^2 = .23$ . Additionally, there was a moderate, negative correlation between persistence (TATC score) and the number of times children asked to give up on the puzzle, which was statistically significant,  $r(31) = -.52$ ,  $p = .002$ . See **Table 3.4** for a summary of reliability and validity scores for each subscale.

**Table 3.4.**

Summary of reliability and validity scores for each subscale.

Temperament	Self-reports consistency ( $n = 202$ )	Parent forms consistency ( $n = 62$ )	Test-retest correlations ( $n = 33$ )	Interrater consistency correlations ( $n = 62$ )
Activity*	.72	.82	.88	.78
Adaptability	.64	.54	.54	.51
Distractibility	.79	.79	.81	.65
Emotionality	.81	.74	.71	.42
Persistence**	.49	.60	.73	.70
Shyness	.79	.87	.73	.68

Note. \*For each subscale, construct validity was also presented via inter-judge agreement.

\*\* Criterion validity for persistence was also presented via the results of the persistence task.

### 3.3.3 Individual differences

In order to better understand the scale, patterns between gender were investigated to see if results are similar to those of other scales. The reliabilities suggest that children have meaningful distinct understandings of their individual temperament dimensions, and that these understandings are generally stable, at least over a period of four months. The mean scores were near the midpoint for each temperament dimension, and the standard deviations suggest considerable variation (see **Table 3.5**). It would have been possible for children's responses to be both consistent and stable, but for the responses to not have differed from child to child; for example, all the children could have selected all eight statements representing the high end of shyness. This was not the case. However, some differences were significantly related to gender.

#### 3.3.3.1 Temperament and gender

This study found that boys were significantly higher in distractibility using TATC scores than girls,  $t(200) = -7.23$ ,  $p < .001$ ,  $d = 1.02$  (see **Table 3.5**). No other differences were statistically significant based upon self-reports ( $ps > .05$ ). These findings were replicated by using both the data from the test-retest, during which girls ( $M = 21.80$ ,  $SD = 6.41$ ) were significantly less distractible than boys ( $M = 26.78$ ,  $SD = 4.75$ ),  $t(31) = -2.56$ ;  $p = .015$ ;  $d = .90$ , and from parent forms, during which girls ( $M = 23.35$ ,  $SD = 5.18$ )

were significantly less distractible than boys ( $M = 27.07$ ,  $SD = 6.59$ ),  $t(60) = -2.49$ ;  $p = .016$ ;  $d = .64$ . Using the scores from the parent forms, there was also a statistically significant difference in activity between girls ( $M = 22.06$ ,  $SD = 5.47$ ) and boys ( $M = 26.11$ ,  $SD = 6.28$ ),  $t(60) = -2.71$ ;  $p = .009$ ;  $d = .69$ . All other gender differences were non-significant ( $ps > .05$ ). Overall, there is evidence to suggest that young boys are significantly more distractible than young girls.

**Table 3.5.**

*Mean score (between 8 and 40) for each temperament trait for both females and males.*

Category	Females ( $n = 110$ )	Males ( $n = 92$ )	Mean ( $n = 202$ )
Activity	23.85 ( $SD = 5.26$ )	25.26 ( $SD = 5.48$ )	24.49 ( $SD = 5.39$ )
Adaptability	29.33 ( $SD = 4.02$ )	28.75 ( $SD = 4.65$ )	29.06 ( $SD = 4.31$ )
Distractibility	18.79 ( $SD = 4.84$ )	24.34 ( $SD = 6.06$ )	21.32 ( $SD = 6.09$ )
Emotionality	18.89 ( $SD = 4.87$ )	17.52 ( $SD = 5.37$ )	18.27 ( $SD = 5.14$ )
Persistence	28.50 ( $SD = 4.45$ )	29.16 ( $SD = 4.78$ )	28.80 ( $SD = 4.61$ )
Shyness	23.83 ( $SD = 7.50$ )	24.36 ( $SD = 6.82$ )	24.07 ( $SD = 7.18$ )

### 3.3.3.2 Temperament and age

This study found that children in the older group were significantly higher in adaptability using TATC scores than those in the younger group,  $t(200) = -3.29$ ,  $p < .001$ ,  $d = .46$  (see **Table 3.6**). No other differences were statistically significant ( $ps > .05$ ). Reliability and validity scores for each age group are displayed in **Table 3.7**.<sup>7</sup>

<sup>7</sup>Retest scores are not reported due to there only being 9 children in the younger age group who completed the retest.

**Table 3.6.**

*Mean score (between 8 and 40) for each temperament trait for both younger and older children.*

<b>Category</b>	<b>Younger (n = 102)</b>	<b>Older (n = 100)</b>
Activity	24.57 (SD = 5.83)	24.41 (SD = 4.94)
Adaptability	28.16 (SD = 4.14)	30.11 (SD = 4.30)
Distractibility	21.83 (SD = 6.09)	20.73 (SD = 6.04)
Emotionality	18.46 (SD = 5.00)	18.05 (SD = 5.30)
Persistence	28.44 (SD = 4.32)	29.11 SD = (4.78)
Shyness	24.31 (SD = 8.01)	23.90 (SD = 6.40)

**Table 3.7.**

*Reliability and validity scores for younger and older children.*

<b>Category</b>	<b><math>\alpha</math> (Child's reports)</b>	
	<b>Younger (n = 102)</b>	<b>Older (n = 100)</b>
Activity	.74	.68
Adaptability	.59	.67
Distractibility	.76	.81
Emotionality	.78	.83
Persistence	.40	.57
Shyness	.85	.71

### **3.4 Discussion**

The aim of this study was to develop a self-report tool for measuring temperament in young children (4 to 8 years old), and to then test the reliability and validity of this scale.

#### **3.4.1 Reliability**

The TATC has fulfilled many of the requirements of any multifactorial psychological test instrument. The internal consistency of .71 is of a high order overall (Hinton, McMurray, & Brownlow, 2004), and the internal consistencies of the various subscales range from .49 to .81. According to cut-offs by Hinton et al. (2004), persistence reported unsatisfactory internal consistency (.49) and adaptability was moderate (.64). Interestingly, when using the TABC, Martin's original sample of parent ratings also reported persistence (.57) and adaptability (.60) to be the least internally consistent of the six subscales (Martin, 1988). Every other subscale in the current study showed high reliability (Hinton et al., 2004). The internal consistency of self-reports for activity could be increased if item 13 ('I usually run everywhere, rather than walk' vs. 'I usually walk everywhere, rather than run') and item 39 ('When I sit, I usually sit still, rather than swing my legs or hands' vs. 'When I sit, I usually swing my legs or hands, rather than sit still') were deleted. These two items were inspired from Martin's (1988) TABC, while the other activity items of the TATC were inspired from various other scales. This may explain why these two items were not as internally consistent as the others. Generally, the activity items from Martin's TABC focuses on hyperactivity. The researcher, however, decided to focus on a more general conception of activity, believing that Martin's conception of activity somewhat overlapped with distractibility, evident from the strong, positive correlations between activity and distractibility in Martin's (1988) results. Within the TATC, there was no significant correlation between activity and distractibility based upon self-reports (see **Appendix D**). Future research should replace these two items (13 and 39). The internal consistency of the parent form ranges from .54 to .87 for the various subscales, and averages at .73 overall. Similar to the self-reports, adaptability (.54) and persistence (.60) were found to have moderate internal consistencies, and the other subscales all reported high reliability (Hinton et al., 2004).

Both the internal consistencies of the self-reports and of the parent forms are generally comparable to that of similar scales. They compare well to the BSQ's range of .47 to .80, and average internal consistency of .65 (Baydar, 1995; McDevitt & Carey, 1978, 1996). The TABC's internal consistencies of the same subscales as the TATC typically ranges between .65 to .86, and averages at .73 for parent forms (Martin, 1988). The TABC, however, has been tested on far more participants. Buss and Plomin's EASI-III (1975) has typically been found to have an average internal consistency of between .62 and .78 (Boer & Westenberg, 1994; Buss & Plomin, 1984; Gasman et al., 2002; Mathiesen & Tambs, 1999; Rowe & Plomin, 1977) with tests on a similar sample size to that of the present study. The CBQ's parent forms reported an average internal consistency of .77 in the original paper (Ahadi, Rothbart, & Ye, 1993), with the internal consistencies of the subscales ranging from .67 to .94. The self-reported version of the CBQ (i.e., Simond's (2006) TMCQ) reported internal consistency scores ranging from .62 to .83, averaging at .74. Overall, the internal consistencies of the TATC's subscales are generally comparable to that of similar scales.

On the whole, the internal consistencies provide statistical evidence that the items within most of the temperament categories are measuring the same dimension. The most notable exception is persistence, which reported .49 using self-reports and .60 using parent forms. The most likely explanation for this is how conceptually broad the subscale is. The items within persistence examined to what extent children continue with activities, desires, and frames of mind in various contexts, including when doing difficult tasks and during conversations. In order for the internal consistency to be higher, a narrower definition of persistence would likely have to be used. The consequence of this, however, could be that content validity is reduced. Adaptability may also have been conceptually broad. For example, there is some disagreement over what individual characteristics influence an individual's adaptability (Le Pine et al, Colquitt, & Erez, 2000). Adaptability can be conceptualised in terms of learning a task that is complex, novel, or ill-defined (Mumford, Baughman, Threlfall, Uhlman, & Costanza, 1993), which has been linked to openness in the five-factor model of personality (Le Pine et al., 2000), or as an element of resilience (Folke et al., 2010).

Similar to persistence, therefore, the conceptually broad definition of adaptability may help explain the moderate reliabilities. This is supported by the fact that adaptability significantly correlated with activity, emotionality, and shyness, while persistence significantly correlated with activity. Removing these two dimensions, there were no significant correlations between the remaining four subscales (see **Appendix D**).

The test-retest reliability of .73 is satisfactory, ranging from .54 to .88 for the various subscales, after a period of four months. The stability of adaptability (.54) was questionable, but every other subscale reported satisfactory stability (Hinton et al., 2004). This compares to an average test-retest reliability of .89 for the BSQ (McDevitt & Carey) after a period of one month, ranging from .67 to .94. Test-retest reliability of Martin's TABC (1988) has been studied extensively, including on samples that were tested as much as two years apart. The TABC's teacher form has reported a test-retest reliability of between .70 and .85 after a period of six months, and a test-retest reliability of between .40 and .65 after a period of 12 months (Martin, 1988). The parent form has reported a test-retest reliability of between .43 and .70 after a period 12 months in respect to ratings by mothers, and a test-retest reliability of between .37 and .62 in respect to ratings by fathers (Martin, 1988). After a period of 24 months, the parent form has reported a test-retest reliability of between .48 and .75 in regard to ratings by mothers, and of between .31 and .65 in regard to ratings by fathers (Martin, 1988). The CBQ's parent forms have test-retest reliability scores ranging from .49 to .79, averaging at .64 after a period of 24 months (Rothbart et al., 2001). There is currently no test-retest reliability data available for the TMCQ as far as the researcher is aware. More data are required from both children and adult raters before a conclusion can be drawn over how stable the responses of the TATC might be over a longer duration of time, but the current results seem promising in regard to the stability of children's self-reported temperament.

### **3.4.2 Validity**

High levels of consistency between different raters are not expected when it comes to measuring elements of one's temperament or personality as people do not always view behaviour the same way, and raters view people within different contexts (Martin, 1988). Still, at least a degree of consistency is expected so there is an

indication that both measures are measuring the same construct. Pearson correlations were used to examine interrater consistency so that the results could be more directly compared to results of the TABC (e.g., Pfeffer & Martin, 1984). Average interrater consistency was moderate (.62), with consistency between the various subscales ranging from .42 to .78. As expected from previous research (e.g., Eisenberg et al., 1998; Measelle et al., 1998; Rudasill et al., 2014), the interrater consistency of the TATC between children and parents was generally higher than that of the TABC between teachers and parents, which typically ranges between 0 and .50. Relationships between clinician forms, parent forms, and teacher forms using the TABC vary from .05 to .33. Teachers and teacher aides range from .06 to .42 in consistency (averaging at .37) when both using the teacher forms of the TABC, and consistency between fathers and mothers centres around .50 when both using the parent forms of the TABC (Martin, 1988). Fathers and mothers range from .28 to .71 in consistency using the CBQ's parent forms, averaging at .51 (Rothbart et al., 2001). As far as the researcher is aware, there is currently no interrater consistency data available for the BSQ or TMCQ. Overall, interrater consistency scores of the TATC compare very well to other measures, with the lowest consistency score (emotionality; .42) being higher than the highest consistency score of some other measures. This is perhaps due to the parent form being completed by the parent who spends the most time with the child, and thus their opinions over the child's temperament being more in sync with the child's own opinions than, for example, the opinions of a parent and a teacher, who interact with the child in different contexts. This is just one possible explanation, as this information was not gathered. Future research could investigate this by asking the rater questions about how much time they spend with the child. Another explanation is that the items on the TATC's parent form were identical to the ones on the self-report form, just altered so they could be completed by a parent. This is in contrast to the TABC's various forms, for example, in which the list of items is completely unique for each form.

The lowest interrater consistency figures stemmed mostly from differences of opinion over adaptability, distractibility, emotionality, and shyness. The inconsistency over shyness ratings between self-reports and parent ratings is consistent with the findings

of Spooner et al. (2005). The interrater consistency between children and parents was considerably higher in the current study however, perhaps due to variations of the same items being used for both children and parents. In the study by Spooner et al. (2005), parents used Rothbart et al.'s (1994) CBQ, while children used Crozier's Shyness Questionnaire (1995). The reasons for lower interrater consistency of shyness than some other temperament dimensions is likely due to the fact that some children feel shy without displaying behaviours that are typical to shyness. For example, an 8-year-old girl in the present study spoke to the researcher for far longer than any of the other children during their respective research sessions (for almost 45 minutes), and was rated extremely low in regard to shyness by her mother. During self-reporting, however, the girl reported almost the maximum level of shyness. When asked why she responded this way, she explained that she quite often volunteers to speak in front of the class or to perform in front of an audience, but that she feels anxious during these instances. Overall, it seems that some items enquire about feelings that may only be clear to the child, and not to an external observer. There is also a possibility that children generally rated themselves as more adaptable, less distractible, and less emotional than their parents due to social desirability. This seems unlikely to be the reason for the difference regarding shyness as research suggests that children do not want to be seen as shy (e.g., Hwang, 2002).

Additional construct validity was presented in the form of the persistence task. Even when controlling for age, more persistent children were less likely to give up on the persistence task. Additionally, less persistent children asked to give up more frequently than children rated as being more persistent. This adds validity to the notion that the persistence scale of the TATC is measuring persistence. There was an attempt to present additional validity in the form of a distraction task (Forster & Lavie, 2016), but the children performed too poorly on the task to use the results. Content validity was also presented for each temperament subscale in the form of inter-judge agreement. Overall, there is a fair amount of evidence presented to support the validity of the TATC's subscales. The most notable exception would be emotionality as there was unsatisfactory consistency (.42) between scores by children and their parents. The most likely explanation is due to social desirability on the side of the children, as in

these cases it was quite clear which option would be considered more socially acceptable (e.g., 'I just whine quietly when I'm angry') and which might be considered as acting spoiled or emotionally immature (e.g., 'I cry, yell, or stomp my foot when I'm angry'). In contrast to the results of this study, using the TABC, Martin (1988) found that interrater consistency between parents and between teachers was higher for emotionality than for any of the other subscales. More research will have to be completed to assess the validity of children's self-reports of emotionality using the TATC. It may be that the items have to be rephrased so socially desirable options are not so clear.

### **3.4.3 Gender differences in temperament**

In the current study, boys were found to be more distractible than girls according to both self-reports and parent forms. This is generally in accordance with previous studies (e.g., Eisenberg et al., 1997, 2003; Martin, 1988; Walker et al., 2001), including a large meta-analysis (Else-Quest et al., 2006). However, some studies have revealed boys to be less distractible than girls (e.g., Eaton, 1983; McDevitt & Carey, 1978). The inconsistency likely stems from the fact that studies have used different measurements. Using the parent forms, boys were also rated as more active than girls. This is consistent with the literature (Else-Quest et al., 2006; Gartstein et al., 2010; Martin, 1988; Martin et al., 1997; McDevitt & Carey, 1978; Walker et al., 2001). Interestingly, there was no significant difference between self-reported activity levels of boys and girls in this study.

It is currently unclear why girls were rated as less active than boys on average by their parents, but no gender differences emerged from the children's self-reports. It may be because parents sometimes have different expectations and perceptions of their children's activity levels based upon their gender. For example, one study found that protective parents were a barrier to girls engaging in certain sports due to fears that they may get hurt and because of some sports being perceived as 'guy sports' by parents (Vu et al., 2008). Furthermore, active boys in this study were characterised positively, whereas active girls were characterised by terms such as 'tomboys', 'too aggressive', and 'too active'. On the other hand, studies monitoring the average number of steps taken by children on a daily basis have reported boys to be more

active than girls (e.g., Telford, Telford, Olive, Cochrane, & Davey, 2015), suggesting that perhaps reports by parents of boys being generally more active than girls is not without merit. Since parents likely compare their own children to other children while assessing activity levels that may also explain why their reports resulted in gender differences, while self-reports did not, since the statements did not have children compare themselves to others, but only think about their own behaviour. Even if children did compare themselves to others without being prompted to, they likely would have compared themselves to their friends and those they spend most of their time with. Considering that the vast majority of peer activity at this age is with same-sex peers (Kail, 2012, p. 424), this may also help to explain the lack of gender differences in self-reports, since girls were less likely to compare themselves to boys, but rather to other girls, and vice versa.

#### **3.4.4 Strengths and weaknesses**

As only a small number of 4-year-old participants were recruited, it is not clear if the tool can be used reliably with this age group. Additionally, 58% of the test-retest sample was 8 years old, and so there is limited evidence presented for the stability of self-reports for younger children. There is currently not enough data to calculate reliability for each age group, but this is something that future research should aim to do. Overall, however, based upon the responses of participants in the study, the findings offer good evidence for the reliability and validity of the TATC.

The biggest strength of the TATC is that it is the only tool available for measuring these six temperament dimensions via children's self-reports. This will provide researchers interested in studying these six areas of temperament with richer insight than is currently available. One scale also brings consistency in how traits are measured, as opposed to using multiple scales to measure these six traits, which may bring potential difficulties due to methodology variation. Researchers are encouraged to use both the self-reports and the parent forms. The TATC is a less subjective process than the Berkeley Puppet Interview and has reported higher reliabilities (e.g., Measelle et al., 1998). The TATC also contains all the benefits of the BPI. First, children are asked items that cover several contexts, rather than be asked for general descriptions or evaluations. Second, children are not presented with yes or no options. Third, children

are not merely presented with Likert-scale items, which young children tend to respond to extremely and with little differentiation (Chambers & Johnston, 2002). Additionally, steps have been taken to reduce the effects of social desirability by having all the characters smiling and seeming proud of their statements. This way, it is less likely to seem to the child that one statement could be considered positive and the other negative. The cartoon characters are also unfamiliar to the child, so they are not tempted to choose a character due to favouritism.

Even if the child does pick a favourite character or animal, counterbalancing has ensured that if a character represents the positive side of an item for any subscale then it will represent the negative side of another item for the same subscale. Even with all this considered, it is crucial to build rapport with the child before conducting the TATC and to explain to the child that there are no right or wrong responses. The child must feel comfortable to answer honestly. This is necessary for researchers getting accurate responses. Future research could investigate which methods of rapport-building are most effective within this context. This study followed the National Institute of Child Health and Human Development's structured interview protocol when building rapport, as this has been shown to enhance the quantity of details given by children in response to free-recall prompts (Sternberg et al., 2001). Potential interviewers should also be aware of the fact that several factors can impact social desirability during interviews with children, including the age, emotional tone, and perceived status of the interviewer (Ceci et al., 1987; Goodman et al., 1991; Tobey & Goodman, 1992).

There is scope for improvement of reliability scores for adaptability and persistence. Future research should make revisions to these items. The problem with persistence is that the researcher has chosen a conceptually broad definition. The items include both matters relating to one's persistence during challenging tasks, and the extent to which one continues with a thought or frame of mind when challenged during conversations. Future research could choose a narrower definition of persistence or divide the subscale into two separate subscales. More data are required to investigate the

reliability of the TATC's scores and to assess with whom it is most reliable. For example, there is not currently enough data to suggest that the TATC can be used reliably with 4-year-olds. It is hoped that the TATC will provide researchers and clinicians with a method for measuring temperament in young children which does not rely solely on caregiver ratings. The TATC has fulfilled many of the requirements of any multifactorial psychological test instrument. The researcher has suggested possible modifications to the TATC for future research. Such research will help lay the foundations for improvements to be made, thereby allowing future researchers to gather as reliable and valid responses from children as possible in regard to their temperamental proclivity.

### **3.5 Summary**

This chapter describes the development of a temperament assessment tool for 4- to 8-year-old children. While this technique could likely be used successfully with older children also, children aged 7 and older are able to reliably report their behavioural characteristics via traditional Likert-scale response items (e.g., Danielson & Phelps, 2003; Ding et al., 2014; Scanlon & Ollendick, 1986; Spooner, 2005), and there are temperament self-reports available for children over the age of 9 (Ellis & Rothbart, 2002; Luby et al., 1999). In this study, children were presented with pairs of statements about temperament dimensions and were then asked to indicate which one of the statements best described themselves. The process can be completed in about 15 minutes and scored in about 10 minutes. Alpha reliability was .71, indicating that responses generally were internally consistent and psychologically meaningful. Additionally, children's responses varied considerably, demonstrating individual differences. The test-retest reliability was .73, indicating that children's responses were stable. There is also preliminary evidence presented of content and criterion validity. The instrument should provide temperament data for research and clinical purposes that is derived from the responses of children themselves, rather than relying on caregiver ratings. There are also recommendations given for future researchers so that small modifications can be made to the temperament assessment tool for improved results. Now that a reliable and valid tool has been devised, it is possible to move on to the next aim of this thesis: to investigate the influence of temperament on eyewitness performance.

## **Chapter 4: Study two – How temperament and interviewing technique affect children as eyewitnesses**

---

### **4.1 Introduction**

In the USA, child protective services annually investigate over 2 million cases of child abuse or neglect (Child Welfare Information Gateway, 2019). A smaller but substantial number are called to criminal courtrooms to testify as witnesses. It is not possible to get an accurate approximation of the current number, but Ceci and de Bruyn (1993) estimated it to be at least 100,000 based upon the figures they had available at the time. The reliability of child eyewitnesses and understanding the factors that may affect their performance are especially important given the large number of children who become involved in the legal system every year. As mentioned in the Introduction, in Scotland, children were asked to give testimony in a criminal court 4,297 times during 2017 (Crown Office and Procurator Fiscal Service, 2018). In England and Wales, the BBC uncovered in 2009 that 47,817 child witnesses testified in court the previous year (Crawford, 2009).

#### **4.1.1 Question types**

In cases involving child witnesses, the child's testimony is often the only available evidence (Brewer et al., 1997; Lamb & Brown, 2006). In these instances, witness testimony is especially crucial to understanding what happened. As mentioned in the Literature Review, investigators are typically encouraged to apply a 'funnel approach', which involves exhausting free recall invitations and other open-ended prompts before reverting to more focused questions (Lamb et al., 2009). During this stage of the interview, even young children can be just as accurate as adults, but they often miss many details (e.g., Hershkowitz et al., 2012; Lamb et al., 2007; Walker, 2013). Police will then use follow-up questions to elicit more details. Open-ended follow-up questions (e.g., 'Tell me more about what happened') generate more accurate and coherent responses than any other type of follow-up questions (e.g., Feltis et al., 2010; Hershkowitz et al., 2012; Lamb et al., 2008; Lyon, 2014; Orbach, Hershkowitz, Lamb, Sternberg, & Esplin, 2000; Snow et al., 2009).

In order to elicit the largest quantity of information possible, police will often use a mixture of open-ended and closed-ended questions (Shapiro, 2006). According to Swerdlow-Freed (2018), closed questions can be considered as option-posing, forced-choice, multiple-choice, or yes/no. Problematically, while they may increase the amount of information given, option-posing often leads children to answer questions they do not know the answer to and so overall accuracy is typically decreased (e.g., Peterson et al., 1999; Walker, 2013; Waterman et al., 2000, 2001, 2004). In fact, children between 5 and 7 years are likely to answer even nonsensical questions if they are presented in a way that poses options (e.g., 'Is red heavier than yellow?'; Hughes & Grieve, 1980; Pratt, 1990; Waterman et al., 2000). This is perhaps because young children feel compelled to provide an answer in order to please the interviewer (Ceci & Bruck, 1995). Even when at least half of the closed-ended questions are correct-leading (i.e., the correct answer is suggested via the question), accuracy is still lower overall than when compared to witnesses who only receive open-ended questions, even though the open-ended questions will typically produce answers that are less complete (e.g., Dodd & Bradshaw, 1980; Loftus et al., 1978). It is therefore important to consider the quality of information given (i.e., accurate or inaccurate), not just overall quantity.

Due to the fact that young children may produce little information in response to open-ended questions, research has found that investigators struggle to maintain best practice, and often overuse improper, including suggestive, questioning (e.g., Luther et al., 2015; Roberts & Cameron, 2015). The problem with police asking suggestive questions is that they have no way of knowing if the question is correct-leading or misleading (i.e., those that lead the witness in the wrong direction). Misleading questions may lead to suggestibility when witnesses wrongly comply with the suggestions of the interviewer and then incorporate that misinformation into their subsequent accounts of the crime. As mentioned in the Introduction, suggestibility is defined by Ceci and Bruck (1995, p. 195) as 'the extent to which individuals come to accept and subsequently incorporate post event information into their memory recollections.' According to Bruck and Melnyk's (2004) review, inaccuracy due to misleading questions can be considered a form of suggestibility. Younger children are

typically more susceptible to inaccuracy from misleading questions than older children and adults (e.g., Cassel et al., 1996; Poole & White, 1993).

### **4.1.2 Age-related differences in suggestibility**

Age is the most reliable predictor of children's memory recall and witness abilities (Ceci & Bruck, 1993, 1995). As outlined in the Literature Review, young children (especially below the age of 6 years) typically provide fewer details than older children and adults. During free recall, there are often no age-related differences in accuracy (e.g., Sutherland & Hayne, 2001), though older children provide more details (e.g., Hershkowitz et al., 2012). In some cases, older children still perform significantly more accurately than younger children even during free recall (e.g., Eisen, Qin, Goodman, & Davis, 2002). As explained in the Literature Review, younger children are typically more vulnerable than older children to external suggestive influences (e.g., Ceci & Bruck, 1993; Ceci & Huffman, 1997; Otgaar, Candel, Smeets, & Merckelbach, 2010; Otgaar, Howe, Brackmann, & Smeets, 2016; Paz-Alonso & Goodman, 2016; Sutherland & Hayne, 2001). This may be especially true when it comes to interrogative suggestibility (e.g., Gudjonsson, Vagni, Maiorano, & Pajardi, 2016). Nevertheless, age-related differences in suggestibility are not always found (e.g., AfHjelmsater, Granhag, Stronmwall, & Memon, 2008, as cited in Paz-Alonso & Goodman, 2016; Lee, 2004; Warren et al., 1991), suggesting there are other individual differences at play.

### **4.1.3 Temperament**

As stated, age does not always account for differences in suggestibility. There are cognitive and personality factors that can influence how suggestible children are. Currently, as described in the Literature Review, there is strong evidence for some of these factors, such as intelligence, memory ability, language ability, attachment style, and creativity (for reviews, see Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020). Other factors, particularly temperament, however, still require further investigation. It is theorised that certain temperament characteristics or behavioural styles may affect one's understanding, interpretation, and processing of an event, as well as one's ability to resist leading questions when providing eyewitness testimony (Shapiro et al., 2005). Ornstein et al. (1997) proposed that particular elements of temperament affect the perception of eyewitnesses as the events take place (activity level, emotionality, and persistence), while other elements (adaptability, shyness, and distractibility) impact on

their performance during forensic interviews. The TABC (Martin, 1988) measures these six temperament characteristics in 4- to 8-year-old children via caregiver ratings, and is the most popular method of assessing childhood temperament in eyewitness psychology (e.g., Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Chae et al., 2014, 2018; Geddie et al., 2000; Gordon et al., 1993; Greenhoot et al., 1999; Imhoff & Baker-Ward, 1999; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Shapiro et al., 2005).

These six temperament traits have been theorised by researchers to be associated with eyewitness performance and suggestibility in the following ways (see Literature Review for a more detailed account):

- ◇ *Activity*. More energetic children are more likely to divert their attention to multiple stimuli and so encode fewer details during the witnessed event (e.g., Chen, 2002; Melnyk, 2002; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Pezdek & Roe, 1995; Roberts & Powell, 2001; Shapiro, et al., 2005; Quas et al., 1999; Warren et al., 1991).
- ◇ *Adaptability*. Less adaptable children may feel more uncomfortable during a forensic interview than more adaptable children (e.g., Geddie et al., 2000; Greenhoot et al., 1999; Shapiro et al., 2005).
- ◇ *Distractibility*. More distractible children divert their attention both during the witnessed event and interview (e.g., Benedan et al., 2020; Melnyk, 2002; Purdy, 2001).
- ◇ *Emotionality*. More emotional children are typically more emotional when witnessing a crime and this impedes the process of encoding (Chae & Ceci, 2005; Chen, 2002; Geddie et al., 2000; Greenhoot et al., 1999; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Scullin, 1997; Shapiro, et al., 2005).
- ◇ *Persistence*. Less persistent children are less resistant to misleading questions due to being less likely to continue with a frame of mind when challenged (e.g., Chen, 2002; Greenhoot et al., 1999).
- ◇ *Shyness*. Shyer children are more sensitive to the social aspects of an interview and may be more easily swayed by suggestions from the interviewer (e.g., Cotterill, 2017; Endres et al., 1999; Gilstrap & Papierno, 2004).

Overall, there is evidence to suggest that temperament characteristics that limit attention can decrease eyewitness performance, and that witnesses that are more sensitive to social pressure, experience greater difficulty during encoding and/or retrieval of information (Shapiro, 2006). Despite this, a relationship between temperament and suggestibility is not always found (e.g., Blackford & Shapiro, 1999, as cited in Bruck & Melnyk, 2004; Burgwyn-Bailes et al., 2001; Gordon & Ornstein et al. 2001; Imhoff & Baker-Ward, 1999; Young et al., 2003). In fact, Klemfuss and Olaguez's (2020) review found there was no further evidence for an association between temperament and suggestibility in children since Bruck and Melnyk's (2004) review. The decline in interest is likely due to the mixed results of previous research and the increased difficulty in working with child participants (e.g., extra clearance being required before conducting research with children makes it a more time-consuming process). It remains to be understood if the difference in results between the previous studies is due to the different ages of participants, the temperament measures, and/or the different stimuli used across studies (Shapiro, 2006).

As explained in the previous chapter, one possible problem is that studies have often determined the temperament of child participants via ratings by teachers, even though research has consistently shown teachers have low consistency with other raters when it comes to rating the personality traits of their students (e.g., Eisenberg et al., 1998; Measelle et al., 1998; Rudasill et al., 2014; Spooner et al., 2005). To the researcher's knowledge, no published studies examining temperament and suggestibility have used self-reported measures, even though young children can provide reliable and valid self-reports so long as age-appropriate techniques are used (e.g., Hwang, 2002; Measelle et al., 2005; Roth et al., 2004). The TATC is a new, self-report measure designed to measure these six temperament traits in young children (see Chapter Three).

#### **4.1.4 Research aims**

The aim of the study was to investigate the impacts of question types (i.e., open-ended, closed-ended, and misleading) and temperament traits on the accuracy of young child eyewitnesses (4 to 8 years old). Given that the association between

temperament and child eyewitness performance has produced mixed results when temperament is measured with caregiver ratings (for reviews, see Bruck & Melnyk, 2004; Klemfuss & Olaguez, 2020), the research aimed to see if there would be an effect when the data is derived from the views and experiences of the children themselves.

The *first hypothesis* is that, overall, accuracy will be significantly higher in response to open-ended questions than closed-ended or misleading questions, especially in the case of younger children compared to older children (e.g., Ceci & Bruck, 1995; Feltis et al., 2010; Hershkowitz et al., 2012; Lamb et al., 2018; Lyon, 2014; Orbach et al., 2000; Snow et al., 2009). The *second hypothesis* is that during closed-ended questions and misleading questions, certain temperament traits will significantly reduce accuracy further, including activity, low adaptability, distractibility, emotionality, low persistence, and shyness.

## 4.2 Methods

### 4.2.1 Design

The dependent variables of this study were: (1) the quantity of correct descriptors given by interviewees, (2) the quantity of report errors given by interviewees, and (3) the overall accuracy of the interviewee's report. The study tested for effects from three independent variables. First, interview condition, which had three levels (open, closed, and misleading). Second, temperament (activity, adaptability, distractibility, emotionality, persistence, and shyness) as measured by the TATC, a self-report technique made specifically for this thesis. Third, age, which had two levels (older and younger children).

### 4.2.2 Participants

A total of 202 participants (92 boys and 110 girls) took part in the study, recruited from seven different primary schools. These were the same participants as the previous study. They were between the ages of 4 and 8 years old ( $M = 5.88$ ,  $SD = 1.27$ ). There were five 4-year-olds, 44 5-year-olds, 53 6-year-olds, 47 7-year-olds, and 53 8-year-olds. For age-group comparisons, children were split into two groups: younger (4 to 6 years old;  $M = 5.47$ ,  $SD = 0.59$ ) and older children (7 to 8 years old;  $M = 7.52$ ,  $SD = 0.50$ ). There were 102 children in the younger group (42 boys and 60 girls) and 100

children in the older group (49 boys and 51 girls). For the conservative estimate of  $\eta^2 = .25$  for the effect size, a power of .80, and a Type I error threshold of  $p = .05$ , the total sample size needed to detect the effect is 158, calculated with the software G\*Power (Faul, Erdfelder, Buchner, & Lang, 2009).

### **4.2.3 Materials**

#### *4.2.3.1 Stimulus*

Footage was used from Shapiro et al.'s (2005) trip to the zoo video as it has been used in a number of previous eyewitness studies (e.g., Chen, 2002; Shapiro, 2006). There is a two-minute sequence at the beginning of the film, shown to the children on a laptop, in which two teenage girls witness a bike theft. During the sequence, a teenage boy asks repeatedly to use a preteen girl's bike. After the girl continuously turns him down, the boy leaves, sneaks back, and then steals the bike.

#### *4.2.3.2 Interview schedule*

Memory over what happened during the video was assessed via one of three interview formats (open-ended (**Appendix E**), closed-ended (**Appendix F**), or misleading (**Appendix G**)). The memory assessment focused on three aspects of the criminal event: (a) characteristics of the stolen bike, (b) actions portrayed by the actors during the video, (c) and the physical characteristics and clothing of the actors within the video. All interview schedules followed closely the introduction of the steps of Scotland's Stepwise Protocol and the NICHD investigative interview protocol (Lamb et al., 2007). All interviews were recorded for coding and transcription. The interview began with the researcher introducing himself to the child, and by making sure the child understood the difference between telling the truth and a lie (e.g., 'If I said my shoes were red, would that be true or not true?'). Rules were also explained to the child, including that they should only answer questions they know the answer to and that they should correct the interviewer if the interviewer said anything incorrect. The researcher then built rapport with the child by asking the child to talk about themselves and about something they like doing. Children were given a practice interview by being asked to tell the researcher about something that occurred recently (e.g., summer holidays, Christmas). Once the practice interview was over, all interview schedules initially asked the children for free recall, inviting them to tell the researcher everything they remembered from the video about the bike theft.

After providing the researcher with information via free recall, the open-ended format contained unbiased follow-up questions (e.g., 'Tell me who the bike belonged to.'). The closed-ended interview format included follow-up questions that provided two options for the child to choose from (e.g., 'Did the bike belong to the girl or the boy?'). In this condition, children were presented with the correct option first during the first question, and then last during the second question, and so on. The misleading interview format consisted of only misleading follow-up questions (e.g., 'The bike belonged to the boy, didn't it?'). Half of these questions asked the child if an incorrect fact was true (e.g., 'Was the bike red?'), and the other half asked the child to confirm an incorrect statement (e.g., 'The handlebars on the bike were curved, weren't they?'). If any of the children said, 'I think', 'I'm not sure', or 'I don't know' at any point then they were reminded to only answer questions for which they were certain about the answer.

#### *4.2.3.3 Temperament*

In order to assess temperament, the researcher read to the children 48 statements and 48 opposing statements from the TATC (see Chapter Three). Each of the statements (e.g., 'I like talking to children I don't know' and 'I don't like talking to children I don't know') appeared in a speech bubble above the head of a cartoon character in a wire bound document. Children were then asked whether or not they agreed with one statement more than the other, as well as to what extent they agreed with their chosen statement (generating a score from 1-5). This self-report measure is designed to measure six temperament traits (activity, adaptability, distractibility, emotionality, persistence, and shyness). It is the only tool available to assess young children's self-reported temperament. The 48-page TATC was read to all of the children in the same order, and this exercise acted as a buffer between children watching the video and the interview examining what children remembered from the video. The average time to complete the measurement with each child was 15 minutes, typically ranging from about 12 to 25 minutes. Psychometric details of the measure can be found in Chapter Three.

#### **4.2.4 Procedure**

Approval was granted from Edinburgh Napier University's School of Applied Science Research Integrity Committee. Each child was introduced to the researcher by their teacher and then brought to a quiet room where the researcher explained the procedure of the study, asked for demographic information, and for consent to continue with the study. If consent was given, the child was asked to watch the bicycle theft video. Afterwards, the TATC was read to the child and the researcher recorded the responses into a log book. During this process, it was made clear to the participants that there were no correct or incorrect responses, and that it was important they answer honestly. The researcher then assessed the memory of the child via one of the three interview formats (children were allocated at random), each of which began with a brief introduction over the nature of eyewitnesses, and asked for permission to either audio or video record the interview (thus allowing the interviewer to later check the accuracy and completeness of what was written during the interview). Afterwards, the children were debriefed (**Appendix H**), but were not told the correct answers and were asked not to talk about the study with classmates. The average time spent with each participant was 25 minutes.

#### **4.2.5 Coding**

The events of the video contained a total of 32 features (see **Table 4.1**) that could be organised into 14 central descriptors (i.e., essential and salient aspects of an event) and 18 peripheral descriptors (i.e., less important aspects of the crime). These descriptors have been chosen because they either establish guilt of the boy or support the credibility of the eyewitness, and thus would be important parts of an eyewitness testimony (Shapiro et al. 2005). Details have been allocated as central or peripheral (**Appendix I**) based upon the coding scheme of previous research (e.g., Chen, 2002; Shapiro, 2006).

**Table 4.1.**

*List of central and peripheral descriptors.*

<b>14 Central Descriptors</b>	<b>18 Peripheral Descriptors</b>
<i>The Crime</i>	<i>The Crime</i>
The bike owner.	The girl sitting on a bench.
The colour of the bike.	The first song sung by the girl.
The model and handlebars of the bike.	The boy touching the bike.
The argument between the girl and boy.	The boy touching the girl.
The girl and boy pulling the bike back and forth.	The girl moving the bike.
The boy sneaking up and stealing the bike.	The second song sung by the girl.
The location of the crime.	The boy slitting his throat with his finger.
	The boy calling the girl a name.
	The girl's reaction to the bike theft.
	The girl's father arriving afterwards.
<i>Person</i>	<i>Person</i>
The boy's name.	The girl's name.
The boy's hair colour	The girl's hair.
The boy's hair length.	The father's hair colour.
The height of the boy and girl.	The girl's clothing.
The age of the boy and girl.	The girl's shoes.
The boy's clothing.	The father's clothing.
The boy's shoes.	The father's shoes.
	The boy's watch.

A score was generated for the number of correct descriptors given by the child during the interview, and a separate score was produced for the number of incorrect descriptors made by the child. These were total scores including both the information provided during recall and follow-up questions. For each descriptor (correct or incorrect), there were three possible points that could be scored by the child. A partially correct or incorrect response (i.e., a description that was incomplete) would score one point, while a complete correct or incorrect response would score two points, and an elaborative correct or incorrect response would score three points. For

example, if the question was in regard to the colour of the bike, a partial response would be 'dark', a complete response would be 'black', and an elaborative response would be 'black with silver writing on it'. The three levels of possible responses to each of the questions were determined before data collection began. In some instances, a response to a question could only have been correct or incorrect (i.e., it could not be partial or elaborate). For instance, the bike could only have belonged to the girl or to the boy. In this case, a child could only score two points for correct information, two points for incorrect information, or zero points for no information. Total correct and incorrect scores were generated by marking the children's interview transcripts against this coding scheme (see **Appendix J**). Six of the transcripts, chosen at random, were coded by a different researcher to check for interrater reliability, producing an average reliability score of 94% across the six transcripts.<sup>8</sup> This was done following Miles and Huberman's (1994) calculation to measure consistency between coders (i.e.,  $\text{Reliability} = \text{Number of agreements} / (\text{Number of agreements} + \text{disagreements})$ ). This method has been used before in studies with child witnesses to code descriptive data (e.g., Magnusson, Ernberg, & Landström, 2017). A final score of overall accuracy from each child was created by adding their correct descriptor score with their incorrect descriptor score to generate a total score, and then by dividing their correct descriptor score against their total score (i.e.,  $\text{Overall accuracy} = \text{Correct} / (\text{Correct} + \text{Errors})$ ). This allowed the researcher to calculate the proportion of accurate information provided by each child.

## **4.3 Results**

### **4.3.1 Interview format and age**

The first hypothesis of the study was that overall accuracy would be higher in response to open-ended questions than closed-ended or misleading questions, especially in the case of younger children compared to older children. The mean results for overall quantity of correct information, overall quantity of errors, and overall accuracy are displayed in **Table 4.2** based upon interview format and age. A 2 x 3 ANOVA was run for each dependent variable, with age (younger vs. older) and interview format (open

---

<sup>8</sup> Only a small number of transcripts were checked for interrater reliability due to the time-consuming nature of a second researcher getting used to the coding scheme. Given the detail of the coding scheme, the researcher was confident that transcripts were marked consistently.

vs. closed vs. misleading) as between-subject factors. All assumptions were met. No outliers were identified using boxplots. The three dependent variables were approximately normally distributed for each interview condition, according to Shapiro-Wilk test of normality ( $ps > .05$ ), and there was homogeneity of variances according to Levene's test ( $ps > .05$ ).

There was a statistically significant difference in correct descriptors between interview conditions,  $F(2, 196) = 118.99, p < .001$ , partial  $\eta^2 = .55$ . A Tukey post hoc test revealed that children who received the misleading format provided significantly fewer correct descriptors than those who received the open-ended ( $p < .001$ ) or closed-ended questions ( $p < .001$ ). There was also a statistically significant difference in correct descriptors between age groups,  $F(1, 196) = 24.76, p < .001$ , partial  $\eta^2 = .11$ . There was no significant interaction between interview conditions and age group on correct descriptors,  $F(2, 196) = .19, p = .83$ .

There was a statistically significant difference in errors between interview conditions,  $F(2, 196) = 65.74, p < .001$ , partial  $\eta^2 = .40$ . A Tukey post hoc test revealed that children who received the open-ended format made significantly fewer errors than those who received the closed-ended ( $p < .001$ ) or misleading questions ( $p < .001$ ), and those who received the closed-ended questions made significantly fewer errors than those who received the misleading questions ( $p < .001$ ). There was also a statistically significant difference in errors between age groups,  $F(1, 196) = 5.30, p = .02$ , partial  $\eta^2 = .03$ . There was no significant interaction between interview conditions and age group on errors,  $F(2, 196) = .76, p = .47$ .

Lastly, there was a statistically significant difference in overall accuracy between interview conditions,  $F(2, 196) = 166.39, p < .001$ , partial  $\eta^2 = .63$ . A Tukey post hoc test revealed that there was a significant difference between open and closed conditions ( $p < .001$ ), open and misleading conditions ( $p < .001$ ), and closed and misleading conditions ( $p < .001$ ). There was also a statistically significant difference in overall accuracy between age groups,  $F(1, 196) = 20.42, p < .001$ , partial  $\eta^2 = .09$ . There was no

significant interaction between interview conditions and age group on errors,  $F(2, 196) = .10, p = .91$ .

**Table 4.2.**

*Mean correct descriptors, errors, and overall accuracy based upon and interview format and age.*

	<b>Interview format</b>	<b>Younger (n = 102)</b>	<b>Older (n = 100)</b>	<b>Total (n = 202)</b>
Correct	Open	48.65	53.88	51.67
	(n = 66)	(SD = 10.83)	(SD = 9.94)	(SD = 10.79)
	Closed	52.00	56.76	54.31
	(n = 70)	(SD = 9.99)	(SD = 7.24)	(SD = 8.47)
	Misleading	31.13	37.38	34.35
	(n = 66)	(SD = 5.25)	(SD = 5.07)	(SD = 6.01)
Errors	Total	44.33	49.57	46.93
		(SD = 12.54)	(SD = 11.61)	(SD = 12.34)
	Open	10.00	8.00	9.03
	(n = 66)	(SD = 5.04)	(SD = 4.34)	(SD = 4.78)
	Closed	16.31	13.32	14.86
	(n = 70)	(SD = 5.88)	(SD = 6.01)	(SD = 6.09)
Accuracy (%)	Misleading	20.69	20.12	20.39
	(n = 66)	(SD = 5.04)	(SD = 6.54)	(SD = 6.29)
	Total	15.58	13.93	14.76
		(SD = 7.11)	(SD = 7.54)	(SD = 7.36)
	Open	83.35	87.29	85.26
	(n = 66)	(SD = 7.33)	(SD = 6.43)	(SD = 7.14)
Accuracy (%)	Closed	76.46	81.29	78.81
	(n = 70)	(SD = 6.47)	(SD = 7.41)	(SD = 7.31)
	Misleading	60.67	65.62	63.22
	(n = 66)	(SD = 7.27)	(SD = 8.09)	(SD = 8.04)
	Total	73.80	77.88	75.82
		(SD = 11.67)	(SD = 11.72)	(SD = 11.84)

### 4.3.2 Temperament

The second hypothesis of the study was that accuracy would be predicted by temperament. Multiple regressions were run to predict correct descriptors, errors, and overall accuracy from the six temperament dimensions. Assumptions were met for running multiple regressions. The bivariate Pearson's correlations between the predictors were examined and there was no evidence of multicollinearity (all  $r$  values < .8). Tolerance values were also in excess of .20 for all predictors. Finally, for each of the dependent variables, the Durbin-Watson test statistic was between 1 and 2 for errors, indicating independence of residuals. Overall, temperament significantly predicted the number of errors,  $F(6, 195) = 9.65$ ;  $p < .001$ ;  $R^2 = .23$ , and the overall accuracy,  $F(6, 195) = 5.93$ ;  $p < .001$ ;  $R^2 = .15$ . Distractibility was a significant, positive predictor for errors ( $B = .58$ ,  $p < .001$ ) and significant, negative predictor for overall accuracy ( $B = -.76$ ,  $p < .001$ ). All other temperament dimensions were non-significant predictors ( $ps > .05$ ).

Pearson correlations were performed for each age group (younger and older) and interview condition (open-ended, closed-ended, misleading) separately (see **Appendix K**), due to there not being enough participants in each condition for separate regressions without losing power (VanVoorhis & Morgan, 2007). In the open-ended condition, there was a weak, positive correlation between distractibility and number of errors,  $r(64) = .26$ ,  $p = .036$ . When only examining results of the younger children in the open-ended condition, there was a weak, negative correlation between adaptability and number of errors,  $r(32) = -.35$ ,  $p = .045$ . When only examining the results of the older children in the open-ended condition, distractibility weakly, positively correlated with errors,  $r(30) = .38$ ;  $p = .033$ ; and weakly, negatively correlated with overall accuracy,  $r = -.41$ ;  $p = .020$ .

In the closed-ended condition, there was a moderate correlation between distractibility and number of errors,  $r(68) = .53$ ;  $p < .001$ , and overall accuracy,  $r = -.46$ ;  $p < .001$ . When only examining results of the younger children in the closed-ended condition, there was a weak, positive correlation between distractibility and number of errors,  $r(34) = .40$ ,  $p = .015$ . When only examining the results of the older children in

the closed-ended condition, distractibility moderately correlated with errors,  $r(32) = .61$ ,  $p < .001$ ; and with overall accuracy,  $r = -.52$ ;  $p = .002$ . Furthermore, there was a weak, negative correlation between activity and overall accuracy,  $r = -.36$ ,  $p = .037$ .

In the misleading condition, correct descriptors positively correlated with adaptability,  $r(64) = .37$ ;  $p = .002$ , and persistence,  $r = .48$ ;  $p < .001$ . Furthermore, errors negatively correlated with adaptability,  $r = -.25$ ;  $p = .043$ , and positively correlated with distractibility,  $r = .56$ ;  $p < .001$ , in the misleading condition. Also, overall accuracy in the misleading condition positively correlated with adaptability,  $r = .45$ ;  $p < .001$ , and negatively correlated with distractibility,  $r = -.55$ ,  $p < .001$ . All other correlations were non-significant ( $ps > .05$ ). When only examining results of the younger children in the misleading condition, there was a moderate, positive correlation between persistence and correct descriptors,  $r(30) = .64$ ,  $p < .001$ . Furthermore, distractibility moderately correlated with errors,  $r = .63$ ,  $p < .001$ ; and with overall accuracy,  $r = -.62$ ;  $p < .001$ . When only examining results of the older children in the misleading condition, distractibility correlated with correct descriptors,  $r(32) = -.36$ ;  $p = .039$ , errors,  $r = .50$ ;  $p = .003$ , and overall accuracy,  $r = -.55$ ;  $p = .001$ .

## 4.4 Discussion

The aim of this study was to better understand the factors potentially responsible for impacting child eyewitness performance and suggestibility, specifically interview format, temperament, and age.

### 4.4.1 Interview format and age

The first hypothesis of the study was that overall accuracy would be higher in response to open-ended questions than closed-ended or misleading questions, especially in the case of younger children compared to older children. The first part of this hypothesis was formed on the basis of past research, finding that children are less accurate when provided with incorrect information, whether it be in the form of an incorrect option from option-posing questions or in the form of a misleading question (e.g., Feltis et al., 2010; Hershkowitz et al., 2012; Lamb et al., 2018; Lyon, 2014; Orbach et al., 2000; Peterson et al., 1999; Snow et al., 2009; Walker, 2013; Waterman et al., 2000, 2001, 2004). Overall, children provided significantly more correct descriptors in response to

open-ended questions than misleading questions in the current study, but not in response to closed-ended questions. This is consistent with research finding that children provide less accurate, but more complete reports in response to closed-ended questioning than open-ended (e.g., Dodd & Bradshaw, 1980; Loftus et al., 1978). There were certain descriptors within the current study—such as the boy's name or the name of the songs that the girl was singing—that were perhaps more likely to be remembered by the child if the interviewer jogged their memory via options in the closed-ended condition.

In the current study, children reported significantly more incorrect information in response to closed-ended and misleading questioning compared to open-ended questions, and significantly more in response to misleading questions compared to closed-ended. Furthermore, overall accuracy was significantly higher in the open-ended condition than in the closed-ended or misleading conditions, and significantly higher in the closed-ended condition than the misleading. The pattern of overall accuracy being worse for closed-ended questions compared to open-ended questions, and then worse still for misleading questions, is consistent with that of previous research (e.g., Memon, Holley, Wark, Bull, & Koehnken, 1996). It may depend on context. For instance, in cases of sexual touching, specific questions are often necessary to elicit disclosures (for a review, see Dupree, 2016). In the present study, however, there were no significant benefits of any type of questions other than open-ended.

The results revealed that older children provided significantly more correct descriptors than younger children. This is consistent with findings revealing that older children generally report more descriptors in total than younger children (e.g., Davies et al., 1989; Karageorge & Zajac, 2011; Kuehn, 1974; Lindsay et al., 1994; Pozzulo & Warren, 2003; Zajac & Karageorge, 2009). Older children also provided significantly more errors than younger children (most notably in the closed-ended condition, though there was no significant interaction effect) and were overall significantly less accurate in their reports. Therefore, responses of older children were both more complete and more

accurate than younger children. Congruent with these findings, past studies have often reported younger children (especially those aged 6 years and younger) to be more suggestible than older children (especially those aged 8 years and over; e.g., Cassel et al., 1996; Ceci & Bruck, 1993; Ceci & Huffman, 1997; Otgaar et al., 2010; Otgaar et al., 2016; Paz-Alonso & Goodman, 2016; Poole & White, 1993; Sutherland & Hayne, 2001; Poole & White, 1993). Even though younger children performed significantly poorer overall than older children, the first hypothesis was not supported, as there was no significant interaction effect, meaning younger children did not perform significantly poorer than older children as a consequence of any particular interview condition.

#### **4.4.2 Temperament**

The second hypothesis of the study was that temperament would predict accuracy. This hypothesis was also supported as accuracy could be predicted in the study to some extent by temperament traits. Children who were more distractible made more errors and were overall less accurate compared to less distractible children. Distractibility also positively correlated with errors in all three conditions and negatively correlated with overall accuracy in the closed-ended and misleading conditions. Furthermore, children within the misleading condition higher in adaptability and persistence gave a higher number of correct details (i.e., they were more resistant to being misled). Children lower in adaptability in the misleading condition also made more errors and were overall less accurate in their reports (i.e., they were more easily misled).

The researcher did believe that shyness would also have been a significant predictor of eyewitness performance during the misleading condition in the study based upon previous research (e.g., Cotterill, 2017; Gilstrap & Papierno, 2004). Since there is a significant, negative correlation between adaptability and shyness (Martin, 1988), this may have been what previous results were tapping into. This is supported by the fact that Johnston et al. (2021) found that children higher in social flexibility, a construct that combines high adaptability and low shyness, correctly disclosed significantly more transgressions during free recall than other children.

These are similar findings to some past research (e.g., Benedan et al., 2020; Blackford, 2000; Burrell et al., 1999; Melnyk, 2002; Purdy, 2001). Greenhoot et al. (1999), for example, found that more persistent children were more resistant to misleading questions. This could be because they are more likely to continue with a frame of mind even when challenged. The current study, however, is the only instance known to the researcher in which these findings have been found with self-reported temperament scores. The fact that by using children's own ratings this study found results similar to studies using parent ratings, supports the validity of children's self-reported temperament scores. Future studies should aim to use children's self-reports where possible, and also have the ratings of parents to verify the findings and to provide comparisons with previous research that has used parent ratings.

Overall, the results of the study suggest that witnesses who are less adaptable to new environments experience greater difficulty during retrieval of information when given misleading questions. The fact that adaptability had an impact in the misleading condition but not others, nor overall, suggests that adaptability does not relate to children's memory abilities generally. It may have been the case that participants with low adaptability scores felt uncomfortable correcting the interviewer in the misleading condition. Importantly, adaptability is a broadly defined concept, tapping into resilience (Folke et al., 2010), and also one's ability to learn a task that is complex, novel, or ill-defined. The latter element has been linked to openness in the five-factor model of personality (Le Pine et al., 2000). Some past research has shown a positive association between openness and memory and eyewitness performance (e.g., Curley, MacLean, & Murray, 2017; Rasmussen & Berntsen, 2010). Particularly, it is theorised that those higher in openness make better eyewitnesses due to the positive association between openness and implicit learning (DeYoung et al., 2012), which is also a characteristic of adaptability (Miyamoto, Wang, & Smith, 2020).

Furthermore, as distractibility impacted on performance across all conditions it suggests that temperament characteristics limiting attention may cause greater difficulty during encoding and/or retrieval of information. Distractibility likely impacts

both the stages of encoding and retrieval. For example, more distractible children may pay less attention during an event, therefore encoding fewer details. This is supported by findings suggesting that hyperactive children perform less well on memory tests (e.g., Melnyk, 2002; Pezdek & Roe, 1995; Warren et al., 1991). Their attention may also be more easily averted during interviews, and, therefore, they will recall fewer details to an interviewer. More research is needed to investigate what can be done to overcome these difficulties, and how interviews might be able to be tailored to suit child witnesses based upon their temperament. For example, distractible children may perform better having several short breaks during the interview, and less adaptable children may provide more information having fewer environmental changes to adapt to (e.g., having their favourite toy present during interviews, conducting the interview in their bedroom). These are avenues for future research to investigate. So far, few witness studies have examined the use of breaks during forensic interviews with young children. Studies have shown, however, that attention in the classroom can be improved with young children by the use of breaks (Saywitz & Camparo, 2014).

#### **4.4.3 Conclusions**

Any laboratory research on eyewitness memory is limited to some extent in its validity. Children are frequently called to testify about cases of sexual abuse and about having witnessed intrafamilial murder (McDonald et al., 2006). Clearly, no laboratory study can ever come close to mimicking the trauma and personal significance connected to such cases. Additionally, children in the study were asked to describe unfamiliar actors, while most crimes against children are committed by offenders who are familiar to the child (e.g., Davies & Noon, 1991). Still, there are important findings here about the general descriptive capabilities of children that have significant implications for police investigations and future research. Statistics suggest that between 10% and 25% of suspects involved in cases with child eyewitnesses are strangers to the child (Davies & Noon; Finkelhor & Jones, 2012, as cited in Carroll County Child Advocacy Centre, 2015; Whealin, 2007, as cited in Carroll County Child Advocacy Centre, 2015, so this is certainly not an insignificant amount.

Based upon the findings of the current study, future research should investigate what can be done to accommodate more distractible children during interview processes, as

well as to accommodate less adaptable children and less persistent children. Though adaptability and persistence were only significant predictors of eyewitness performance during the misleading condition, these are still important findings since interviewers often struggle to maintain best interview practices with young children (e.g., Luther et al., 2014; Roberts & Cameron, 2015; Waterhouse et al., 2018). Another important avenue for future research to explore is how interviewing techniques and temperament may impact the perceived accuracy and confidence of child witnesses, as confidence from an eyewitness in their own accounts remains one of the most persuading factors for jurors (Nicholson et al., 2014), even though recorded confidence by child-witnesses does not always reflect actual accuracy (Brewer & Day, 2005), even less so than in cases of adult witnesses (Keast et al., 2007; Tenney, Small, Kondrad, Jaswal, & Spelman, 2011). Having now established that eyewitness temperament affects eyewitness performance, the focus in the next chapter shifts to consider how eyewitness temperament influences perceptions of jurors.

## **Chapter 5: Study three – How temperament and interviewing techniques affect jurors' perceptions of witness credibility**

---

### **5.1 Experiment one: Introduction**

The third study contained two experiments. Experiment One aimed to investigate the influence of interview question styles and shyness in child witnesses on the interpretation of reliability, confidence, and compelling-nature of child-witness testimony by mock-jurors.

In cases of child sexual abuse, the child's testimony is often the only available evidence (Brewer et al., 1997; Lamb & Brown, 2006). In these instances, there is about a 33% chance of a case being prosecuted (Cross et al., 1994; Walsh et al., 2008). A case is significantly more likely to be prosecuted if there is additional evidence available, such as independent witnesses to the child's disclosure or a suspect's confession (Ernberg & Landström, 2016; Walsh et al., 2008). As mentioned in the previous chapter, children are also frequently required to testify in a number of other cases, such as domestic violence, kidnapping, and intrafamilial murder (e.g., McDonald et al., 2006).

On the whole, child witnesses are generally viewed as being honest by jurors (e.g., Nunez et al., 2011; Ross et al., 1990, 2003), but they are also typically perceived as having poorer memories than adult witnesses (e.g., Bottoms & Goodman, 1994; Goodman et al., 1987; Newcombe & Bransgrove, 2007; Nikonova & Ogloff, 2005; Pozzulo & Dempsey, 2009; Ross, et al., 1990; 2003). However, these findings generally come from studies asking participants to compare the reliability of a child witness's statement to that of an adult-witness. In reality, jurors are likely to have only the child witness's testimony available (Brewer et al., 1997; Lamb & Brown, 2006). Therefore, it is imperative to understand what factors impact the perceived reliability of testimony from child eyewitnesses, and not just how that perceived reliability compares to that of an adult eyewitness. Furthermore, different results are often found depending on the type of crime and on whether the witness is also the victim in the case. Older children and adults are typically viewed as more reliable than younger children

(especially those under 6 years old), but younger children can be viewed as more reliable in certain situations, such as being a victim-witness in a sexual abuse case (e.g., Goodman et al., 1987; Pozzulo & Dempsey, 2009).

In child maltreatment trials, jurors consider child testimony the most important evidence (Myers et al., 1999). Furthermore, confidence from an eyewitness in their own account remains one of the most persuading factors for jurors overall (Nicholson, et al., 2014). This is despite the fact that confidence does not always reflect actual accuracy (Brewer & Day, 2005), even less so than it does in cases of adult witnesses (Keast et al., 2007; Tenney et al., 2011). Nevertheless, jurors are likely to perceive a response as correct if an eyewitness expresses certainty in being correct (Brewer & Burke, 2002; Brigham & Wolfskeil, 1983; Cutler, Penrod & Dexter, 1990). Furthermore, the time taken for an eyewitness to respond to a question may be taken as another indicator of confidence and reliability (e.g., Dunning & Peretta, 2002; Robinson, Johnson, & Herndon, 1997; Shaw, 1996). Jurors may be more sceptical of eyewitness testimony in child sexual abuse cases if they have knowledge about suggestibility and cognitive deficits, especially if the child is a bystander-witness and not a victim-witness (Goodman, Golding, & Haith, 1984).

### **5.1.1 Questioning style**

Another factor that has been found to influence eyewitnesses is the style of questioning. Specifically, leading questions are often used in cross-examination to the detriment of witness confidence and accuracy (e.g., Gous & Wheatcroft, 2020; for a review, see Morrison, Forrester-Jones, Bradshaw, & Murphy, 2019; Zajac, Westera, & Kaladelfos, 2018). In a study by Kebbell et al. (2010), adult participants viewed a video of a woman being attacked by a man and were individually questioned about the incident. Half of the participants were asked questions using six categories of confusing questions often used by lawyers, while the other half were asked for the same information using simply phrased alternatives. In a second study by the authors, they investigated how the question styles impacted on the interpretations of mock-jurors. They found that lawyers' use of confusing questions significantly reduced accuracy, speed of response, and jurors' ability to determine accuracy.

Researchers studying child eyewitness testimony have also highlighted the importance of the questioning styles used to evoke children's responses. Goodman et al. (1999) claimed that it is a technique used by lawyers to discredit the testimony of child witnesses. This is because when leading questions are used by lawyers, it can be difficult for jurors to discern whether the child actually knows the answer or is simply complying with the lawyer's suggestion. Using the scenario of a murder within their study, Kalra and Heath (1997) found that 6-year-old bystander-witnesses were rated as significantly less credible when leading questions were used, and mock-jurors were also significantly less confident in their guilty verdicts. Similarly, Tubb et al. (1999) found that a 9-year-old's reports of sexual abuse were perceived as significantly less credible and the defendant was significantly less likely to be perceived as guilty when the disclosure was elicited by police through suggestive questioning (participants were presented with one of two types of hearsay evidence: written transcripts or second-hand testimony by a police officer). In a study by Olaguez and Klemfuss (2020), mock-jurors read a transcript from either a direct examination or a cross-examination with a child witness in a sexual abuse case. Those who read the direct examination rated the child as being significantly more credible. Overall, there is evidence to suggest that how child-witnesses' statements are obtained influences the interpretation of mock-jurors.

### **5.1.2 Perceived confidence and shyness of witnesses**

As mentioned in the Literature Review, some research has found that leading questioning can decrease a witness's confidence (Wheatcroft et al., 2004), but findings in this area are generally mixed (Wheatcroft, 2002). The researcher is not aware of any studies that have focused specifically on children regarding the relationship between leading questioning and perceived confidence. Additionally, the researcher is not aware of any studies that have investigated the connection between witness shyness and perceived credibility. Since shyer individuals tend to have a lack of confidence in their everyday life and are more likely to question themselves, as well as rely on others for information (Crozier, 1982), as witnesses, shyness may have an effect on juror perceptions of their confidence, thereby making them less credible to potential jurors. In support of this, Brodsky, Griffin, and Cramer (2010) found that witnesses rated

higher in shyness were perceived as less confident when testifying by a panel of judges, while witnesses higher in assertiveness were rated as significantly more credible than less assertive witnesses in a separate study (Larson & Brodsky, 2014). Furthermore, expert witnesses with high eye contact typically receive significantly higher credibility ratings from jurors than those with medium and low eye contact, and shyer individuals usually report decreased eye contact preference (Neal & Brodsky, 2008). Lastly, when looking at child witnesses, in an unpublished master's thesis, those regarded as more talkative were rated significantly more favourably than less talkative child witnesses by mock-jurors (Pierce, 2019). Even though shyness was not a significant predictor of eyewitness performance in the previous study, it is not necessarily the case that performance and perception are linked.

### 5.1.3 Research aims

The aim of the experiment was to investigate the impacts of question types (i.e., open-ended, closed-ended, and misleading) and child witness shyness on the interpretation of mock-jurors, who watched a video of the eyewitnesses from the previous study. The *first hypothesis* is that child witnesses (4 to 8 years old) will be perceived as less reliable, confident, and compelling by mock-jurors if they undergo closed-ended questions or misleading questions compared to those who undergo open-ended questions (Castelli, Goodman, & Ghetti, 2005; Kalra & Heath, 1997; Kebbel et al., 2010; Tubb et al., 1999). The *second hypothesis* is that high shy child witnesses will be interpreted as less reliable, confident, and compelling by mock-jurors than low shy children, despite the fact that shyness was not a significant predictor of eyewitness performance in the previous study as perceived reliability and actual reliability are not always in agreement (e.g., Bidrose & Goodman, 2000). This is because shy children may be perceived as less confident and, therefore, as less accurate (Brodsky et al., 2010; Pierce, 2019) due to shyer individuals having less confidence in their own testimonies and finding it harder to maintain eye contact (Crozier, 1982; Neal & Brodsky, 2008).

## 5.2 Experiment one: Methods

### 5.2.1 Design

A 2 x 3 mixed factorial design was used. The between-subjects factor was the questioning style (open-ended, closed-ended, or misleading). Participants were

allocated to one condition at random. The within-subjects factor was witness shyness (low, moderate, or high), measured in a previous study (see Chapter Four). Participants viewed a witness statement from each of the three conditions. The dependent variables were how reliable, confident, and compelling the mock-jurors perceived each witness to be measured on a 9-point Likert scale.

### 5.2.2 Participants

A total of 97 participants (76 females, 18 males, 3 did not provide a gender identity) took part in the study, ranging from 18 to 73 years old ( $M = 26.21$ ,  $SD = 9.84$ ). Fifteen of the participants watched videos of the child witnesses in the presence of the researcher. Due to the outbreak of COVID-19, face-to-face recruitment was cancelled. All subsequent participants read transcripts of the interviews online.<sup>9</sup> This meant that some aspects of shyness (e.g., body language) might not have been detected by participants who read the transcripts, though analysis revealed no significant differences in responses between those who watched the videos and those read the transcripts (see **Appendix L**). Participants responded to requests shared by the researcher and his supervisors on social media. For a moderate effect size ( $\eta^2 = .50$ ), a power of .80, and a threshold of  $p = .05$ , the total sample size needed to detect the effect is 42, calculated with the software G\*Power (Faul et al., 2009). The sample size in the study is comparative to previous research (e.g., Karla & Heath, 1997).

### 5.2.3 Materials

Initially, participants were asked to read an information sheet and sign a consent form, as well as provide demographic information (**Appendix M**). Participants then watched videos or read transcripts of interviews with child participants from the previous study (see Chapter Four). The interviews presented to participants were selected on the basis of being as comparable as possible in terms of accuracy, age, number of descriptors reported during the interview, and gender (see **Table 5.1** for similarities and differences between videos). The length of the video was not controlled for due to the small selection of videos available to the researcher and since videos had been matched on number of descriptors, which was thought to be more important (as

---

<sup>9</sup> Online participants did not view videos of the interviews due to ethical concerns that they might be able to record the interviews with a mobile device and, therefore, retain copies of the interviews.

longer videos may have included pauses). During the interviews, children were questioned on what they remembered from a bicycle theft video using one of three interview formats (open-ended, closed-ended, or misleading). The interviews ranged from 7 to 12 minutes. The participants in the video group were shown the interviews as videos, while the transcript group were presented with verbatim transcripts of the same interviews. After watching each video or reading each transcript, participants were asked to complete a juror perception questionnaire (**Appendix N**), reporting how reliable, detailed, confident, compelling, focused, shy, and articulate they believed the child to be, as well as how biased the interviewer was perceived to be. Responses were recorded on a 9-point Likert scale. The questionnaire was inspired by Wheatcroft et al.'s (2004) study, which had a similar design to the current study.

**Table 5.1.**

*Interviews with child witnesses used for the study.*

	Open condition	Closed condition	Misleading condition
<b>Low-shyness child (scored 11-15 on shyness scale)</b>	7-year-old boy (46 correct descriptors, 11 errors, 81% accurate)	7-year-old girl (38 correct descriptors, 8 errors, 83% accurate)	7-year-old girl (48 correct descriptors, 12 errors, 80% accurate)
<b>Moderate-shyness child (scored 22)</b>	8-year-old girl (49 correct descriptors, 11 errors, 82% accurate)	8-year-old girl (48 correct descriptors, 9 errors, 84% accurate)	8-year-old girl (44 correct descriptors, 13 errors, 77% accurate)
<b>High-shyness child (scored 30-33)</b>	8-year-old boy (44 correct descriptors, 10 errors, 81% accurate)	6-year-old girl (38 correct descriptors, 13 errors, 75% accurate)	8-year-old girl (31 correct descriptors, 9 errors, 78% accurate)

### 5.2.4 Procedure

Approval for the study was received from Edinburgh Napier University's School of Applied Science Research Integrity Committee. Furthermore, permission was received from the parents of participants from the previous study (see Chapter Four) for their video interviews to be used. Participants were allocated to one of three conditions at

random (open-ended, closed-ended, or misleading; see **Table 5.1**.) Within each condition, participants viewed three separate interviews with three child witnesses, differing in levels of shyness (low, moderate, high), determined in a previous study (see Chapter Four). The order of interviews was randomised for each participant. Subsequent to each interview, participants were asked to complete a juror perception questionnaire, in order to provide details over their interpretation of the interview. Participants were then debriefed (**Appendix O**) and thanked for their time after all three interviews and completing three juror perception questionnaires.

## 5.3 Experiment one: Results

### 5.3.1 Manipulation checks

#### 5.3.1.1 *Perceived bias of interview questions*

A manipulation check was conducted to ensure mock-jurors were able to detect the interviewer as biased in the closed-ended and misleading conditions (see **Table 5.2**). Since the rationale for the hypothesis was that mock-jurors would perceive child-witnesses as less credible when given leading questions, due to mock-jurors then no longer being sure if the child was answering because they knew the answer or were simply agreeing with the interviewer, it was imperative that participants were able to detect these questions as more biased than open-ended questions. Indeed, there was a statistically significant difference in perceived bias between groups as determined by one-way ANOVA,  $F(2, 278) = 36.46, p < .001$ , partial  $\eta^2 = .21$ . A Tukey post hoc test revealed that the interviewer was perceived as significantly less biased in the open-ended condition compared to the closed-ended ( $p < .001$ ) and misleading ( $p < .001$ ) conditions. There was no significant difference between the closed-ended and misleading ( $p > .05$ ) conditions.

#### 5.3.1.2 *Perception of shyness levels*

Similarly, a manipulation check was conducted to ensure the mock-jurors were able to accurately detect the shyness of the child witnesses based upon self-reported shyness levels (see **Table 5.3**). In fact, there was a statistically significant difference in perceived shyness by mock-jurors between shyness groups as determined by one-way ANOVA,  $F(2, 278) = 3.16; p = .044$ , partial  $\eta^2 = .03$ . A Tukey post hoc test revealed that the children determined to be low in shyness based upon their self-reports were

perceived as significantly less shy by mock-jurors than those rated as being high in shyness ( $p = .037$ ). All other differences were non-significant ( $p > .05$ ).

**Table 5.2.**

*Perceptions of child witnesses based upon interview condition.*

	Open condition	Closed condition	Misleading condition
<b>Reliable (1-9)</b>	4.78 ( $SD = 1.94$ )	5.00 ( $SD = 1.95$ )	5.37 ( $SD = 2.01$ )
<b>Confident (1-9)</b>	4.52 ( $SD = 2.14$ )	4.58 ( $SD = 2.20$ )	5.35 ( $SD = 2.21$ )
<b>Compelling (1-9)</b>	4.30 ( $SD = 2.03$ )	4.36 ( $SD = 2.16$ )	4.99 ( $SD = 2.02$ )
<b>Biased (1-9)</b>	3.71 ( $SD = 1.76$ )	5.82 ( $SD = 2.11$ )	5.60 ( $SD = 2.20$ )

**Table 5.3.**

*Perceptions of child witnesses based upon shyness levels.*

	Low-shyness	Moderate-shyness	High-shyness
<b>Reliable (1-9)</b>	5.22 ( $SD = 1.94$ )	4.97 ( $SD = 2.14$ )	4.86 ( $SD = 1.82$ )
<b>Confident (1-9)</b>	5.13 ( $SD = 2.22$ )	4.72 ( $SD = 2.21$ )	4.49 ( $SD = 2.15$ )
<b>Compelling (1-9)</b>	4.80 ( $SD = 2.01$ )	4.49 ( $SD = 2.12$ )	4.27 ( $SD = 2.08$ )
<b>Perceived shyness (1-9)</b>	4.05 ( $SD = 2.06$ )	4.56 ( $SD = 2.25$ )	4.83 ( $SD = 2.16$ )

### 5.3.2 Interview format and shyness

Three two-way ANOVAs were run to determine if interview condition and shyness level had a significant effect on any of the dependent variables (i.e., how reliable, confident, and compelling the child witness was perceived to be by mock-jurors). A MANOVA was not appropriate due to the strong, positive correlations between the dependent variables. All assumptions were met for two-way ANOVA. No outliers were identified using boxplots. The three dependent variables were approximately normally distributed for each interview condition, according to Shapiro-Wilk test of normality ( $ps > .05$ ), and there was homogeneity of variances according to Levene's test ( $ps > .05$ ). Descriptive statistics are displayed in **Tables 5.2** and **5.3**. There was no statistically significant difference in reliability between interview conditions as determined by two-

way ANOVA,  $F(2, 272) = 2.45$ ;  $p = .088$ , or between shyness levels,  $F(2, 272) = 1.33$ ;  $p = .267$ . However, there was a statistically significant interaction (see **Figure 5.1**) between interview condition and shyness level on perceived reliability,  $F(4, 272) = 8.91$ ,  $p < .001$ , partial  $\eta^2 = .12$ . The interactions are explored further in the following sections.

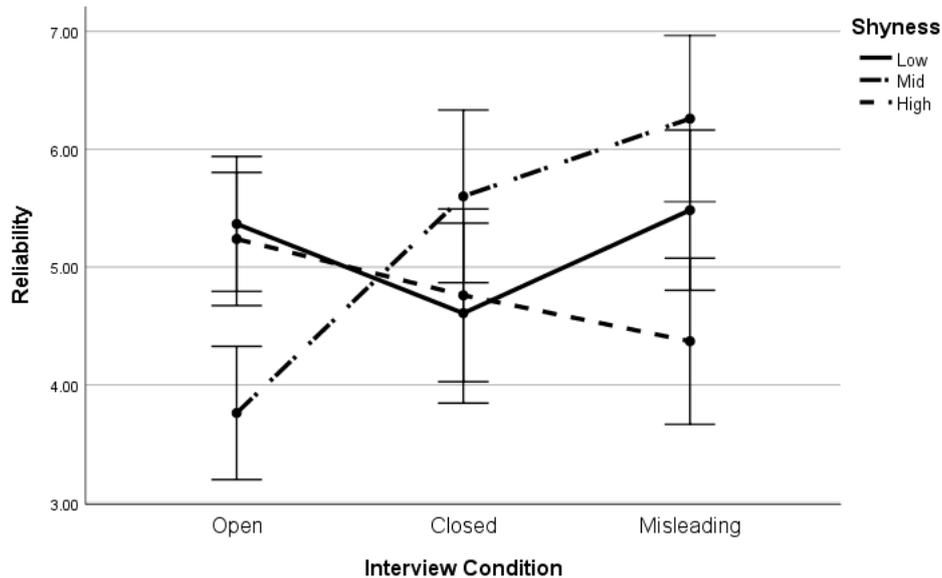


Figure 5.1. Interaction between interview on condition and shyness level on mean reliability.

There was no statistically significant difference in perceived confidence between shyness levels,  $F(2, 272) = 2.00$ ;  $p = .137$ , but there was a statistically significant difference between interview conditions,  $F(2, 272) = 4.55$ ;  $p = .011$ , partial  $\eta^2 = .03$ . A Tukey post hoc test revealed that the children in the misleading condition were perceived as significantly more confident than those in the open-ended ( $p = .012$ ) and closed-ended ( $p = .049$ ) conditions. Furthermore, there was a statistically significant interaction see (**Figure 5.2**) between interview condition and shyness level on perceived confidence,  $F(4, 272) = 10.24$ ,  $p < .001$ , partial  $\eta^2 = .13$ .

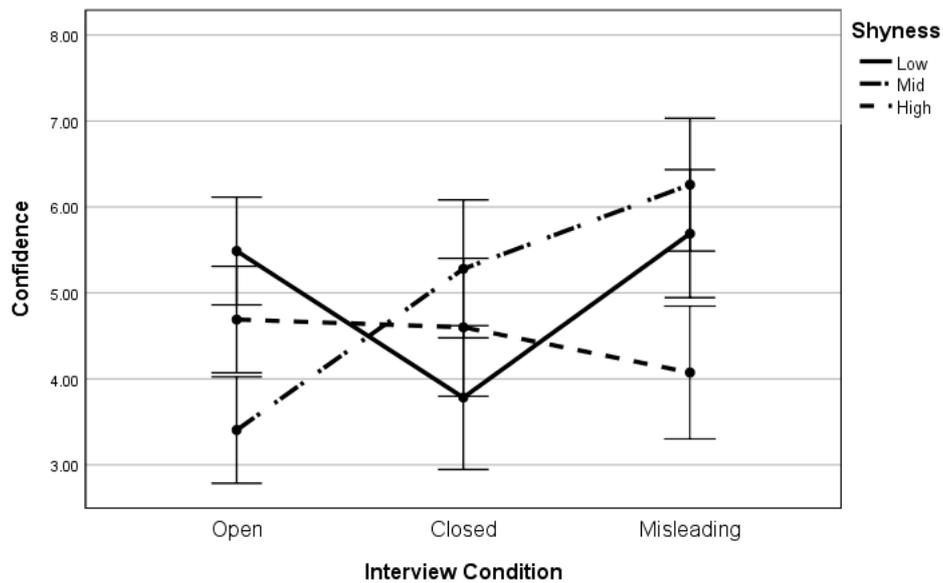


Figure 5.2. Interaction between interview on condition and shyness level on mean confidence.

There was no statistically significant difference in how compelling witnesses were perceived to be between shyness levels,  $F(2, 272) = 2.52$ ;  $p = .082$ , but there was a statistically significant difference between interview conditions,  $F(2, 272) = 3.57$ ;  $p = .030$ , partial  $\eta^2 = .03$ . A Tukey post hoc test revealed that the children in the misleading condition were perceived as significantly more compelling than those in the open-ended condition ( $p = .030$ ). There was a statistically significant interaction (see **Figure 5.3**) between interview condition and shyness level on how compelling the witnesses were,  $F(4, 272) = 11.54$ ,  $p < .001$ , partial  $\eta^2 = .15$ .

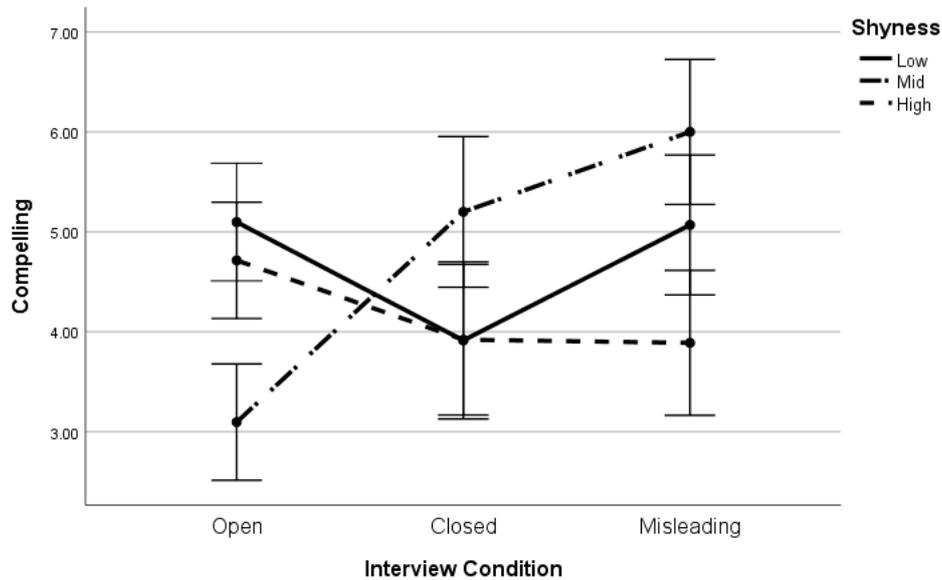


Figure 5.3. Interaction between interview on condition and shyness level on mean compelling.

5.3.2.1 *Shyness: Open-ended condition*

To investigate further the interaction effects between interview condition and shyness level on the three dependent variables, three separate ANOVAs were run for each interview condition. In the open-ended condition (see **Table 5.4**), there was a statistically significant effect of shyness level on perceived reliability,  $F(2, 122) = 10.16$ ;  $p < .001$ ; partial  $\eta^2 = .14$ , perceived confidence,  $F(2, 122) = 11.74$ ;  $p < .001$ ; partial  $\eta^2 = .16$ , and how compelling the witness was found to be by mock-jurors,  $F(2, 122) = 13.76$ ,  $p < .001$ , partial  $\eta^2 = .18$ . Tukey post hoc tests revealed that the moderate-shy child was perceived as significantly less reliable, confident, and compelling than the low-shy child ( $p < .05$ ) and high-shy child ( $p < .05$ ). All other differences were non-significant ( $p > .05$ ).

**Table 5.4.**

*Perceptions of child witnesses based upon shyness levels in the open-ended condition.*

	Low-shyness	Moderate-shyness	High-shyness
<b>Reliable (1-9)</b>	5.37 ( $SD = 1.96$ )	3.76 ( $SD = 1.88$ )	5.24 ( $SD = 1.56$ )
<b>Confident (1-9)</b>	5.49 ( $SD = 2.34$ )	3.40 ( $SD = 1.84$ )	4.69 ( $SD = 1.72$ )
<b>Compelling (1-9)</b>	5.10 ( $SD = 2.06$ )	3.10 ( $SD = 1.62$ )	4.71 ( $SD = 1.85$ )

5.3.2.2 *Shyness: Closed-ended condition*

Within the closed-ended condition (see **Table 5.5**), there were no statistically significant differences reported based on shyness level for any of the dependent variables ( $ps > .05$ ).

**Table 5.5.**

*Perceptions of child witnesses based upon shyness levels in the closed-ended condition.*

	Low-shyness	Moderate-shyness	High-shyness
<b>Reliable (1-9)</b>	4.61 ( <i>SD</i> = 1.70)	5.60 ( <i>SD</i> = 2.10)	4.76 ( <i>SD</i> = 1.94)
<b>Confident (1-9)</b>	3.78 ( <i>SD</i> = 1.59)	5.28 ( <i>SD</i> = 2.23)	4.60 ( <i>SD</i> = 2.48)
<b>Compelling (1-9)</b>	3.91 ( <i>SD</i> = 1.86)	5.20 ( <i>SD</i> = 2.08)	3.92 ( <i>SD</i> = 2.31)

5.3.2.3 *Shyness: Misleading condition*

In the misleading condition (see **Table 5.6**), there was a statistically significant effect of shyness level on perceived reliability,  $F(2, 80) = 3.53$ ;  $p = .002$ ; partial  $\eta^2 = .147$ , perceived confidence,  $F(2, 80) = 8.46$ ;  $p < .001$ ; partial  $\eta^2 = .175$ , and how compelling the witness was found to be by mock-jurors,  $F(2, 80) = 8.87$ ,  $p < .001$ , partial  $\eta^2 = .182$ . Tukey post hoc tests revealed that the high-shy child was perceived as significantly less confident and compelling than the low-shy child ( $p \leq .05$ ) and the moderate-shy child ( $p < .05$ ). Furthermore, the high-shy child was perceived as significantly less reliable than the moderate-shy child ( $p = .001$ ). All other differences were non-significant ( $p > .05$ ).

**Table 5.6.**

*Perceptions of child witnesses based upon shyness levels in the misleading condition.*

	Low-shyness	Moderate-shyness	High-shyness
<b>Reliable (1-9)</b>	5.48 ( <i>SD</i> = 2.05)	6.26 ( <i>SD</i> = 1.51)	4.37 ( <i>SD</i> = 2.02)
<b>Confident (1-9)</b>	5.69 ( <i>SD</i> = 2.09)	6.26 ( <i>SD</i> = 1.43)	4.07 ( <i>SD</i> = 2.43)
<b>Compelling (1-9)</b>	5.07 ( <i>SD</i> = 1.91)	6.00 ( <i>SD</i> = 1.41)	5.00 ( <i>SD</i> = 2.02)

**5.3.3 Perceived shyness**

Pearson correlations were run to determine the relationship between perceived shyness and the dependent variables. There were weak, negative correlations between

perceived shyness and perceived reliability,  $r(95) = -.18$ ;  $p = .002$ , perceived shyness and perceived confidence,  $r = -.32$ ;  $p < .001$ , and between perceived shyness and how compelling mock-jurors found the witness,  $r = -.21$ ;  $p < .001$ .

### **5.3.4 Summary: Results**

Overall, there was no significant effect of questioning style on the perceived reliability of the witnesses. Those in the misleading condition, however, were perceived as more confident and compelling than those in the open-ended condition, as well as more confident than those in the closed-ended condition. There were no overall effects of shyness on the perceived reliability, confidence or compelling nature of the witnesses. There were interaction effects between questioning style and shyness level for all three of the dependent variables. In the open-ended condition, the moderate-shy child was rated as being less reliable, confident, and compelling than the low-shy and high-shy child. There were no significant effects in the closed-ended condition. In the misleading condition, the high-shy child was rated as being less reliable, confident, and compelling than the moderate-shy child, as well as less confident and compelling than the low-shy child.

## **5.4 Experiment one: Discussion**

The aim of this study was to better understand potential influences on perceptions of child-witnesses, specifically children's shyness levels and the interviewer's questioning style.

### **5.4.1 Questioning style**

The first hypothesis of the study was that child witnesses would be perceived as less reliable, confident, and compelling by mock-jurors if they underwent closed-ended or misleading questions as opposed to open-ended questions. The results revealed there were no statistically significant differences based on interview conditions in terms of how reliable the witnesses were perceived to be by mock-jurors. Instead, participants overall rated all witnesses as being near the midpoint. However, child witnesses in the misleading condition were rated as being significantly more confident and compelling than those in the open-ended condition, as well as more confident than those in the closed-ended condition.

These are surprising findings given that Kebbell et al. (2010) found the use of lawyers' confusing questions significantly reduced accuracy, speed of response, and jurors' ability to determine accuracy compared to simply phrased alternatives. Furthermore, Kalra and Heath (1997) found that 6-year-old bystander-witnesses were rated as significantly less credible when leading questions were used. Of course, ecological validity is limited since the researcher in the study may have been perceived differently by participants than lawyers would be by child witnesses. However, the previously mentioned studies have also used experimenters as questioners, generating significant results. Within the current study, mock-jurors rated the interviewer as being significantly more biased within the closed-ended and misleading conditions compared to within the open-ended condition. Therefore, mock-jurors did regard questions within the misleading condition as leading but this did not affect their perception of the witnesses in terms of reliability.

Given that interviews were selected for the study on the basis of being as comparative as possible in terms of accuracy, the child witnesses within the misleading condition's interviews were particularly resistant to misleading questions. It is possible that participants picked up on this, and, therefore, found child witnesses to be more reliable because of it. This may explain why these witnesses were rated as significantly more confident than those in the open-ended and closed-ended conditions. Additionally, it may be the case that the current study was not tapping into all categories of confusing questions from Kebbell et al.'s (2010) study. For example, when asking leading questions, these authors also used negatives (e.g., 'Did the attacker not grab the woman around the waist?') and double negatives (e.g., Complex: 'Is it not true that the woman did not go into the house?'). It may be that future studies have to better reflect questions used by lawyers, much like Kebbell et al. (2010) managed to do. In the present study, however, unlike Kebbell et al.'s (2010), the witnesses included young children, and so the questions had to be both understandable and answerable to all age groups. Evidence suggests that questions containing double negatives are particularly difficult for young children to understand (e.g., Brennan & Brennan, 1988, as cited in Hanna, Davies, Crothers, & Henderson,

2011; Perry et al., 1995, as cited in Hanna et al., 2011; Saywitz, Camparo, & Romanoffz, 2010).

Furthermore, it may be the case that mock-jurors were sceptical of children's reports, regardless of interview format. Some of the participants were psychology students, who were possibly educated on the dangers of eyewitness testimony. For example, one participant within the open-ended condition, commented that, '[It is] hard to answer questions on the reliability of children as they are an unreliable witness due to age, ability to communicate, and memory etc.' While other studies have also asked participants to rate how leading the questions in the study were (e.g., Karla & Heath, 1997), it is possible this allowed the participants in the current study to place the blame on the interviewer rather than the child since they were familiar with eyewitness interview procedures (i.e., they were unwilling to say the child was less credible because they believed the child's credibility was directly affected by the questions they had been asked).

Additionally, the seriousness of the crime may have had an impact. In some previous studies (e.g., Karla & Heath, 1997), child witnesses have been describing cases such as murder. Comparatively, participants in the current study may have thought that a bike theft was not very important, and therefore there was less chance of witnesses paying attention to it.

#### **5.4.2 Shyness**

The second hypothesis of the experiment was that shyer witnesses would be perceived as less reliable, confident, and compelling by mock-jurors than less shy witnesses.

There is limited evidence to support this prediction. Overall, level of shyness had no significant effect on the perceived reliability, confidence or compelling nature of child witnesses according to mock-jurors. This is despite the fact that less shy children were rated as being less shy, meaning the mock-jurors were detecting the shyness of the children, at least to some extent.

This is the first study that the researcher is aware of that has examined whether the shyness of child witnesses has an impact on the interpretation of mock-jurors. In the open-ended condition, the moderate-shy child was rated as being less reliable, confident, and compelling than both the low-shy and high-shy child. This finding may have been related to shyness. For example, mock-jurors may have regarded the low-shy in the open-ended condition as being overconfident. However, it is also possible that the moderate-shy was rated as being more credible due to reasons not controlled for by the researcher. In the misleading condition, the high-shy child was rated as being less reliable, confident, and compelling than the moderate-shy child, as well as less confident and compelling than the low-shy child.

It is unclear, however, how generalisable these findings are. As mentioned, there are other factors potentially responsible for the results. For example, the low-shy child in the open-ended condition at one point in the interview said, 'that's what I imagine him to be wearing' when describing the clothes worn by the man in the video. The researcher interpreted this as a poor choice of verb by the child and understood the child to mean that they had reported the clothing details that they remembered. However, mock-jurors may have interpreted it differently, believing the child to mean they were imagining details in their report. Therefore, even if mock-jurors regarded this witness to be confident, they may have considered the witness to be unreliable, regardless. In hindsight, this is something that the researcher should have ascertained with the child during the interview, as it was explained in the interview instructions to children that they should only report what they remember.

Furthermore, the moderate-shy child within the open-ended condition, rated as being significantly less reliable, confident, and compelling than the low-shy child, and as being significantly less reliable and compelling high-shy child, was the only female child-witness within the open-ended condition. It is unlikely, however, that this would have had an impact on mock-juror interpretations given the majority of participants were reading transcripts. While the literature is conflicting in regard to whether or not witness-gender impacts juror interpretations, prior research has focused on sexual

abuse cases (e.g., Bottoms & Goodman, 1994; Haegerich & Bottoms, 2000; Waterman & Foss-Goodman, 1984). In these studies, boys have been regarded as less credible witnesses when victims to sexual abuse, with mock-jurors reporting to believe that boys are better able to avoid being victims of such crimes than girls. It does not seem therefore that these findings would apply to the present study, despite recent findings that mock-jurors still view male sexual assault victims less favourably than female victims (e.g., Sommer, Reynolds, & Kehn, 2016). To err on the side of caution, however, future research should aim to counterbalance witness-gender if possible.

## **5.5 Experiment two: Introduction**

During the first experiment, it was observed that there was significant within-group variation in scores (i.e., differences in scores were not only stemming from factors the researcher had manipulated or that were related to the children, but also from the mock-jurors themselves). For instance, standard deviations were often between two and three for responses to items on the 9-point questionnaire, even when looking only at participants in the same condition as each other, suggesting large disagreements in general attitudes towards child-witnesses. To investigate this further, a follow-up experiment was set up, examining personality questionnaire responses from some of the mock-jurors to determine if this accounted for a significant portion of the variance.

### **5.5.1 Personality and the Big Five model**

Personality is what creates patterns of behaviours, thoughts, and feelings (Allport, 1937), often characterised in terms of traits, which refer to labels that can be applied to relatively stable patterns of behaviour (Kreitler & Kreitler, 1990, p. 4). Personality, therefore, can have a significant influence on our perception. For example, people's political beliefs can be predicted based upon their personality. Findings indicate that more orderly people are significantly more likely to tilt towards political conservatism than less orderly people (Hirsch et al., 2010). This suggests that our inbuilt personality constitutes a set of filters through which we view the world, and that this alters the manner in which we process information, including the way in which we vote.

Currently, the Big Five is the most widely accepted model of personality (Larsen & Buss, 2001; Funder, 2019), stemming from Allport's lexical hypothesis that natural

language has evolved terms for all fundamental personality differences (McCrae & Costa, 1985a). The Big Five is the result of various researchers over the course of several decades identifying clusters of adjectives that tend to go together in describing particular personality traits, consistently showing the same five categories of descriptions across time and cultures (e.g., Digman, 1990, as cited in Goldberg, 1993; Digman & Inouye, 1986; Goldberg, 1990, 1992, 1993; Fiske, 1949, as cited in Schultz & Schultz, 2005; Norman, 1963; Tupes & Christal, 1961, as cited in Goldberg, 1993; Saucier & Goldberg, 1996). These five traits are openness (tendency to be curious and unconventional), conscientiousness (tendency to be organised and disciplined), extraversion (tendency to be outgoing and enthusiastic), agreeableness (tendency to be sympathetic and cooperative), and neuroticism (tendency to be anxious and insecure). They can be remembered using the acronym OCEAN.

Some researchers have suggested that the Big Five be referred to by Roman numerals I–V, because the labels are not useful (Funder, 2019). This is because they are necessarily oversimplified and potentially misleading. Importantly, each of the five factors are 'not so much one thing but more of a collection of many things that have something in common' (Saucier & Goldberg, 2003, p.14). In order to combat this problem, Costa and McCrae divided each of the Big Five into six facets. Costa and McCrae's Revised NEO Personality Inventory (or NEO PI-R; 1992) is currently the most popular method of assessing the Big Five (Soto & John, 2009). It contains 240 items and six subcategories for each of the five factors. The subcategories are referred to by the authors as lower order facets, representing the distinct, though co-varying, elements within a trait (Costa & McCrae, 1995). They limited the number of facets to six because they felt 'more than six would soon lead to intellectual overload' (Costa & McCrae, 1995, p.26-27).

The NEO PI-R has proven to have high internal consistency, and test-retest reliability scores, as well as to be valid with children (Markey, Market, & Tinsley, 2004). A great deal of cross-cultural research has also been carried out using the NEO PI-R, finding robust evidence through factor analysis of the Big Five in China, Estonia, Finland,

France, German-speaking countries, India, Mexico, Portugal, Russia, South Korean, the Philippines, Turkey, Vietnam, and Zimbabwe, as well as an additional 51 countries (McCrae & Terracciano, 2005). However, it should be noted that studies have failed to replicate the five-factor structure in some parts of the world, such as with an indigenous society in the Amazon (Gurven, von Rueden, Massenkoff, Kaplan, & Vie, 2013).

Although the same factors are common to many cultures, there are major differences in regard to their relative importance (Allik & McCrae, 2004; McCrae & Terracciano, 2005). Japanese residents consider conscientiousness to be more important than all the other factors. In Hong Kong and India, agreeableness was found to be the most important factor. In Australia, extraversion and agreeableness are considered to be more desirable than the other three factors. Overall, Europeans and Americans tend to score higher in extraversion and openness and lower in agreeableness compared to Asians and Africans. The distribution of personality traits also varies depending on geographical location in the United States. For example, agreeable people are more likely to be found in the South-eastern United States, and openness is highest in areas near Denver, Los Angeles, Miami, New York City, Portland, San Antonio, Seattle, and San Francisco (Florida, 2008).

The Big Five personality traits can be identified through factor analysis using questionnaires not even designed to measure the five factors, including Cattell's 16PF and Eysenck Personality Questionnaire. For instance, the traits extraversion and neuroticism of Eysenck's Personality Questionnaire are represented in the traits extraversion and neuroticism of the Big Five, while the third personality trait of Eysenck's model, psychoticism, is represented in the traits agreeableness and conscientiousness (Aluja, Garcia, Garcia, 2002; Digman, 1997; Goldberg, 1993; Markon, Krueger, & Watson, 2005; McCrae & Costa, 1985b, 1987). There is also a considerable degree of consistency between self-ratings and observer-ratings of the Big Five traits, further supporting their reliability (Connolly, Kavanagh, & Viswesvaran, 2007; Costa & McCrae, 1986, 1988; McCrae et al., 2004).

An original idea behind the Big Five is that they are orthogonal, meaning that being higher or lower in one of the traits is not predictive of being higher or lower in another. However, DeYoung (2006) has shown that the five factors are not as orthogonal as originally believed. In fact, there is a small, positive correlation between conscientiousness, agreeableness, and emotional stability (the reverse of neuroticism), and there is a small, positive correlation between openness and extraversion. In DeYoung's (2006) model, extraversion and openness create plasticity (to explore and engage flexibly with novelty), while conscientiousness, emotional stability, and agreeableness create stability (to maintain stability and avoid disruption in emotional, social, and motivational domains). DeYoung (2010) suggests that these meta-traits may have a biological basis. Specifically, evidence is accumulating to suggest that plasticity is related to dopamine and dopamine-related brain structures, while stability is related to serotonin and serotonin-related brain structures (Allen & DeYoung, 2017). This is because serotonin facilitates the stabilisation of emotion and motivation, as well as the inhibition of aggressive and impulsivity, while dopamine facilitates exploration, approach behaviour, and flexible cognitive functioning (Shiner & DeYoung, 2013).

Furthermore, DeYoung (2006) empirically and conceptually separated each of the Big Five into two lower-level aspects. According to this approach openness is made up of openness-to-experience and intellect, conscientiousness is made up of orderliness and industriousness, extraversion is made up of enthusiasm and assertiveness, agreeableness is made up of politeness and compassion, and neuroticism is made up of withdrawal and volatility. This model provides more independent predictive power (DeYoung, Quilty, & Peterson, 2007). For example, orderliness is a stronger and more reliable predictor of social conservatism than conscientiousness, and the same is true with compassion as a predictor of social liberalism as opposed to agreeableness (Hirsch et al., 2010).

Though not all psychologists accept this model, the Big Five remains the most popular trait approach among personality researchers (Buss & Larsen, 2001; Funder, 2019). Some researchers have proposed more than five personality dimensions, and others have argued that no list of factors can fully describe the complexity of human personality (Ashton et al., 2004; Ashton & Lee, 2008; Di Blas, 2005; Lanning, 1994; Lee & Ashton, 2008; Paunonen, 2002, as cited in Buss & Larsen, 2001; Paunonen, Haddock, Forsterling, & Keinonen, 2003; Saucier, Georgiades, Tsaousis, & Goldberg, 2005). Nonetheless, the Big Five has been widely replicated and has inspired a great deal of research. There is evidence that the five factors of personality emerge early in childhood (Shiner & DeYoung, 2013), are relatively stable (for a review, see Roberts, Walton, & Viechtbauer, 2006), appear across cultures, have neurological correlates (DeYoung, 2010), and have strong heritable components (Bouchard, 1994; Bouchard & Loehlin, 2001; DeFruyt et al., 2006; Jang, McCrae, Angleitner, Riemann, & Livesley, 1998; McGue, Bacon, & Lykken, 1993; Larsson, Andershed, & Lichtenstein, 2006; Riemann, Angleitner, & Strelau, 1997; Yamagata et al., 2006).

The biggest strength of the model though is its ability to predict future behaviour. For example, conscientiousness is a strong predictor of having a healthy lifestyle and living longer (Connor-Smith & Flashbart, 2007; Tucker, Elliott, & Klein, 2006; Walton & Roberts, 2004). This is because conscientious people are less likely to engage in negative health behaviours, such as drinking alcohol, drug use, and smoking. Furthermore, they are more likely to engage in positive health behaviours, such as exercising and dieting. Longitudinal studies, some investigating the same people for nearly 70 years, have shown that children scoring higher in conscientiousness are more likely to be physically healthier in adulthood, and to live longer than children who scored lower in conscientiousness (Booth-Kewley & Vickers, 1994; Friedman et al., 1993, 1995; Marshall, Wortman, Vickers, Kusulas, & Hervig, 1994).

A number of other outcomes can be predicted using the Big Five, including grades in school (Chamorro-Premuzic & Furnham, 2003), risky sexual behaviour (Allen & Walter, 2018; Miller et al, 2004; Trobst, Herbst, Masters, & Costa, 2002; Zietsch, Verweij,

Bailey, Wright, & Martin, 2010), alcohol consumption (Paunonen, 2003), pathological gambling (Bagby et al., 2007), volunteering (Carlo, Okun, Knight, de Guzman, 2005), declining to become a union member (Parkes & Razavi, 2004), forgiveness (Brose, Rye, Lutz-Zois, & Ross, 2005), leadership effectiveness (Hassan, Asad, & Hoshino, 2016; Silverthorne, 2001), and compliance with social distancing guidelines during a pandemic (Abdelrahman, 2020). The potential impact that the Big Five can have on jurors will now be discussed.

### **5.5.2 Juror personality**

Early research examined the influence of jury personality traits on attitudes related towards the legal system, such as on death penalty trials (Nietzel, McCarthy, & Kern, 1999) and insanity defence cases (Cutler, Moran, & Narby, 1992). The traits examined included locus of control (Phares & Wilson, 1972; Sosis, 1974), empathy (Moran & Comfort, 1986), and authoritarianism (Devine, Clayton, Dunford, Seying, & Pryce, 2001). This area of research was largely abandoned due to small effect sizes (Greene et al., 2002). However, almost all of this research was conducted before the emergence of the Big Five personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism).

Only a small number of studies have examined the impact of the Big Five personality traits on jury decision making, specifically focusing on jury deliberations. On a sample of 17 juries ( $n = 285$ ), Clark et al. (2007) found that extraversion was associated with being chosen as jury foreperson. Furthermore, the more extraverted the foreperson, the longer deliberations lasted and the more influence the foreperson was perceived to have by other jurors. Extraversion was also associated with being influential in jury studies by Marcus, Lyons, and Guyton (2000) and Rotenberg, Hewlett, and Siegart (1998). These studies investigated the impact of personality on jury deliberation, but as far as the researcher is aware, no research has examined the potential influence of juror personality on perceptions of child witnesses.

Though not jury studies, a significant amount of research surrounding the Big Five has generated relevant findings. As mentioned previously in this chapter, more recent

research has found that each of the Big Five personality traits can be empirically divided into two distinct aspects in order to provide more independent predictive power (though much of the research is in the realm of political psychology, offering greater predictive power of political orientations than previous Big Five assessments). The aspects of the Big Five are measured by the Big Five Aspect Scale (DeYoung et al., 2007).

#### *5.5.2.1 Openness*

Openness reflects aesthetic sensitivity, creativity, and intellectual curiosity (McCrae, 1987; Ziegler, Cengia, Mussel & Gerstorf, 2015; Weisberg et al., 2011). It is a predictor of divergent thinking and creative achievement (Batey & Furnham, 2006; Carson, Peterson, & Higgins, 2003; Feist, 1998; Feist & Barron, 2003; King, Walker, & Broyles, 1996; Kaufman et al., 2016; McCrae, 1987; Silvia, Nusbaum, Berg, Martin, & O'Connor, 2009). Individuals high in openness tend to enjoy novelty and fiction, as well as pastimes like art, movies, poetry, philosophical discussions, and unconventional music. Individuals higher in openness are more likely to enjoy engaging in fantasies, to be curious, and seek new experiences and cultures (Schwaba, Luhmann, Denissen, Chung, & Bleidorn, 2017; Zhiyan & Singer, 1997).

According to DeYoung's (2006) model, openness is divided into two aspects: openness-to-experience and intellect. The two aspects independently predict separate outcomes. For instance, intellect predicts general intelligence, as well as verbal and nonverbal intelligence, while openness-to-experience is only associated with verbal intelligence (DeYoung, Quilty, Peterson, & Gray, 2014). Furthermore, openness-to-experience is positively associated with increased interest in novel stimuli, whilst intellect is predictive of increased understanding of such stimuli (DeYoung et al., 2014). While social liberalism is associated with openness in general, a finding that holds across cultures and measurements, it is only independently associated with openness-to-experience, not intellect (Burton et al., 2015; Carney et al., 2008; Hirsch et al., 2010; Sibley et al., 2012; Xu et al., 2013). According to Carnet et al. (2008), this is because individuals adopt viewpoints that provide the best fit with the needs rooted in their psychological dispositions. Those higher in openness are typically lower in dogmatism

and a need for cognitive closure (Mondak & Halperin, 2008; Onraet, Van Hiel, Roets, & Cornelis, 2011).

The latter may hold relevance for jury perceptions. For example, if a witness's statement seems incomplete, this may be of less concern to a jury member higher in openness. Importantly, however, those higher in openness are not more persuadable, despite earlier predictions to the contrary, in general, or even when that information is tailored toward motivations rooted in openness (Bekker, 2014). Openness strongly, negatively correlates with ageism, anti-feminism, ethnocentrism, Right Wing Authoritarianism, Social Dominance Orientation, transphobia, and prejudice (Allan, Johnson, & Emerson, 2013; Galton, Hammond, & Stinchcombe, 2020; McCrae & Costa, 1997; Nagoshi et al., 2008; Platt & Szoka, 2019; Sibley & Duckitt, 2000). Therefore, jurors higher in openness should be less likely to discriminate against witnesses, defendants, or victims based on age, race, sexual orientation, or gender.

#### *5.5.2.2 Conscientiousness*

Conscientiousness is associated with self-discipline and organisation (Weisberg et al., 2011). Conscientiousness is the best personality predictor of success in education (Dumfart & Neubauer, 2016), sometimes even accounting for a larger portion of the variance than intelligence (Kappe & van der Flier, 2012). This is likely because conscientious people are more likely to attend classes and to submit assignments on time. They tend to make efficient use of their time, finding inactivity adverse, and are more likely to experience guilt as a result of using their time unproductively (Kertechian, 2018).

According to DeYoung's (2006) model, conscientiousness is divided into orderliness and industriousness. Orderliness is predictive of social conservatism (though low openness is a stronger predictor), but industriousness is not (Hirsch et al., 2010). Individuals who score higher in this aspect are more likely to rely upon the moral foundations of loyalty, authority, and purity, rather than care and fairness, than those who score lower (Graham, Nosek, Haidt, Iyer, Koleva, & Ditto, 2011). Therefore, they are more likely to view it as wrong for a soldier to disobey orders they disagree with or

for a child to not show respect to their parent, even if that parent has done something terrible to the child (Haidt, 2012). Orderly people are more likely to be higher in disgust sensitivity, meaning they find mess and dirt intolerable (Xu, Plaks, & Peterson, 2016).

They are more likely to view violations of group-orientated rules as deserving of punishment and to be higher in disgust sensitivity (Alford, Hibbing, & Smith, 2013; Hirsch et al., 2010; Xu et al., 2016). There are a number of potential implications here for jurors. For example, jurors higher in orderliness may believe violations of certain moral foundations are deserving of harsher punishments, particularly those related to loyalty, authority, and purity. It is not so clear how conscientiousness may be related to the perceptions of witnesses.

### *5.5.2.3 Extraversion*

The concept of extraversion was first introduced by Carl Jung (1921, as cited in Jung, 2014), referring to whether one orientates themselves towards external (e.g., socialising) or internal experiences (e.g., reading, thinking). However, Jung's observations have been criticised for only being based upon those in his own life (Boeree, 2006). Nowadays, extraversion is commonly characterised in terms of sociability and outgoingness (Costa & McCrae, 1992), though it also involves assertiveness (DeYoung et al., 2007; Goldberg, 1992; Shiner & DeYoung, 2013) and one's tendency to experience positive emotions (Depue & Collins, 1999; Gray & McNaughton, 2003; Robinson, Moeller, & Ode, 2010; Tellegen et al., 1988).

Extraversion is relatively stable across the lifespan and is a powerful impactor on behaviour (Roberts et al., 2006). For example, it is quite difficult for introverts to act like extraverts, and vice versa. Extraverts report higher levels of happiness and are more sensitive to rewards (Gray, 1970, as cited in Gray & McNaughton, 2003). They are also more likely to have a higher number of sexual partners and to have children at a younger age (Eysenck & Eysenck, 1975, as cited in Larsen & Buss, 2001). Extraverts are more likely to process their thoughts externally, seek out time for socialising, and to have a larger variety of friends (Asendorpf & Wilpers, 1998). Introverts, on the other

hand, typically dislike being the centre of attention, and avoid large groups. Introverts are also more emotionally reserved, and more likely to seek time alone to think and recharge (Duffy et al., 2018).

According to DeYoung's (2006) model, extraversion is divided into assertiveness and enthusiasm. The distinction was made previously by Depue and Collins (1999), in order to reflect the difference in dopamine- and opiate-mediated reward functioning, described by DeYoung (2010) as the difference between 'wanting' and 'liking' a reward. The association between extraversion and jury deliberation has been discussed previously in this chapter, but it is not so clear how extraversion may be connected to the perception of witnesses. Though the researcher is not aware of research directly tied to perceptions of child witnesses, research suggests jurors higher in extraversion are significantly more likely to view expert witnesses favourably (Cramer, Brodsky, & DeCoster, 2009). The authors theorised this was due to extraverted jurors being more likely to seek out positive emotions and therefore being more likely to focus on aspects they like of a witness, meaning this could go beyond perceptions of expert witnesses and potentially apply to perceptions of child witnesses too.

#### *5.5.2.4 Agreeableness*

Agreeableness is associated with altruism and cooperation (Weisberg et al., 2011). Agreeable people are self-sacrificing, compassionate, and polite (DeYoung et al., 2007). Highly agreeable people will put another person's concerns ahead of their own. They are generally altruistic, compliant, cooperative, empathetic, modest, non-competitive, and conflict-averse (Costa & McCrae, 1992). Low agreeableness is a predictor of deviant behavior, as well as low conscientiousness (Salgado, 2002; Vize, Lynam, Lamkin, Miller, & Pardini, 2016). High agreeableness is an indicator of lower income (Matz & Gladstone, 2020). This may be because they are less likely to ask for pay rises in the first place and are also more likely to be less competent negotiators. Disagreeable people are argumentative and like to have everything their own way without regard for others. Extremely low agreeableness is predictive of many personality disorders, including narcissism and psychopathy (Vernon, Villani, Vickers, & Harris, 2008). As is the case with any trait, very few people score extremely high or

low. Instead, most people score near the middle of the spectrum, leaning one way a little more than the other (Funder, 2019).

According to DeYoung's (2006) model, agreeableness is divided into compassion and politeness. Agreeableness overall has been closely linked to higher activity levels and volume within several structures of the brain (e.g., left dorsolateral prefrontal cortex, superior temporal sulcus, posterior cingulate cortex; DeYoung, 2010; Allen & DeYoung, 2017). Politeness has been associated with a lower baseline testosterone level (DeYoung et al., 2013; Montoya, Terburg, Bos, & van Honk, 2012; Turan et al., 2014). The combination likely allows for the ability to suppress aggression and regulate emotions. Along with neuroticism, it is one of the two Big Five dimensions with the largest and most reliably reported gender differences, with women tending to score higher than men on both aspects of each dimension (Costa, Terracciano, & McCrae, 2001; Feingold, 1994; Weisberg et al., 2011), though considerable gender differences in neuroticism do not appear until adolescence (Parker & Brotchie, 2010).

Agreeableness is positively associated with economic liberalism, but not social liberalism until one gets down to the aspect level (Hirsch et al., 2010). Specifically, compassion is positively associated with social and economic liberalism, whilst politeness is positively associated with social conservatism (Gerber et al., 2010, 2011; Hirsch et al., 2010; Riemann, Grubich, Hempel, Mergl, & Richter, 1993). Compassion reflects empathy, sympathy, and prosocial behaviour, while polite individuals are more likely to be compliant, cooperative, and respectful (DeYoung, Weisberg, Quilty, & Peterson, 2013). Jurors higher in compassion may, therefore, have greater sympathy for a victim-witness.

#### *5.5.2.5 Neuroticism*

Neuroticism is associated with tendencies to experience negative emotion, including anger, anxiety, depression, and self-consciousness (Kale et al., 2020; Weisberg et al., 2011). Individuals who score high on neuroticism are more likely to report lower self-esteem, and to be more self-conscious and shy (Judge et al., 1997). They are also more at risk for developing common mental disorders, such as mood disorders and anxiety

disorders (Jeronimus et al., 2016). Highly neurotic people are particularly sensitive to threats, including social threats, such as a fear that they may not be accepted or liked by others. There is a negative correlation between neuroticism and happiness, well-being, and physical health (Connor-Smith & Flashbart, 2007; DeNeve & Cooper, 1998).

In DeYoung's (2006) model, neuroticism is made of the aspects withdrawal and volatility. The distinction is linked to Gray and McNaughton's (2000, as cited in Shiner & DeYoung, 2013) theory that neuroticism is linked to two systems—the behavioural inhibition system (BIS) and the fight-flight-freeze system (FFFS). The FFFS, involving the amygdala and lower regions of the brain, responds to a threatening stimulus in terms of flight (panic) or fight (reactive aggression). The BIS, on the other hand, centred around the hippocampus, responds to a stimulus that one desires but that contains threat, creating an approach-avoidance conflict (Shiner & DeYoung, 2013), such as wanting to ask someone on a date but fearing rejection or wanting to impress your teacher with a presentation but being worried about judgement from peers. Clearly, the anger or panic, connected to the FFFS, is reflected in volatility, while anxiety and self-consciousness, connected to the BIS, is reflected in withdrawal. It is not yet clear how neuroticism may be connected to perceptions of jurors.

### **5.5.3 Research aims**

To investigate further the varied responses in Experiment One, the research aims were to investigate the impact of personality on the perceptions of child witnesses according to mock-jurors.

## **5.6 Experiment two: Methods**

### **5.6.1 Design**

After completing Experiment One, participants completed a personality assessment (DeYoung et al.'s (2007) Big Five Aspect Scale) to examine if there was a significant correlation between personality traits and measures of juror perception, completed by 21% of participants overall.

### **5.6.2 Participants**

The final 20 participants in Experiment One (15 females, 5 males), averaging at 23.71 years old ( $SD = 7.52$ ), completed a personality assessment. Importantly, for a moderate

effect size and a Type 2 error threshold of  $p = .05$ , this sample size only gives a power of .70, calculated with the software G\*Power (Faul et al., 2009). Therefore, any significant findings will have to be replicated with a larger sample, but the analysis may still help understand how perceptions of the participants in the current study were impacted.

### **5.6.3 Materials**

Means and internal consistencies for each aspect of the Big Five Aspect Scale (DeYoung et al., 2007) are reported in **Table 5.7**. See **Appendix P** for items.

### **5.6.4 Procedure**

Permission for a revised version of Experiment One was received by Edinburgh Napier University's School of Applied Science Research Integrity Committee. Participants completed the stages of Experiment One, and then were asked to complete the Big Five Aspect Scale (DeYoung et al., 2007). Afterwards, they were fully debriefed (see **Appendix O**).

## **5.7 Experiment two: Results**

Pearson correlations were run to determine the relationship, if any, between each of the personality aspects and how reliable, confident, compelling, and shy mock-jurors regarded the child witnesses (see **Appendix Q**). One respondent's data was removed due to being identified as an outlier by box plots in several dimensions. There were moderate, positive relationships between withdrawal and three of the dependent variables, including perceived reliability,  $r(17) = .47$ ;  $p = .048$ , perceived confidence,  $r(17) = .55$ ;  $p = .019$ , and how compelling the child witness was,  $r(17) = .49$ ;  $p = .041$ . There were also significant correlations between perceived shyness and three of the personality dimensions. This included positive correlations with agreeableness,  $r(17) = .49$ ;  $p = .039$ , and compassion,  $r(17) = .59$ ;  $p = .010$ , and a negative correlation with assertiveness,  $r(17) = -.51$ ;  $p = .032$ . All other correlations were non-significant ( $ps > .05$ ).

**Table 5.7.**

*Big Five Aspect Scale means and internal consistencies (Cronbach's Alpha).*

	Mean	SD	A
Openness	3.71	.50	.86
Openness	3.97	.49	.71
Intellect	3.45	.62	.84
Conscientiousness	3.25	.40	.74
Orderliness	3.44	.57	.77
Industriousness	3.05	.52	.75
Agreeableness	4.15	.54	.92
Compassion	4.28	.54	.88
Politeness	4.02	.61	.86
Extraversion	3.22	.45	.78
Assertiveness	3.01	.61	.78
Enthusiasm	3.44	.64	.84
Neuroticism	3.03	.55	.86
Withdrawal	3.18	.62	.77
Volatility	2.88	.71	.88

## 5.8 Experiment two: Discussion

The aim of the experiment was to investigate if personality traits of mock-jurors would impact how they perceived the child witnesses. In fact, mock-jurors higher in withdrawal, an aspect of neuroticism, were more likely to view child witnesses as being reliable, confident, and compelling than mock-jurors lower in withdrawal. It is not surprising that judgements over the characteristics of a witness may be impacted by the personalities of jurors as one's personality influences their typical social perceptions (e.g., Hannuschke, Gollwitzer, Geukes, Nestler, & Back, 2020). For example, higher neuroticism typically predicts one being more likely to display heightened sensitivity to imagined threats, in social situations and non-social situations (e.g., Denissen & Penke, 2008).

In the current study, those higher in withdrawal may have perceived child witnesses more positively because they identified with any perceived anxiety or discomfort on the part of the witnesses. Alternatively, there may have been a contrast effect, as found by Hannuschke et al. (2020), when 110 psychology freshmen were more likely to make positive judgements over the perceived sociability and warmth of others if they were higher in neuroticism. According to this explanation, those higher in withdrawal hold less favourable views of their own abilities and so perceive others as having more positive abilities compared to themselves. In support of this, those higher in neuroticism are more likely to underestimate how their own self-esteem will be rated by others (e.g., Barrett & Pietromonaco, 1997), as well as to overestimate the self-esteem of others (e.g., Kilianski, 2008).

Furthermore, mock-jurors higher in agreeableness, specifically compassion, and lower in assertiveness, an aspect of extraversion, were more likely to view child witnesses as being shy. Those higher in extraversion typically perceive more positive qualities in others than those lower in extraversion (e.g., Barrett & Pietromonaco, 1997), including when determining the credibility of expert witnesses (Cramer et al., 2009). This may be because they are more likely to detect characteristics that are congruent with their own. It may also be the case that one has to be at least somewhat shy themselves in order to pick up on cues of shyness in others (see Chapter Three), explaining why those higher in assertiveness were less likely to rate the children as being shy, regardless of the child's self-report level of shyness. Those higher in compassion are more likely to be empathic (Shiner & DeYoung, 2013), therefore more sensitive to the pain of others (Loggia, Mogil, & Bushnell, 2008). Of course, these explanations are ones developed post hoc. Therefore, they require further validation.

## **5.9 Conclusions**

Overall, in Experiment One, even though mock-jurors rated the interviewer as being more biased within the closed-ended and misleading conditions than the open-ended condition, they did not rate child witnesses within these conditions as being any less reliable. In fact, they rated the witnesses in the misleading condition as being more confident than those in the open-ended and closed-ended conditions, as well as more compelling than those in the open-ended condition. Therefore, though jurors may

have awareness over the potentially negative effects of closed-ended and leading questioning, there was no evidence to suggest that these questioning styles lessen a child's credibility. This is inconsistent with prior research (Castelli et al., 2005; Karla & Heath, 1997; Kebbel et al., 2010; Tubb et al., 1999). It may be the case that participants in this study were sceptical of child witnesses, though questionnaire scores were generally at the midpoint, regardless of whether other factors were present or not. Future research should focus on gathering a broader sample, including those who are not psychology students and may have less knowledge of shortcomings of eyewitness memory.

By focusing only on closed-ended questions and misleading questions, the study may also have not been tapping into other questions reflective of those used by lawyers, such as confusingly worded questions. It may also be the case that the open-ended questions in this study were not open-ended enough, seeing as how they were mostly follow-up questions (e.g., 'Tell me the colour of the bike.'). Additionally, in an attempt to have interviews with children that were comparable in terms of accuracy, the children in the misleading condition may have been regarded as being resistant to leading questions by mock-jurors. To investigate whether or not this is a sufficient explanation of the study's findings, the study could be repeated, or a similar study could be employed, using a range of accuracy scores across conditions.

Furthermore, the study did not support the hypothesis that shyer children may be regarded as less credible by mock-jurors than less shy children. Within the open-ended and misleading conditions, the low-shy child was rated as significantly more reliable than some of the more shy children, but there may have been other factors that made these specific children more credible as witnesses than the other children. In order to produce more generalisable results, future research should aim to counterbalance as many factors as possible, such as witness-age and witness-gender. The ecological validity of the study is also limited since the majority of participants read transcripts of interviews with witnesses, rather than see them testify. It is worth noting though that analysis revealed no significant differences in results between those who watched the

videos and those who read the transcripts (see **Appendix L**). Furthermore, since the majority of previous studies have used transcripts (e.g., Castelli et al., 2005; Karla & Heath, 1997; Tubb et al., 1999), it means the results of this study can be compared to the findings of previous research more directly.

Importantly, the perceptions of the child witnesses were impacted by the personality traits of the mock-jurors themselves. Those higher in withdrawal were more likely to rate the child witnesses as being reliable, confident, and compelling. Higher agreeableness, specifically compassion, and lower assertiveness predicted mock-jurors rating child witnesses as being shy. Since the personality questionnaire was only completed by 21% of the overall participants, however, it cannot be said with certainty that these findings can be applied to the entire sample. Still, the main point here is that, in some ways, the self-reported shyness of the child, in this context, is not important if it is not detected by jurors. Either way, a juror will have their own constructed reality, including their own subjective opinion over the characteristics of a witness, and this reality is influenced by the juror's own psychological dispositions.

## **Chapter 6: General discussion**

---

The purpose of this research was to investigate children's eyewitness memory and the factors that may impact accuracy, as well as mock-juror perceptions of those witnesses. Specifically, the research aimed to examine how temperament and interviewing styles affect children as witnesses and mock-juror perceptions of child witnesses.

### **6.1 Research objectives**

The research objectives were as follows:

- ◇ To develop a new, self-report method for determining temperament traits in 4 to 8-year-old children (Study One);
- ◇ To investigate how interviewing techniques and temperament affect children as eyewitnesses (Study Two);
- ◇ To examine how interviewing techniques, shyness of child witnesses, and the personality traits of jurors impact mock-juror perceptions of child witnesses, (Study Three).

#### **6.1.1 Research objective 1**

There were three aims of the first study: (1) to investigate whether 4- to 8-year-old children have consistent and meaningful understandings regarding their own temperament characteristics, (2) to determine whether a new tool (the Temperament Assessment Tool for Children; TATC) discriminates between different children, and (3) to assess if these differences between children are stable over time. Two hundred and two 4 to 8-year-old children were presented with 48 pairs of statements that each represented the high- and low-end points of six temperament dimensions (activity, adaptability, distractibility, emotionality, persistence, and shyness). For example, 'I like talking to children I don't know' and 'I don't like talking to children I don't know' represented the low and high ends of shyness respectively. Children were then asked to pick the statement that best described themselves. Internal consistency was found to be satisfactory overall. Furthermore, the tool was able to discriminate between different children. Lastly, differences were found to be sufficiently stable over a 4-month period. Ratings were also gathered from parents, and interrater correlation was satisfactory, supporting the validity of TATC scores. The TATC provides temperament

data for research and clinical purposes that is derived from the views and experiences of young children themselves, rather than relying on questionnaire data from external raters.

### **6.1.2 Research objective 2**

The aim of this research was to better understand how the temperament of a child-witness and the questions they were asked may affect their eyewitness performance during an interview. Two-hundred and two 4-8-year-old children watched a video of a theft and were interviewed using open, closed, or misleading questions. The children also completed the TATC in Study One, a self-report method for determining temperament traits in young children, developed specifically for this study. Children made significantly more errors to closed-ended and misleading questions compared to open-ended questions and provided significantly more correct information to open-ended questions compared to misleading questions. Witnesses in the open-ended condition were significantly more accurate overall than the other interview conditions. Children who were more distractible made more errors and had an overall lower accuracy rate. Furthermore, children who were more adaptable and more persistent were less likely to be misled when asked misleading questions. Temperament characteristics that limit attention (e.g., distractibility) may cause greater difficulty during encoding and/or retrieval, and misleading questions are particularly detrimental to the accuracy of less adaptable and less persistent children. As far as the researcher knows, this was the first study to demonstrate self-reports of temperament are predictors of children's eyewitness performance.

### **6.1.3 Research objective 3**

This study aimed to demonstrate the influence of interview question styles and shyness in child witnesses on the interpretation of reliability, confidence, and compelling-nature of their interviews by mock-jurors. Ninety-seven mock-jurors (*Mean age* = 26.21, *SD* = 9.84) were randomly assigned to one of three interview conditions (open-ended, closed-ended, or misleading). The child-witnesses varied in levels of shyness (i.e., highly shy, moderately shy, and lowly shy) according to self-reports, gathered in Study One. Contrary to the researcher's predictions, the findings indicated that interview condition had no significant impact on the perceived reliability of children witnesses, even though mock-jurors identified closed-ended and misleading

interview styles to be significantly more biased than open-ended. Furthermore, child witnesses in the misleading condition were rated as being significantly more confident than those in the open-ended and closed-ended conditions, opposite to what was expected. This may be because children in this condition were regarded as being resistant to leading questions. There was no difference overall between shyness-levels on any of the dependent variables. Within the open-ended and misleading conditions, the low-shy child was rated as significantly more reliable than some of the more shy children, but there may have been other factors that made these specific children more credible as witnesses. The perceptions of jurors were impacted by their own personality traits. Particularly, those higher in withdrawal were more likely to rate child witnesses favourably, and those higher in compassion and lower in assertiveness were more likely to rate them as shy. This study may be the first to suggest that personality traits of mock-jurors impact their perceptions of child witnesses.

## **6.2 Impactors of witness performance**

As expected, in Study Two, closed-ended and misleading questions had an overall negative impact on children's performance as witnesses. In the open-ended condition, children's accuracy rates were 85% on average. This was statistically significantly higher than accuracy rates in both the closed-ended (79%) and misleading (63%) conditions. Children provided significantly more correct pieces of information in the open-ended and closed-ended conditions than the misleading condition. They provided more correct descriptors in the closed-ended condition than the open-ended condition, but this difference was not statistically significant. Furthermore, as indicated above, despite children providing more amounts of correct information in the closed-ended condition, this was not reflected in their overall accuracy rates. Children provided significantly more errors in both the closed-ended and misleading conditions than the open-ended condition, and more still in the misleading condition compared to the closed-ended condition. This is in agreement with prior research, demonstrating that overall accuracy is worse for closed-ended questions than open-ended questions, and then worse still for misleading questions (e.g., Memon et al., 1996).

In Study Two, age was a strong impactor on witness performance. Specifically, older children provided significantly more correct descriptors than younger children, as well

as significantly more errors and were overall less accurate. This is in line with a number of studies demonstrating age trends among child witnesses (e.g., Davies et al., 1989; Gagnon & Cry, 2017; Karageorge & Zajac, 2011; Kuehn, 1974; Lindsay et al., 1994; Pozzulo & Warren, 2003; Zajac & Karageorge, 2009). There are a number of developmental changes occurring between these age groups, discussed in more detail during the Literature Review, that may explain these findings, including the development of language abilities, theory of mind, memory, and social skills. Given that the older children provided more information than younger children in all conditions, it does support the idea that there may have been a progression in language abilities and memory in particular.

Though there was no significant interaction effect between age and interview format, the biggest differences in terms of overall accuracy were in the closed-ended and misleading conditions. Importantly, there were no age effects reported within the open-ended condition. This is in accordance with research showing that younger children (especially those aged 6 years and younger) are typically more suggestible than older children (especially those aged 8 years and over; e.g., Cassel et al., 1996; Ceci & Bruck, 1993; Gudjonsson et al., 2016; Otgaar et al., 2010, 2016; Paz-Alonso & Goodman, 2016; Poole & White, 1993; Sutherland & Hayne, 2001). According to some research, young children can be as accurate as older children and adults during free recall (e.g., Hershkowitz et al., 2012; Lamb et al., 2008; Marchant, 2013; Walker, 2013). This would suggest that interviewers should be especially careful about not using closed-ended and misleading questions with younger children.

One of the most interesting findings of the research was that temperament, measured via children's self-reports, had an impact on witness performance when they were questioned about an event. While previous research has found temperament traits to impact eyewitness performance (e.g., Benedan et al., 2020; Blackford, 2000; Burrell et al., 1999; Chen & Shapiro, 2000, as cited in Purdy, 2001; Greenhoot et al., 1999; Johnston et al., 2021; Melnyk, 2002; Memon et al., 1996; Palmer et al., 1998, as cited in Bruck & Melnyk, 2004; Purdy, 2001), results have been inconsistent, and this is the

first research the researcher is aware of to have used self-reports. Overall, children who were more distractible made significantly more errors and had significantly lower accuracy rates than less distractible children. Distractibility also positively correlated with errors in all three conditions and negatively correlated with overall accuracy in the closed-ended and misleading conditions. In the misleading condition, adaptability and persistence both positively correlated with the number of correct details reported, whilst adaptability negatively correlated with errors, and positively correlated with accuracy rate. Therefore, more adaptable and persistent children were more resistant to being misled.

Future research should further investigate the role of temperament using self-reports. Since Bruck and Melnyk's (2004) review on the impact of temperament on suggestibility in children, at least two other significant findings have emerged directly relevant to the present research. First, Benedan et al. (2020) found, using teacher ratings, that children with attention problems were significantly more suggestible. This is a similar finding to in the present study that more distractible children provided significantly more errors in all conditions. Second, Johnston et al. (2021) found that children lower in social flexibility, also rated by teachers, were significantly less likely to report transgressions during an interview. This construct is a combination of low adaptability and high shyness (Keogh, 1982). Since the two have been found to significantly correlate (Martin, 1988), and both have been inconsistently linked with eyewitness performance and suggestibility (for a review, see Bruck & Melnyk, 2004; Cotterill, 2017; Study Two in Chapter Four), it may make sense to investigate the two as one construct.

The present study also found that more persistent children were significantly more resistant to misleading questions, similar to previous findings using caregiver ratings (Chen, 2002; Greenhoot et al., 1999). Future research may, therefore, want to focus on three areas in particular—distractibility (or attention problems), social flexibility (or low adaptability and high shyness), and persistence—using a combination of self-reports, parent ratings, and teacher ratings. The present study finding that self-reports

were significant predictors of eyewitness performance is an important step, as it means that forensic interviewers can potentially gather self-reports of temperament from children. If forensic interviewers are able to identify highly distractible or socially inflexible children they may then be able to tailor the interview to make it more appropriate, such as by including many short breaks for distractible children (Saywitz & Camparo, 2014), though more research needs to be done to measure the potential effectiveness of this. This is information that may be useful to Registered Intermediaries in England and Wales; following these findings, Registered Intermediaries may be able to make recommendations to police during forensic interviews with children after challenging behaviour has been identified.

In Study Three, questioning style had no significant impact on the perceived reliability of the witnesses according to mock-jurors. Those in the misleading condition, however, were perceived as more confident than those in the open-ended and closed-ended conditions, as well as more compelling than those in the open-ended condition. This is contrary to prior research (e.g., Karla & Heath, 1997; Kebbell et al., 2010; Tubb et al., 1999). It may be because children were regarded as being highly resistant to the leading questions by mock-jurors. There were also no overall differences between the shyness levels of the child witnesses in the perceived reliability, confidence or compelling-nature according to mock-jurors. As far as the researcher knows, this is the first study to investigate this.

In the open-ended condition, the moderate-shy child was rated as being less reliable, confident, and compelling than the low-shy child. Similarly, in the misleading condition, the high-shy child was rated as being less reliable, confident, and compelling than the moderate-shy child, as well as less confident and compelling than the low-shy child. This would support the researcher's predictions that higher shyness makes a child witness less credible, in accordance with research finding that shyer children are more likely to doubt themselves and rely upon others for information (e.g., Crozier, 2000). However, there may have been other variables at play, such as perceived likability and sincerity (Cramer et al., 2009), that made certain children appear more

reliable, confident, and compelling than others. A wider range of variables should be investigated and controlled for by future researchers.

### **6.3 Methodological considerations**

Study One provided initial evidence of reliability and validity for a new, self-report measure of temperament. Internal consistencies were satisfactory, but recommendations were given to potentially improve the reliability scores in future studies. Test-retest reliability scores showed the self-reports to be stable. However, most of the participants who completed the self-report for a second time were older children. Furthermore, only about a third of the participants completed the self-report a second time. Therefore, further research is required to demonstrate the stability of self-report scores. Interrater correlation between the child ratings and a parent version was satisfactory, supporting initial validity. A persistence task demonstrated further validity. Furthermore, as self-report scores were able to predict performance during Study Two, this would also suggest high external validity. However, there may have been cases of social desirability during this study. For example, the children generally rated themselves as being less emotional than their parents rated them. Further research is required to fully demonstrate the validity of children's self-reports.

Study Two provides findings about the general descriptive capabilities of children. However, the validity is limited since it cannot mimic the complicated, and often traumatic, cases children are often witness to (e.g., McDonald et al., 2006).

Furthermore, the nature of the between-group design of the study may have limited the validity of the findings further. In real life, investigators will typically ask children a mix of open-ended, closed-ended, and leading questions (e.g., Lamb et al., 2006, 2018; Sternberg et al., 2001), rather than sticking to only one type of follow-up question. The misleading condition especially may have had high demand characteristics for the children (Horowitz, 2009); after all, once they were asked a number of misleading questions, the children may eventually have formed an interpretation of what the interviewer was doing. If they formed the interpretation that the interviewer was asking misleading questions, they may have become more resistant. On the other hand, they could have formed a different interpretation, such as that they were answering incorrectly and so the interviewer was challenging them, in which case they

may have been more likely to agree with the interviewer's suggestions. In future, when possible, researchers should ask a mix of questions to see whether effects remain or are diluted. This would also achieve a more ecologically valid interview style.

The researcher followed guidelines used by Police in Scotland, however, in which a number of steps are followed, beginning with requests for free recall, and only resorting to follow-up questions once free recall prompts have been exhausted. All interviews in Study Two, therefore, regardless of interview condition, began with requests for free recall. The results, therefore, demonstrate the impacts of pursuing with only one type of follow-up question once free recall has been given. The steps taken by the researcher also included rapport building and a practice interview in order to mirror true police procedures. These steps added to the ecological validity of the research.

The researcher made every effort to counterbalance as many variables as possible in Study Three based upon the interviews that were available. However, there are a number of limitations with this study. For example, the moderate-shy child within the open-ended condition was the only female child witness within that condition. It is unlikely this had an impact, given that participants were largely reading transcripts and the literature does not suggest witness-gender effects outside of sexual abuse cases (e.g., Bottoms & Goodman, 1994; Haegerich & Bottoms, 2000; Waterman & Foss-Goodman, 1984). The results of Study Three would be more though convincing had all other variables been successfully controlled for.

#### **6.4 Future recommendations**

Research in the field over the past few decades has been dedicated to understanding factors that impact eyewitness accuracy and suggestibility. Future research should build from these findings to understand what can be done to overcome these obstacles and test out procedures that may decrease the likelihood of suggestibility as much as possible.

A number of studies have demonstrated the potential impactors on suggestibility and witness performance (e.g., Alexander et al., 2002; Burgwyn-Bailes et al., 2001; Chae, 2004; Clarke-Stewart et al., 2004; Imhoff & Baker-Ward, 1999; Quas et al., 1999; Roebbers & Schneider, 2001; Rossi et al., 2011; Young et al., 2003). Future research should build on these findings and the results of Study Two in order to test what system variables during forensic interviews can be altered in order to overcome these impactors. For example, the findings of Study Two suggested that witnesses who are less adaptable to new environments experience greater difficulty during retrieval of information when given misleading questions. It may be the case that less adaptable children will be less vulnerable to suggestion if they have fewer environmental changes to adapt to. Future research could test if variables such as having the interview at the child's house or allowing the child to have their favourite toy present during the interview will lessen their vulnerability to misleading questions. Interviewers should be cautious though about using misleading questions during forensic interviews with children based on the results of Study Two and similar research (e.g., Lamb et al., 2006; Sternberg et al., 2001).

The researcher used age as discrete data in Study Two due to having a wide age range of participants and there being different witness capabilities between these age groups, as established by prior research (e.g., Ceci & Bruck, 1993). However, in retrospect, the researcher would have preferred to have recorded the ages of the participants in months, rather than years, so as to have the option of using age as continuous data. It is possible that information was lost when recorded in ranges, such as possible differences in witness capabilities between younger 5-year-olds and older 5-year-olds. Future research should, therefore, record age in terms of months so that it may be used as continuous data, thereby being more sensitive and carrying more statistical power.

Study Three was the first attempt the researcher is aware of to use videoed child-witness testimonies, as opposed to transcripts, in a mock-juror study. Unfortunately, face-to-face data collection was prohibited during the research, and Study Three was

completed using transcripts as an alternative. Future research should aim to use videoed testimonies as there may be impactors of juror decision making that transcript studies fail to tap into, such as how quickly a witness responds to a question. Though there were no significant differences between the results of those who watched the videos in Study Three and those who read the transcripts and this may suggest that the use of videos is not worth the logistical difficulties, this is based on a small sample. It is worth researchers exploring potential differences further, as using videos would be thought to improve ecological validity, given that jurors would not normally be asked to read transcripts of witness statements. Future research could explore a wider range of possible impactors on perceptions of child witnesses. For example, there is little research exploring effects, if any, of witness-gender outside of sexual abuse cases. Furthermore, most research in this field has investigated how individuals acting as jurors make their decisions, rather than groups of people (i.e., juror decision making, rather than jury decision making). When possible, researchers should aim to use mock juries to improve ecological validity, as there are instances of jurors changing their decisions due to the influence of other jurors (e.g., Clark et al., 2007; Golding et al., 2007).

Future research should also explore the wider implications of personality factors influencing the perceptions of mock-jurors. In Study Three, more neurotic mock-jurors, specifically those higher in withdrawal, were more likely to view witnesses positively, for example, and those higher in assertiveness were less likely to detect the shyness of a witness. The researcher has constructed post hoc explanations for these findings (e.g., more assertive jurors are less adept at detecting shyness because they have less experience themselves of what it feels like), however these require further validation. These results are also stemming from a small sample size. Additional studies should investigate the potential connection between personality traits and detecting shyness in a wider range of situations with a larger sample of participants. There are also implications here for how personality may impact mock-juror perceptions of witnesses more broadly. For example, jurors higher in dogmatism, associated with low openness (Mondak & Halperin, 2008), in one study were more likely to show racial bias towards

black defendants (Foley & Chamblin, 1982). Future research should investigate the impact of juror personality on perceptions of a wider range of witnesses.

## **6.5 Conclusions**

The development of the TATC (Study One) provides the first indication of reliable and valid self-reports from young children of the temperament dimensions noted by researchers as being relevant to eyewitness performance and suggestibility (e.g., Geddie et al., 2000; Greenhoot et al., 1999; Ornstein et al., 1999). With additional work, this tool can contribute to understanding of temperament and how it may impact children's eyewitness performance. The TATC may be used by interviewers, as well as by researchers, to help build rapport and to potentially identify challenging behaviour, such as high distractibility. There are a number of impactors of eyewitness performance identified by previous research. Study Two demonstrates the potential impact of temperament. Future research should aim to investigate what can be done to overcome the negative influences of certain temperament traits in order to tailor forensic interviews to suit children's individual temperament. In line with previous research, Study Two shows that witness capabilities are impacted by age. Specifically, younger children provided fewer correct descriptors, more errors, and were overall less accurate. Study Three demonstrated that the shyness of a child-witness and the personality factors of jurors may impact on the perceived credibility of witnesses according to mock-jurors, however future research is needed to build on these findings. The largest contribution to the field of the current research is the unique evidence that a reliable, self-report method for measuring temperament in young children can be used to predict eyewitness performance and suggestibility during an interview.

## References

---

- Abe, J. A. A., & Izard, C. E. (1999). A longitudinal study of emotion expression and personality relations in early development. *Journal of Personality and Social Psychology, 77*(3), 566-577. doi: [10.1037/0022-3514.77.3.566](https://doi.org/10.1037/0022-3514.77.3.566)
- Abdelrahman, M. (2020). Personality Traits, Risk Perception, and Protective Behaviors of Arab Residents of Qatar During the COVID-19 Pandemic. *International Journal of Mental Health Addiction*. doi: [10.1007/s11469-020-00352-7](https://doi.org/10.1007/s11469-020-00352-7)
- Ablow, J. C., & Measelle, J. R. (1993). *The Berkeley Puppet Interview: Administration and scoring system manuals*. Berkeley: University of California.
- Afshan, A., Askari, I., & Manickam, L. S. S. (2015). Shyness, self-construal, extraversion–introversion, neuroticism, and psychoticism: A cross-cultural comparison among college students. *SAGE Open, 5*(2), 1-8. doi: [10.1177/2158244015587559](https://doi.org/10.1177/2158244015587559)
- Ahadi, S. A., & Rothbart, M. K. (1994). Temperament, development, and the Big Five. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 189-207). Edbaum.
- Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality, 7*(5), 359-378. doi: [10.1002/per.2410070506](https://doi.org/10.1002/per.2410070506)
- Alexander, K. W., Goodman, G. S., Schaaf, J. M., Edelstein, R. S., Quas, J. A., & Shaver, P. R. (2002). The role of attachment and cognitive inhibition in children's memory and suggestibility for a stressful event. *Journal of Experimental Child Psychology, 83*(4), 262-290. doi: [10.1016/S0022-0965\(02\)00149-2](https://doi.org/10.1016/S0022-0965(02)00149-2)

## References

- Alford, J. R., Hibbing, J. R., & Smith, K. B. (2013). *Predisposed: Liberals, conservatives, and the biology of political differences*. Routledge.
- Allan, L. J., Johnson, J. A., & Emerson, S. D. (2014). The role of individual difference variables in ageism. *Personality and Individual Differences, 59*, 32-37. doi: 10.1016/j.paid.2013.10.027
- Allen, T. A., & DeYoung, C. G. (2017). Personality neuroscience and the five factor model. *Oxford Handbook of the Five Factor Model*. doi: 10.1093/oxfordhb/9780199352487.013.26
- Allen, M. S., & Walter, E. E. (2018). Linking big five personality traits to sexuality and sexual health: A meta-analytic review. *Psychological Bulletin, 144*(10), 1081–1110. doi: [10.1037/bul0000157](https://doi.org/10.1037/bul0000157)
- Allik, J., & McCrae, R. R. (2004). Toward a geography of personality traits: Patterns of profiles across 36 cultures. *Journal of Cross-cultural Psychology, 35*(1), 13-28. doi: 10.1177/0022022103260382
- Allport, G. W. (1937). *Personality: a psychological interpretation*. Holt.
- Aluja, A., García, Ó., & García, L. F. (2002). A comparative study of Zuckerman's three structural models for personality through the NEO-PI-R, ZKPQ-III-R, EPQ-RS and Goldberg's 50-bipolar adjectives. *Personality and Individual Differences, 33*(5), 713-725. doi: 10.1016/S0191-8869(01)00186-6
- Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review, 12*(6), 1043-1047. doi: 10.3758/BF03206441

## References

- Andrews, S., & Lamb, M. (2016). How do Lawyers Examine and Cross-Examine Children in Scotland?. *Applied Cognitive Psychology, 30*(6), 953-971. doi: 10.1002/acp.3286
- Ang, R. P. (2006). Effects of parenting style on personal and social variables for Asian adolescents. *American Journal of Orthopsychiatry, 76*(4), 503-511. doi: 10.1037/0002-9432.76.4.503
- Anwar, S., Bayer, P., & Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics, 127*(2), 1017-1055. doi: 10.1093/qje/qjs014
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioural and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213. doi: [10.1037/0033-2909.101.2.213](https://doi.org/10.1037/0033-2909.101.2.213)
- Ashton, M. C., & Lee, K. (2008). The prediction of Honesty–Humility-related criteria by the HEXACO and Five-Factor Models of personality. *Journal of Research in Personality, 42*(5), 1216-1228. doi: 10.1016/j.jrp.2008.03.006
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A Six-Factor Structure of Personality-Descriptive Adjectives: Solutions From Psycholexical Studies in Seven Languages. *Journal of Personality and Social Psychology, 86*(2), 356–366. doi: [10.1037/0022-3514.86.2.356](https://doi.org/10.1037/0022-3514.86.2.356)
- Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74*(6), 1531–1544. doi: [10.1037/0022-3514.74.6.1531](https://doi.org/10.1037/0022-3514.74.6.1531)
- Bagby, R. M., Vachon, D. D., Bulmash, E. L., Toneatto, T., Quilty, L. C., & Costa, P. T. (2007). Pathological gambling and the five-factor model of personality. *Personality and Individual Differences, 43*(4), 873-880. doi: 10.1016/j.paid.2007.02.011

## References

- Baker-Ward, L., Gordon, B. N., Ornstein, P. A., Larus, D. M., & Clubb, P. A. (1993). Young children's long-term retention of a pediatric examination. *Child development, 64*(5), 1519-1533. doi: 10.1111/j.1467-8624.1993.tb02968.x
- Bala, N. (2018, November 27). Children Witnesses in the Criminal Courts: Recognizing Competence and Assessing Credibility. Retrieved March 23, 2020, from <https://www.lawnow.org/children-witnesses-in-the-criminal-courts-recognizing-competence-and-assessing-credibility/>
- Ball, C. T., Pelco, L. E., Havill, V., & Reed-Victor, E. (2001). Confirmatory factor analysis of the temperament assessment battery for children-revised: Parent form. *Journal of Psychoeducational Assessment, 19*(4), 365-379. doi: 10.1177/073428290101900405
- Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs, 132*(4), 355-429. doi: 10.3200/MONO.132.4.355-430
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bartol, C. R. & Costello, N. (1976). Extraversion as a function of temporal duration of electric shock: an exploratory study. *Perceptual and Motor Skills, 42*, 1174.
- Barrett, L. F., & Pietromonaco, P. R. (1997). Accuracy of the five-factor model in predicting perceptions of daily social interactions. *Personality and Social Psychology Bulletin, 23*(11), 1173-1187. doi: 10.1177/01461672972311005
- Baydar, N. (1995). Reliability and validity of temperament scales of the NLSY child assessments. *Journal of Applied Developmental Psychology, 16*(3), 339-370. doi: [10.1016/0193-3973\(95\)90024-1](https://doi.org/10.1016/0193-3973(95)90024-1)

## References

- Bekker, B. N. (2014). *Personality and politics*. (Unpublished doctoral dissertation). University of Southern Denmark
- Benedan, L., Zajac, R., McFarlane, F., & Powell, M. B. (2020). Attentional difficulty is a risk factor for interrogative suggestibility in preschoolers. *Children and Youth Services Review, 119*, 105487. doi: 10.1016/j.chilyouth.2020.105487
- Berk, L. E. (2012). *Child development* (9<sup>th</sup> edition). Pearson.
- Bettenay, C., Ridley, A., Henry, L., & Crane, L. (2015). Changed Responses Under Cross-examination: The Role of Anxiety and Individual Differences in Child Witnesses. *Applied Cognitive Psychology, 29*(3), 485-491. doi: 10.1002/acp.3125
- Bidrose, S. & Goodman, G. S. (2000). Testimony and evidence: A scientific case study of memory for child sexual abuse. *Applied Cognitive Psychology, 14*, 197-213. doi: 10.1002/(SICI)1099-0720(200005/06)14:3<197::AID-ACP647>3.0.CO;2-6
- Bienvenu, O. J., Nestadt, G., Samuels, J. F., Costa, P. T., Howard, W. T., & Eaton, W. W. (2001). Phobic, panic, and major depressive disorders and the five-factor model of personality. *The Journal of Nervous and Mental Disease, 189*(3), 154-161.
- Birbaumer, N., Grodd, W., Diedrich, O., Klose, U., Erb, M., Lotze, M., ... & Flor, H. (1998). fMRI reveals amygdala activation to human faces in social phobics. *Neuroreport, 9*(6), 1223-1226. doi: 10.1097/00001756-199804200-00048
- Blackford, C. (2000). *The effects of age, delayed interval, and type of question on false memory syndrome in childhood*. (Unpublished master's thesis). Emporia State University.

## References

- Brodsky, S. L., Griffin, M. P., & Cramer, R. J. (2010). The Witness Credibility Scale: An outcome measure for expert witness research. *Behavioral Sciences & the Law, 28*(6), 892-907. doi: 10.1002/bsl.917
- Boer, F., & Westenberg, P. M. (1994). The factor structure of the Buss and Plomin EAS Temperament Survey (parental ratings) in a Dutch sample of elementary school children. *Journal of Personality Assessment, 62*(3), 537-551. doi: [10.1207/s15327752jpa6203\\_13](https://doi.org/10.1207/s15327752jpa6203_13)
- Boeree, C.G. (2006). *Personality theories*. Retrieved February 24<sup>th</sup>, 2021 at <http://webpace.ship.edu/cgboer/persintro>.
- Booth-Kewley, S., & Vickers Jr, R. R. (1994). Associations between major domains of personality and health behavior. *Journal of Personality, 62*(3), 281-298. doi: 10.1111/j.1467-6494.1994.tb00298.x
- Borkenau, P., Riemann, R., Angleitner, A., & Spinath, F. M. (2001). Genetic and environmental influences on observed personality: Evidence from the German Observational Study of Adult Twins. *Journal of Personality and Social Psychology, 80*(4), 655-668. doi: 10.1037/0022-3514.80.4.655
- Bottoms, B. L., Golding, J. M., Stevenson, M. C., Wiley, T. R. A., & Yozwiak, J. A. (2007). A review of factors affecting jurors' decisions in child sexual abuse cases. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *The handbook of eyewitness psychology, Vol. 1. Memory for events* (p. 509–543). Lawrence Erlbaum Associates Publishers.
- Bottoms, B., & Goodman, G. (1994). Perceptions of Children's Credibility in Sexual Assault Cases. *Journal of Applied Social Psychology, 24*(8), 702-732. doi: 10.1111/j.1559-1816.1994.tb00608.x

## References

- Bouchard, T. J. (1994). Genes, environment, and personality. *Science-AAAS-Weekly Paper Edition*, 264(5166), 1700-1701. doi: 10.1126/science.8209250
- Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, 31(3), 243-273. doi: 10.1023/A:1012294324713
- Bowcott, O. (2017, October 10). Two-year-old girl gives evidence in UK abuse case. Retrieved March 23, 2020, from <https://www.theguardian.com/law/2017/oct/10/two-year-old-girl-gives-evidence-in-uk-abuse-case>
- Brackmann, N., Otgaar, H., Sauerland, M., & Jelicic, M. (2015). When Children are the Least Vulnerable to False Memories: A True Report or a Case of Autosuggestion? *Journal of Forensic Sciences*, 61, S271-S275. doi: 10.1111/1556-4029.12926
- Brackmann, N., Otgaar, H., Sauerland, M. and Merckelbach, H., 2015. *Children are poor witnesses. Or are they?* The Inquisitive Mind. Retrieved February 28, 2020, from <https://www.in-mind.org/article/children-are-poor-witnesses-or-are-they>
- Van Brakel, A. M., & Muris, P. (2006). A brief scale for measuring “behavioral inhibition to the unfamiliar” in children. *Journal of Psychopathology and Behavioral Assessment*, 28(2), 79-84. doi: 10.1007/s10862-006-7485-7
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford University Press.
- Brainerd, C. J., Reyna, V. F., & Ceci, S. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin*, 134(3), 343-382. doi: 10.1037/0033-2909.134.3.343
- Brainerd, C. J., Reyna, V. F., & Zember, E. (2011). Theoretical and forensic implications of developmental studies of the DRM illusion. *Memory & Cognition*, 39(3), 365-380. doi: 10.3758/s13421-010-0043-2

## References

- Braungart, J. M., Fulker, D. W., & Plomin, R. (1992). Genetic mediation of the home environment during infancy: A sibling adoption study of the HOME. *Developmental Psychology, 28*(6), 1048-1055. doi: 10.1037/0012-1649.28.6.1048
- Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26*(3), 353-364. doi: 10.1023/A:1015380522722
- Brewer, N., & Day, K. (2005). The confidence-accuracy and decision latency-accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology and Law, 12*(1), 119-128. doi: 10.1375/pplt.2005.12.1.119
- Brewer, K. D., Rowe, D. M., & Brewer, D. D. (1997). Factors related to prosecution of child sexual abuse cases. *Journal of Child Sexual Abuse: Research, Treatment, & Program Innovations for Victims, Survivors, & Offenders, 6*, 91-111. doi: 10.1300/j070v06n01\_07
- Briggs, R. S. (1988). Shyness: Introversion or neuroticism? *Journal of Research in Personality, 22*, 290-307. doi: 10.1016/0092-6566(88)90031-1
- Brigham, J. C., & Wolfskeil, M. P. (1983). Opinions of attorneys and law enforcement personnel on the accuracy of eyewitness identifications. *Law and Human Behavior, 7*(4), 337-349. doi: 10.1007/BF01044736
- Bright-Paul, A., Jarrold, C., & Wright, D. (2008). Theory-of-mind development influences suggestibility and source monitoring. *Developmental Psychology, 44*(4), 1055-1068. doi: 10.1037/0012-1649.44.4.1055

## References

- Brose, L. A., Rye, M. S., Lutz-Zois, C., & Ross, S. R. (2005). Forgiveness and personality traits. *Personality and Individual Differences, 39*(1), 35-46. doi: 10.1016/j.paid.2004.11.001
- Brown, G. L., Mangelsdorf, S. C., Agathen, J. M., & Ho, M. H. (2008). Young children's psychological selves: Convergence with maternal reports of child personality. *Social Development, 17*(1), 161-182. doi: [10.1111/j.1467-9507.2007.00421.x](https://doi.org/10.1111/j.1467-9507.2007.00421.x)
- Bruck, M., Ceci, S. J., & Francoeur, E. (2000). Children's use of anatomically detailed dolls to report genital touching in a medical examination: Developmental and gender comparisons. *Journal of Experimental Psychology: Applied, 6*(1), 74-83. doi: 10.1037/1076-898X.6.1.74
- Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18*(8), 947-996. doi: [10.1002/acp.1070](https://doi.org/10.1002/acp.1070)
- Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Science, 12*(1), 1-11.
- Bull, R., & Corran, E. (2002). Interviewing child witnesses: Past and future. *International Journal of Police Science & Management, 4*(4), 315-322. doi: 10.1350/ijps.4.4.315.10872
- Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B.N. & Ornstein, P.A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 15*(7), S25-S48. doi: 10.1002/acp.833

## References

- Burton, C. M., Plaks, J. E., & Peterson, J. B. (2015). Why do conservatives report being happier than liberals? The contribution of neuroticism. doi: 10.23668/psycharchives.1706
- Burt, H., & Gaskill, H. (1932). Suggestibility and the form of the question. *Journal of Applied Psychology*, 16(4), 358-373. doi: 10.1037/h0071438
- Buss, A. H. (1989). Temperaments as personality traits. In G. A. Kohnstamm, J. E. Bates, & M. K. Rothbart (Eds), *Temperament in childhood* (pp. 49 – 58). Wiley.
- Buss, A., & Plomin, R. (1975). *A temperament theory of personality development*. John Wiley.
- Buss, A.H., & Plomin, R. (1984). *Temperament: Early developing personality traits*. Erlbaum.
- Buss, A. H., & Plomin, R. (2014). *Temperament (PLE: Emotion): Early developing personality traits*. Psychology Press.
- Campbell, R., Walker, J., & Baron-Cohen, S. (1995). The development of differential use of inner and outer face features in familiar face identification. *Journal of Experimental Child Psychology*, 59(2), 196-210. doi: 10.1006/jecp.1995.1009
- Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion and prosocial value motivation. *Personality and individual differences*, 38(6), 1293-1305. doi: 10.1016/j.paid.2004.08.012
- Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, 29(6), 807-840. doi: 10.1111/j.1467-9221.2008.00668.x

## References

- Carson, S. H., Peterson, J. B., & Higgins, D. M. (2003). Decreased Latent Inhibition Is Associated With Increased Creative Achievement in High-Functioning Individuals. *Journal of Personality and Social Psychology, 85*(3), 499–506. doi: [10.1037/0022-3514.85.3.499](https://doi.org/10.1037/0022-3514.85.3.499)
- Caspi, A. (1998). Personality development across the life course. In W. Damon & N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development* (p. 311–388). John Wiley & Sons, Inc.
- Caspi, A., Harrington, H., Milne, B., Amell, J. W., Theodore, R. F., & Moffitt, T. E. (2003). Children's behavioral styles at age 3 are linked to their adult personality traits at age 26. *Journal of Personality, 71*(4), 495-514. doi: 10.1111/1467-6494.7104001
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453-484. doi: 10.1146/annurev.psych.55.090902.141913
- Caspi, A., & Silva, P. A. (1995). Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development, 66*(2), 486-498. doi: 10.1111/j.1467-8624.1995.tb00885.x
- Cassel, W. S., Roebbers, C. E., & Bjorklund, D. F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology, 61*(2), 116-133. doi: [10.1006/jecp.1996.0008](https://doi.org/10.1006/jecp.1996.0008)
- Cassidy, G., & MacDonald, R. A. (2007). The effect of background music and background noise on the task performance of introverts and extraverts. *Psychology of Music, 35*(3), 517-537. doi: 10.1177/0305735607076444

## References

- Castelli, P., Goodman, G. S., & Ghetti, S. (2005). Effects of interview style and witness age on perceptions of children's credibility in sexual abuse cases. *Journal of Applied Social Psychology, 35*(2), 297-317. doi: 10.1111/j.1559-1816.2005.tb02122.x
- CBS Los Angeles. (2014, August 5). 30 Years Later, Key Figures Reflect On McMartin Child Abuse Case. Retrieved March 23, 2020, from <https://www.youtube.com/watch?v=WLS-WhiYrGw>.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*(3), 403-439. doi: 10.1037/0033-2909.113.3.403
- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association.
- Ceci, S. J., & de Bruyn, E. (1993). Child Witnesses in Court: A Growing Dilemma. *Children Today, 22*(1), 5-9.
- Ceci, S. J., & Hritz, A., & Royer, C. (2017). Understanding Suggestibility. O'Donohue, W. T., & Fanetti, M (Ed.), *Forensic Interviews Regarding Child Sexual Abuse* (pp. 141-154). Springer.
- Ceci, S. J., & Huffman, M. C. (1997). How suggestible are preschool children? Cognitive and social factors. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(7), 948-958. doi: 10.1097/00004583-199707000-00017
- Ceci, S. J., Huffman, M., Smith, E., & Loftus, E. (1994). Repeatedly Thinking about a Non-event: Source Misattributions among Preschoolers. *Consciousness And Cognition, 3*(3-4), 388-407. doi: 10.1006/ccog.1994.1022

## References

- Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General*, *116*(1), 38-49. doi: 10.1037/0096-3445.116.1.38
- Cederborg, A. C. (2004). Factors influencing child witnesses. *Scandinavian Journal of Psychology*, *45*(3), 197-205. doi: 10.1111/j.1467-9450.2004.00395.x
- Cederborg, A. C., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child abuse & neglect*, *24*(10), 1355-1361. doi: 10.1016/S0145-2134(00)00183-6
- Chae, Y. (2004). *Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions*. (Unpublished doctoral dissertation). Cornell University.
- Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *19*(4), 383-407. doi: [10.1002/acp.1094](https://doi.org/10.1002/acp.1094)
- Chae, Y., Goodman, M., Goodman, G. S., Troxel, N., McWilliams, K., Thompson, R. A., ... & Widaman, K. F. (2018). How children remember the Strange Situation: The role of attachment. *Journal of Experimental Child Psychology*, *166*, 360-379. doi: 10.1016/j.jecp.2017.09.001
- Chae, Y., Goodman, G. S., Larson, R. P., Augusti, E.-M., Alley, D., VanMeenen, K. M., ... Coulter, K. P. (2014). Children's memory and suggestibility about a distressing event: The role of children's and parents' attachment. *Journal of Experimental Child Psychology*, *123*, 90-111. doi: 10.1016/j.jecp.2014.01.005

## References

- Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Paediatric Psychology, 27*(1), 27-36. doi: [10.1093/jpepsy/27.1.27](https://doi.org/10.1093/jpepsy/27.1.27)
- Chamorro-Premuzic, T., & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality, 17*(3), 237-250 doi: 10.1002/per.473
- Chance, J., & Goldstein, A. (1984). Face-Recognition Memory: Implications for Children's Eyewitness Testimony. *Journal of Social Issues, 40*(2), 69-85. doi: 10.1111/j.1540-4560.1984.tb01094.x
- Carroll County Child Advocacy Center (2015). *Child Sexual Abuse Statistics*. Retrieved 24 March 2021, from [https://www.cc-cac.org/wp-content/uploads/2020/04/all\\_statistics\\_20150619.pdf](https://www.cc-cac.org/wp-content/uploads/2020/04/all_statistics_20150619.pdf)
- Chavira, D. A., Stein, M. B., & Malcarne, V. L. (2002). Scrutinizing the relationship between shyness and social phobia. *Journal Anxiety Disorder, 16*(6), 585-598. doi: 10.1016/S0887-6185(02)00124-X
- Cheek, J. M., Melchior, L. A., & Carpentieri, A. M. (1986). Shyness and self-concept. In L. M. Hattman & K. R. Blankstein (Eds.), *Perception of self in emotional disorder and psychotherapy* (pp. 113-131). Plenum Press.
- Cheit, R. E. (2014). *The witch-hunt narrative: Politics, psychology, and the sexual abuse of children*. Oxford University Press.
- Chen, C. F. (2002). *The effects of age, interview style, time delay, and temperament on children's suggestibility*. (Unpublished master's thesis). Emporia State University.
- Child Welfare Information Gateway. (2019.) *Child maltreatment 2017: Summary of key findings*. Department of Health and Human Services, Administration for Children and Families, Children's Bureau.

## References

- Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review*, *15*(6), 1089-1092. doi: 10.3758/PBR.15.6.1089
- Christov-Moore, L., Simpson, E. A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P. F. (2014). Empathy: gender effects in brain and behavior. *Neuroscience & Biobehavioral Reviews*, *46*, 604-627. doi: 10.1016/j.neubiorev.2014.09.001
- Clark, J., Boccaccini, M. T., Caillouet, B., & Chaplin, W. F. (2007). Five factor model personality traits, jury selection, and case outcomes in criminal and civil cases. *Criminal Justice and Behavior*, *34*(5), 641-660. doi: 10.1177/0093854806297555
- Clarke-Stewart, K., Malloy, L., & Allhusen, V. (2004). Verbal ability, self-control, and close relationships with parents protect children against misleading suggestions. *Applied Cognitive Psychology*, *18*(8), 1037-1058. doi: 10.1002/acp.1076
- Collins, K., Harker, N., & Antonopoulos, G. A. (2016). The impact of the registered intermediary on adults' perceptions of child witnesses: Evidence from a mock cross examination. *European Journal on Criminal Policy and Research*, *23*(2), 211-225. doi: 10.1007/s10610-016-9314-1
- Coplan, R. J., Wilson, J., Frohlick, S. L., & Zelenski, J. (2006). A person-oriented analysis of behavioral inhibition and behavioral activation in children. *Personality and Individual Differences*, *41*(5), 917-927. doi: 10.1016/j.paid.2006.02.019
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*(1), 110-117. doi: 10.1111/j.1468-2389.2007.00371.x

## References

- Connor-Smith, J. K., & Flachsbart, C. (2007). Relations between personality and coping: A meta-analysis. *Journal of Personality and Social Psychology, 93*(6), 1080-1107. doi: [10.1037/0022-3514.93.6.1080](https://doi.org/10.1037/0022-3514.93.6.1080)
- Corr, P. J., Pickering, A. D., & Gray, J. A. (1995). Sociability/impulsivity and caffeine-induced arousal: Critical flicker/fusion frequency and procedural learning. *Personality and Individual Differences, 18*(6), 713-730. doi: 10.1016/0191-8869(95)00001-M
- Costa, P. T., & McCrae, R. R. (1986). Cross-sectional studies of personality in a national sample: I. Development and validation of survey measures. *Psychology and Aging, 1*(2), 140-143. doi: 10.1037/0882-7974.1.2.140
- Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: a six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of personality and social psychology, 54*(5), 853-863. doi: 10.1037/0022-3514.54.5.853
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64*(1), 21–50. doi: 10.1207/s15327752jpa6401\_2
- Costa Jr, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology, 81*(2), 322. doi: 10.1037/0022-3514.81.2.322
- Cotterill, B. F. (2017). *How leading questions and shyness affect children as eyewitnesses*. (Unpublished master's thesis). University of York.

## References

- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1), 87-114. doi: 10.1017/S0140525X01003922
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, 169, 323-338.
- Cramer, R. J., Brodsky, S. L., & DeCoster, J. (2009). Expert witness confidence and juror personality: Their impact on credibility and persuasion in the courtroom. *Journal of the American Academy of Psychiatry and the Law Online*, 37(1), 63-74.
- Crawford, A. (2009, November 24). NSPCC urges more support for child court witnesses. Retrieved March 23, 2020, from <http://news.bbc.co.uk/1/hi/uk/8375293.stm>.
- Cronch, L. E., Viljoen, J. L., & Hansen, D. J. (2006). Forensic interviewing in child sexual abuse cases: Current techniques and future directions. *Aggression and Violent Behavior*, 11(3), 195-207. doi: 10.1016/j.avb.2005.07.009
- Crooks, C. V., & Peters, R. D. (2005). Predicting academic difficulties: Does a complex, multidimensional model outperform a unidimensional teacher rating scale?. *Canadian Journal of Behavioural Science*, 37(3), 170-180. doi: 10.1037/h0087254
- Cross, T. P., De Vos, E., & Whitcomb, D. (1994). Prosecution of child sexual abuse: Which cases are accepted? *Child Abuse & Neglect*, 18, 661-677. doi: 10.1016/0145-2134(94)90016-7
- Crossman, A. (2001). *Predicting suggestibility: The role of individual differences and socialization*. (Unpublished doctoral dissertation). Cornell University.
- Crown Office and Procurator Fiscal Service. (2018). Freedom of information request.

## References

- Crown Prosecution Service. (2017, October). *CPS report shows more offenders are being successfully prosecuted for sexual crimes than ever before*. Retrieved 30 December 2020, from <https://www.cps.gov.uk/cps/news/cps-report-shows-more-offenders-are-being-successfully-prosecuted-sexual-crimes-ever>
- Crozier, W. (1981). Shyness and self-esteem. *British Journal Of Social Psychology*, 20(3), 220-222. doi: 10.1111/j.2044-8309.1981.tb00537.x
- Crozier, W. R. (1995). Shyness and self-esteem in middle childhood. *British Journal of Educational Psychology*, 65(1), 85-95. [doi: 10.1111/j.2044-8279.1995.tb01133.x](https://doi.org/10.1111/j.2044-8279.1995.tb01133.x)
- Crozier, W. R. (2000). Shyness and social relationships: Continuity and change. In W. R. Crozier (Ed.), *Shyness: Development, consolidation, and change* (pp. 1–21). New York: Routledge.
- Curley, L. J., MacLean, R., & Murray, J. (2017). The Relationship between the Big 5 Personality Traits and Eyewitness Recognition. *Journal of Articles in Support of the Null Hypothesis*, 13(2).
- Cutler, B. L., Moran, G., & Narby, D. J. (1992). Jury selection in insanity defense cases. *Journal of Research in Personality*, 26(2), 165-182. doi: 10.1016/0092-6566(92)90052-6
- Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, 14(2), 185-191. doi: 10.1007/BF01062972
- Danielson, C. K., & Phelps, C. R. (2003). The assessment of children's social skills through self-report: A potential screening instrument for classroom use. *Measurement and*

## References

- Evaluation in Counselling and Development*, 35(4), 218-229. doi: [10.1080/07481756.2003.12069068](https://doi.org/10.1080/07481756.2003.12069068)
- Danielsdottir, G., Sigurgeirsdottir, S., Einarsdottir, H., & Haraldsson, E. (1993). Interrogative suggestibility in children and its relationship with memory and vocabulary. *Personality and Individual Differences*, 14(3), 499-502. doi: 10.1016/0191-8869(93)90323-u
- Davies, G. M. (1996). Children's identification evidence. *Psychological issues in eyewitness identification*, 233-258. doi:
- Davies, G. and Noon, E. (1991) *Evaluation of the live link for child witnesses*. London: HMSO.
- Davies, G. M., Smith, S., & Blincoe, C. (2008). A "weapon focus" effect in children. *Psychology, crime & law*, 14(1), 19-28. doi: 10.1080/10683160701340593
- Davies, G., Tarrant, A., & Flin, R. (1989). Close encounters of the witness kind: Children's memory for a simulated health inspection. *British Journal of Psychology*, 80(4), 415-429. doi: 10.1111/j.2044-8295.1989.tb02333.x
- Davis, H., & Carr, M. (2002). Gender differences in mathematics strategy use: The influence of temperament. *Learning and Individual Differences*, 13(1), 83-95. doi: 10.1016/S1041-6080(02)00063-8
- Davis, D., Loftus, E. F. (2005). Age and functioning in the legal system: Perception memory and judgment in victims, witnesses and jurors. In Noy, I., Karwowski, W. (Eds.), *Handbook of forensic human factors and ergonomics*. Taylor & Francis.
- Davis, J. M., McKone, E., Dennett, H., O'Connor, K. B., O'Kearney, R., & Palermo, R. (2011). Individual differences in the ability to recognise facial identity are associated with social anxiety. *PloS one*, 6(12), e28800. doi: 10.1371/journal.pone.0028800

## References

- Deal, J. E., Halverson Jr, C. F., Havill, V., & Martin, R. P. (2005). Temperament factors as longitudinal predictors of young adult personality. *Merrill-Palmer Quarterly*, *51*, 315-334.
- De Fruyt, Filip, Meike Bartels, Karla G. Van Leeuwen, Barbara De Clercq, Mieke Decuyper, and Ivan Mervielde. Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, *91*(3), 538-552.  
10.1521/pedi.19.2.171.62627
- Dekle, D. J., Beal, C. R., Elliott, R., & Huneycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology*, *10*(1), 1-12. doi: 10.1002/(SICI)1099-0720(199602)10:1<1::AID-ACP354>3.0.CO;2-Y
- DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin*, *124*(2), 197-229. doi: [10.1037/0033-2909.124.2.197](https://doi.org/10.1037/0033-2909.124.2.197)
- Denissen, J. J., & Penke, L. (2008). Neuroticism predicts reactions to cues of social inclusion. *European Journal of Personality: Published for the European Association of Personality Psychology*, *22*(6), 497-517. doi: 10.1002/per.682
- De Pauw, S. S. (2017). Childhood personality and temperament. *The Oxford handbook of the five factor model*, 243-280.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*(3), 491-517. doi: 10.1017/S0140525X99372046

## References

- de Villiers, J., & de Villiers, P. (2014). The Role of Language in Theory of Mind Development. *Topics In Language Disorders*, 34(4), 313-328. doi: 10.1097/tld.0000000000000037
- Devine, D. J., & Caughlin, D. E. (2014). Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psychology, Public Policy, and Law*, 20(2), 109-134. doi: 10.1037/law0000006
- Devine, D. J., Clayton, L. D., Dunford, B. B., Seying, R., & Pryce, J. (2001). Jury decision making: 45 years of empirical research on deliberating groups. *Psychology, Public Policy, and Law*, 7(3), 622–727. doi: [10.1037/1076-8971.7.3.622](https://doi.org/10.1037/1076-8971.7.3.622)
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, 91(6), 1138–1151. doi: [10.1037/0022-3514.91.6.1138](https://doi.org/10.1037/0022-3514.91.6.1138)
- DeYoung, C. G. (2010). Personality neuroscience and the biology of traits. *Social and Personality Psychology Compass*, 4(12), 1165-1180. doi: 10.1111/j.1751-9004.2010.00327.x
- DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the Personality Inventory for DSM–5. *Personality Disorders: Theory, Research, and Treatment*, 7(2), 113–123. doi: [10.1037/per0000170](https://doi.org/10.1037/per0000170)
- DeYoung, C. G., & Gray, J. R. (2009). Personality neuroscience: Explaining individual differences in affect, behaviour and cognition. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (p. 323–346). Cambridge University Press. doi: [10.1017/CBO9780511596544.023](https://doi.org/10.1017/CBO9780511596544.023)

## References

- DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality, 46*(1), 63-78. doi: 10.1016/j.jrp.2011.12.003
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880-896. doi: 10.1037/0022-3514.93.5.880
- DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to experience, intellect, and cognitive ability. *Journal of Personality Assessment, 96*(1), 46-52. doi: 10.1080/00223891.2013.806327
- DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the aspects of the Big Five, the interpersonal circumplex, and trait affiliation. *Journal of personality, 81*(5), 465-475. doi: 10.1111/jopy.12020
- Di Blas, L. (2005). Personality-relevant attribute-nouns: a taxonomic study in the Italian language. *European Journal of Personality: Published for the European Association of Personality Psychology, 19*(7), 537-557. doi: 10.1002/per.569
- DiCerbo, K. E. (2014). Game-based assessment of persistence. *Journal of Educational Technology & Society, 17*(1), 17-28.
- DiPietro, J. A., Hodgson, D. M., Costigan, K. A., & Johnson, T. R. (1996). Fetal antecedents of infant temperament. *Child development, 67*(5), 2568-2583. doi: 10.1111/j.1467-8624.1996.tb01874.x
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*(6), 1246–1256. doi: [10.1037/0022-3514.73.6.1246](https://doi.org/10.1037/0022-3514.73.6.1246)

## References

- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, *50*(1), 116-123. doi: [10.1037/0022-3514.50.1.116](https://doi.org/10.1037/0022-3514.50.1.116)
- Digman, J. M., & Shmelyov, A. G. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology*, *71*(2), 341-351. doi: [10.1037/0022-3514.71.2.341](https://doi.org/10.1037/0022-3514.71.2.341)
- Dobbs, S., Furnham, A., & McClelland, A. (2011). The effect of background music and noise on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, *25*(2), 307-313. doi: 10.1002/acp.1692
- Dodd, D. H., & Bradshaw, J. M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior*, *19*(6), 695-704. doi: [10.1016/S0022-5371\(80\)90379-5](https://doi.org/10.1016/S0022-5371(80)90379-5)
- Doinita, N. E., & Maria, N. D. (2015). Attachment and parenting styles. *Procedia-Social and Behavioral Sciences*, *203*, 199-204. doi: 10.1016/j.sbspro.2015.08.282
- Downer, J. T., & Mendez, J. L. (2005). African American father involvement and preschool children's school readiness. *Early Education and Development*, *16*(3), 317-340. doi: 10.1207/s15566935eed1603\_2
- Drevets, W. C. (1999). Prefrontal cortical-amygdalar metabolism in major depression. *Annals of the New York Academy of Sciences*, *877*(1), 614-637. doi: 10.1111/j.1749-6632.1999.tb09292.x
- D'Souza, L., Gowda, H. M. R., & Gowda, D. K. S. (2006). Shyness and fear among high school students. *Pakistan Journal of Psychological Research*, *21* (32), 53-60.

## References

- Duffy, K. A., Helzer, E. G., Hoyle, R. H., Fukukura Helzer, J., & Chartrand, T. L. (2018). Pessimistic expectations and poorer experiences: The role of (low) extraversion in anticipated and experienced enjoyment of social interaction. *PloS one*, *13*(7), e0199146. doi: 10.1371/journal.pone.0199146
- Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: A 10- to 12-second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology*, *87*(5), 951–962. doi: [10.1037/0021-9010.87.5.951](https://doi.org/10.1037/0021-9010.87.5.951)
- Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful noncognitive predictor of school achievement in adolescents. *Journal of Individual Differences*, *37*(1), 8–15. doi: [10.1027/1614-0001/a000182](https://doi.org/10.1027/1614-0001/a000182)
- Dupree, L., Patterson, D. A., Nugent, W. R., & White, S. F. (2016). How children verbally reported their experiences of a genital exam. *Journal of Human Behavior in the Social Environment*, *26*(1), 110-118. doi: 10.1080/10911359.2015.1062675
- Durbin, C. E., Hayden, E. P., Klein, D. N., & Olino, T. M. (2007). Stability of laboratory-assessed temperamental emotionality traits from ages 3 to 7. *Emotion*, *7*(2), 388-399. doi: [10.1037/1528-3542.7.2.388](https://doi.org/10.1037/1528-3542.7.2.388)
- Earls, F., & Jung, K. G. (1987). Temperament and home environment characteristics as causal factors in the early development of childhood psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *26*(4), 491-498. doi: 10.1097/00004583-198707000-00005
- Eaton, W. O. (1983). Measuring activity level with actometers: Reliability, validity, and arm length. *Child Development*, *54*(3), 720-726. doi: 10.2307/1130060

## References

- Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development, 265-275*. doi: 10.2307/1130193
- Edelstein, L. (1998). An Accusation Easily to be Made? Rape and Malicious Prosecution in Eighteenth-Century England. *The American Journal Of Legal History, 42(4)*, 351. doi: 10.2307/846040
- Eder, R. A. (1990). Uncovering young children's psychological selves: Individual and developmental differences. *Child Development, 61(3)*, 849-863. doi: [10.2307/1130969](https://doi.org/10.2307/1130969)
- Eisen, M. L., Goodman, G. S., Qin, J., Davis, S., & Crayton, J. (2007). Maltreated children's memory: Accuracy, suggestibility, and psychopathology. *Developmental Psychology, 43(6)*, 1275-1294. doi: 10.1037/0012-1649.43.6.1275
- Eisen, M. L., Qin, J., Goodman, G. S., & Davis, S. L. (2002). Memory and suggestibility in maltreated children: Age, stress arousal, dissociation, and psychopathology. *Journal of Experimental Child Psychology, 83(3)*, 167-212. doi: 10.1016/S0022-0965(02)00126-1
- Eisenberg, N., Duckworth, A. L., Spinrad, T. L., & Valiente, C. (2014). Conscientiousness: Origins in childhood? *Developmental Psychology, 50(5)*, 1331-1349. doi: [10.1037/a0030977](https://doi.org/10.1037/a0030977)
- Eisenberg, N., Guthrie, I. K., Fabes, R. A., Reiser, M., Murphy, B. C., Holgren, R., Maszk, P., & Losoya, S. (1997). The relations of regulation and emotionality to resiliency and competent social functioning in elementary school children. *Child Development, 68(2)*, 295-311. doi: [10.2307/1131851](https://doi.org/10.2307/1131851)
- Eisenberg, N., Shepard, S. A., Fabes, R. A., Murphy, B. C., & Guthrie, I. K. (1998). Shyness and children's emotionality, regulation, and coping: Contemporaneous, longitudinal, and

## References

- across-context relations. *Child Development*, 69(3), 767-790. doi: [10.1111/j.1467-8624.1998.00767.x](https://doi.org/10.1111/j.1467-8624.1998.00767.x)
- Eisenberg, N., Valiente, C., Morris, A. S., Fabes, R. A., Cumberland, A., Reiser, M., Gershoff, E. T., Shepard, S. A., & Losoya, S. (2003). Longitudinal relations among parental emotional expressivity, children's regulation, and quality of socioemotional functioning. *Developmental Psychology*, 39(1), 3–19. doi: [10.1037/0012-1649.39.1.3](https://doi.org/10.1037/0012-1649.39.1.3)
- Elschberger, H. (2005). The effects of prior knowledge on children's memory and suggestibility. *Journal Of Experimental Child Psychology*, 92(3), 247-275. doi: [10.1016/j.jecp.2005.05.002](https://doi.org/10.1016/j.jecp.2005.05.002)
- Ellis, L. K., & Rothbart, M. K. (2001, April). *Revision of the early adolescent temperament questionnaire*. In Poster presented at the 2001 biennial meeting of the society for research in child development, Minneapolis, MN. doi: [10.1037/t07624-000](https://doi.org/10.1037/t07624-000)
- Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin*, 132(1), 33-72. doi: [10.1037/0033-2909.132.1.33](https://doi.org/10.1037/0033-2909.132.1.33)
- Endres, J., Poggenpohl, C., & Erben, C. (1999). Repetitions, warnings and video: Cognitive and motivational components in preschool children's suggestibility. *Legal and Criminological Psychology*, 4(1), 129-146. doi.org/[10.1348/135532599167725](https://doi.org/10.1348/135532599167725)
- Ernberg, E., Tidefors, I., & Landström, S. (2016). Prosecutors' reflections on sexually abused preschoolers and their ability to stand trial. *Child Abuse & Neglect*, 57, 21-29. doi: [10.1016/j.chiabu.2016.06.001](https://doi.org/10.1016/j.chiabu.2016.06.001)
- Evans, A. D., & Lee, K. (2013). Emergence of lying in very young children. *Developmental psychology*, 49(10), 1958-1963. doi: [10.1037/a0031409](https://doi.org/10.1037/a0031409)

## References

- Evans, A. D., & Lyon, T. D. (2020). Evasive Behaviors in 9-12-Year-Olds Concealment of a Minor Transgression, American-Psychology Law-Society, New Orleans, LA, 2020, March.
- Everson, M. D., & Boat, B. W. (1989). False allegations of sexual abuse by children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 28(2), 230-235. doi: 10.1300/j070v06n01\_07
- Everson, M. D., & Boat, B. W. (1994). Putting the anatomical doll controversy in perspective: An examination of the major uses and criticisms of the dolls in child sexual abuse evaluations. *Child Abuse & Neglect*, 18(2), 113-129. doi: 10.1016/0145-2134(94)90114-7
- Everson, M. D., & Boat, B. W. (1997). Anatomical dolls in child sexual abuse assessments: A call for forensically relevant research. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 11(7), S55-S74. doi: 10.1002/(SICI)1099-0720(199712)11:7<S55::AID-ACP521>3.0.CO;2-P
- Ewing, A. R., & Taylor, A. R. (2009). The role of child gender and ethnicity in teacher-child relationship quality and children's behavioral adjustment in preschool. *Early Childhood Research Quarterly*, 24(1), 92-105. doi: 10.1016/j.ecresq.2008.09.002
- Fabes, R. A., Shepard, S. A., Guthrie, I. K., & Martin, C. L. (1997). Roles of temperamental arousal and gender-segregated play in young children's social adjustment. *Developmental Psychology*, 33(4), 693-702. doi: [10.1037/0012-1649.33.4.693](https://doi.org/10.1037/0012-1649.33.4.693)

## References

- Faller, K. C. (2005). Anatomical dolls: Their use in assessment of children who may have been sexually abused. *Journal of Child Sexual Abuse, 14*(3), 1-21. doi: 10.1300/J070v14n03\_01
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. doi: 10.3758/BRM.41.4.1149
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*(3), 429-456. doi: [10.1037/0033-2909.116.3.429](https://doi.org/10.1037/0033-2909.116.3.429)
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review, 2*(4), 290-309. doi: 10.1207/s15327957pspr0204\_5
- Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality, 37*(2), 62-88. doi: 10.1016/S0092-6566(02)00536-6
- Feltis, B. B., Powell, M. B., Snow, P. C., & Hughes-Scholes, C. H. (2010). An examination of the association between interviewer question type and story-grammar detail in child witness interviews about abuse. *Child Abuse & Neglect, 34*(6), 407-413. doi: 10.1016/j.chiabu.2009.09.019
- Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50*(3), 274-290.
- Florida, R. (2008, May 4). Where do all the neurotics live? Boston Globe. Retrieved January 2, 2021, from

## References

[http://www.boston.com/bostonglobe/ideas/articles/2008/05/04/where\\_do\\_all\\_the\\_neurotics\\_live/](http://www.boston.com/bostonglobe/ideas/articles/2008/05/04/where_do_all_the_neurotics_live/)

Fitzgerald, R., & Price, H. (2015). Eyewitness identification across the life span: A meta-analysis of age differences. *Psychological Bulletin*, *141*(6), 1228-1265. doi: 10.1037/bul0000013

Fivush, R. (1997). Event memory in early childhood. In N. Cowan (Ed.), *The development of memory in childhood* (p. 139–161). Psychology Press/Erlbaum (UK) Taylor & Francis.

Fivush, R. (2011). The Development of Autobiographical Memory. *Annual Review Of Psychology*, *62*(1), 559-582. doi: 10.1146/annurev.psych.121208.131702

Fivush, R., McDermott Sales, J., Goldberg, A., Bahrick, L., & Parker, J. (2004). Weathering the storm: Children's long-term recall of Hurricane Andrew. *Memory*, *12*(1), 104-118. doi: 10.1080/09658210244000397

Foley, M. (2014). Children's memory for source. In P. J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (Vol. 1, pp. 427– 452). Wiley Blackwell.

Foley, L. A., & Chamblin, M. H. (1982). The effect of race and personality on mock jurors' decisions. *The Journal of Psychology*, *112*(1), 47-51. doi: 10.1080/00223980.1982.9923533

Folke, C., Carpenter, S. R., Walker, B., Scheffer, M., Chapin, T., & Rockström, J. (2010). Resilience thinking: integrating resilience, adaptability and transformability. *Ecology and society*, *15*(4).

[doi: 10.5751/ES-03610-150420](https://doi.org/10.5751/ES-03610-150420)

## References

- Forster, S., & Lavie, N. (2016). Establishing the attention-distractibility trait. *Psychological Science, 27*(2), 203-212. doi: 10.1177/0956797615617761
- Friedman, W., & Lyon, T. (2005). Development of Temporal-Reconstructive Abilities. *Child Development, 76*(6), 1202-1216. doi: 10.1111/j.1467-8624.2005.00844.x-i1
- Friedman, H. S., Tucker, J. S., Schwartz, J. E., Martin, L. R., Tomlinson-Keasey, C., Wingard, D. L., & Criqui, M. H. (1995). Childhood conscientiousness and longevity: health behaviors and cause of death. *Journal of Personality and Social Psychology, 68*(4), 696-703. doi: 0.1037//0022-3514.68.4.696
- Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Schwartz, J. E., Wingard, D. L., & Criqui, M. H. (1993). Does childhood personality predict longevity?. *Journal of Personality and Social Psychology, 65*(1), 176-185. doi: 10.1037//0022-3514.65.1.176
- Funder, D. C. (2019). *The personality puzzle* (8<sup>th</sup> edition). W. W. Norton & Company.
- Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 11*(5), 445-455. doi: 10.1002/(SICI)1099-0720(199710)11:5<445::AID-ACP472>3.0.CO;2-R
- Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics, 45*(3), 203-217. doi: 10.1080/00140130210121932
- Galton, N., Hammond, N. G., & Stinchcombe, A. (2020). Personality traits and fears of death and dying predict ageism. *Death Studies, 1-7*. doi: 10.1080/07481187.2020.1829746

## References

- Garstein, M. A., Slobodskaya, H. R., Żylicz, P. O., Foztyła, D., & Nakagawa, A. (2010). A cross-cultural evaluation of temperament: Japan, USA, Poland and Russia. *International Journal of Psychology and Psychological Therapy*, 10(1), 55-75.
- Garven, S., Wood, J. M., Malpass, R. S., & Shaw III, J. S. (1998). More than suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, 83(3), 347-359. doi: 10.1037/0021-9010.83.3.347
- Gasman, I., Purper-Ouakil, D., Michel, G., Mouren-Siméoni, M. C., Bouvard, M., Perez-Diaz, F., & Jouvent, R. (2002). Cross-cultural assessment of childhood temperament. *European Child & Adolescent Psychiatry*, 11(3), 101-107. doi: [10.1007/s00787-002-0248-4](https://doi.org/10.1007/s00787-002-0248-4)
- Geddie, L., Fradin, S., & Beer, J. (2000). Child characteristics which impact accuracy of recall and suggestibility in preschoolers: is age the best predictor? *Child Abuse and Neglect*, 24, 223-235. [doi.org/10.1016/S0145-2134\(99\)00133-7](https://doi.org/10.1016/S0145-2134(99)00133-7)
- Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2011). The big five personality traits in the political arena. *Annual Review of Political Science*, 14, 265-287. doi: 10.1146/annurev-polisci-051010-111659
- Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., & Ha, S. E. (2010). Personality and political attitudes: Relationships across issue domains and political contexts. *American Political Science Review*, 24(1), 111-133.
- Ghetti, S., & Angelini, L. (2008). The Development of Recollection and Familiarity in Childhood and Adolescence: Evidence From the Dual-Process Signal Detection Model. *Child Development*, 79(2), 339-358. doi: 10.1111/j.1467-8624.2007.01129.x

## References

- Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38*(5), 705-718. doi: 10.1037/0012-1649.38.5.705
- Gignac, G., & Powell, M. (2006). A direct examination of the nonlinear (quadratic) association between intelligence and suggestibility in children. *Applied Cognitive Psychology, 20*(5), 617-623. doi: 10.1002/acp.1213
- Gilstrap, L. L., & Papierno, P. B. (2004). Is the cart pushing the horse? The effects of child characteristics on children's and adults' interview behaviours. *Applied Cognitive Psychology, 18*(8), 1059-1078. doi: [10.1002/acp.1072](https://doi.org/10.1002/acp.1072)
- Glenn, A. L., & Raine, A. (2008). The neurobiology of psychopathy. *Psychiatric Clinics of North America, 31*(3), 463-475. doi: 10.1016/j.psc.2008.03.004
- Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216-1229. doi: [10.1037/0022-3514.59.6.1216](https://doi.org/10.1037/0022-3514.59.6.1216)
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42. doi: [10.1037/1040-3590.4.1.26](https://doi.org/10.1037/1040-3590.4.1.26)
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*(1), 26-34. doi: [10.1037/0003-066X.48.1.26](https://doi.org/10.1037/0003-066X.48.1.26)
- Golding, J. M., Bradshaw, G. S., Dunlap, E. E., & Hodell, E. C. (2007). The impact of mock jury gender composition on deliberations and conviction rates in a child sexual assault trial. *Child Maltreatment, 12*(2), 182-190. doi: 10.1177/1077559506298995

## References

- Golding, J. M., Fryman, H., Marsil, D., & Yozwiak, J. (2003). Big girls don't cry: the effect of child witness demeanor on juror decisions in a child sexual abuse trial. *Child Abuse & Neglect*, 27(11), 1311-1321. doi: 10.1016/j.chiabu.2003.03.001
- Golding, J. M., Sanchez, R. P., & Segó, S. A. (1997). The believability of hearsay testimony in a child sexual assault trial. *Law and Human Behavior*, 21(3), 299-325. doi: 10.1023/A:1024842816130
- Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., Hinde, R. A., & McCall, R. B. (1987). What is temperament? Four approaches. *Child Development*, 58(2), 505–529. [10.2307/1130527](https://doi.org/10.2307/1130527)
- Goldsmith, H. H., & Gagne, J. R. (2012). Behavioral assessment of temperament. In M. Zentner & R. L. Shiner (Eds.), *Handbook of temperament* (pp. 209–228). Guilford Press.
- Gonzalez, J. J., Hynd, G. W., & Martin, R. P. (1994). Neuropsychology of temperament. In Vemon, P. D. (ed.), *The Neuropsychology of Individual Differences* (pp. 235-256). Academic Press.
- Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child development*, 61(6), 1859-1871. doi: 10.1111/j.1467-8624.1990.tb03570.x
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and life history*, 1(1), 69-99. doi: 10.1075/jnlh.1.1.05chi
- Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill-Palmer Quarterly (1982-)*, 109-157. doi: 10.1037/10097-005

## References

- Goodman, G. S., Goldfarb, D. A., Chong, J. Y., & Goodman-Shaver, L. (2014). Children's eyewitness memory: The Influence of cognitive and socio-emotional factors. *Roger Williams UL Rev.*, *19*, 476.
- Goodman, G. S., Golding, J. M., & Haith, M. M. (1984). Jurors' reactions to child witnesses. *Journal of Social Issues*, *40*(2), 139-156. doi: 10.1111/j.1540-4560.1984.tb01098.x
- Goodman, G. S., Golding, J. M., Helgeson, V. S., Haith, M. M., & Michelli, J. (1987). When a child takes the stand. *Law and Human Behavior*, *11*(1), 27-40. doi: 10.1007/BF01044837
- Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior*, *10*(4), 317-332. doi: 10.1007/BF01047344
- Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, *5*(2), 255. doi: 10.1037/1076-8971.5.2.255
- Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P., & Gordon, D. S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect*, *27*(5), 525-540. doi: 10.1016/S0145-2134(03)00037-1
- Gordon, B. N., Ornstein, P. A., Nida, R. E., Follmer, A., Crenshaw, M. C., & Albert, G. (1993). Does the use of dolls facilitate children's memory of visits to the doctor? *Applied Cognitive Psychology*, *7*(6), 459-474. doi: [10.1002/acp.2350070602](https://doi.org/10.1002/acp.2350070602)
- Gous, G., & Wheatcroft, J. M. (2020). Directive leading questions and preparation technique effects on witness accuracy. *SAGE Open*, *10*(1), 1-12. doi: 10.1177/2158244019899053

## References

- Graham, A. M., Ablow, J. C., & Measelle, J. R. (2010). Interparental relationship dynamics and cardiac vagal functioning in infancy. *Infant Behavior and Development*, 33(4), 530-544. doi: [10.1016/j.infbeh.2010.07.005](https://doi.org/10.1016/j.infbeh.2010.07.005)
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366-385. doi: [10.1037/a0021847](https://doi.org/10.1037/a0021847)
- Gray, J. A. & McNaughton, N. (2003). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. (Second edition). Oxford University Press.
- Greene, E. (1988). Judge's instruction on eyewitness testimony: Evaluation and revision. *Journal of Applied Social Psychology*, 18, 252-276. doi: 10.1111/j.1559-1816.1988.tb00016.x
- Greene, E., Chopra, S. R., Kovera, M. B., Penrod, S. D., Rose, V. G., Schuller, R., et al. (2002). Jurors and juries: A review of the field. In J.R.P. Ogloff (Ed.), *Taking psychology and law into the twenty-first century* (pp. 225-284). Kluwer Press.
- Greenhoot, A. F., Ornstein, P. A., Gordon, B. N., & Baker-Ward, L. (1999). Acting out the details of a paediatric check-up: The impact of interview condition and behavioural style on children's memory reports. *Child Development*, 70(2), 363-380. doi: [10.1111/1467-8624.00027](https://doi.org/10.1111/1467-8624.00027)
- Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review*, 93(2), 216-229. doi: [10.1037/0033-295X.93.2.216](https://doi.org/10.1037/0033-295X.93.2.216)
- Gross, J., & Hayne, H. (1996). Eyewitness identification by 5-to 6-year-old children. *Law and Human Behavior*, 20(3), 359-373. doi: 10.1007/BF01499028

## References

- Guerin, D. W., Gottfried, A. W., & Thomas, C. W. (1997). Difficult temperament and behaviour problems: A longitudinal study from 1.5 to 12 years. *International Journal of Behavioral Development, 21*(1), 71-90. doi: 10.1080/016502597384992
- Gudjonsson, G. H. (1992). Interrogative suggestibility: factor analysis of the Gudjonsson Suggestibility Scale (GSS 2). *Personality and Individual Differences, 13*(4), 479-481. doi: 10.1016/0191-8869(92)90077-3
- Gudjonsson, G.H. (2003). *The Psychology of Interrogations, Confessions, and Testimony: A Handbook*. John Wiley & Sons, Ltd.
- Gudjonsson, G. H., & Clark, N. K. (1986). Suggestibility in police interrogation: A social psychological model. *Social Behaviour, 1*, 83–104. doi: 10.1016/0191-8869(92)90077-3
- Gudjonsson, G., Vagni, M., Maiorano, T., & Pajardi, D. (2016). Age and memory related changes in children's immediate and delayed suggestibility using the Gudjonsson Suggestibility Scale. *Personality and Individual Differences, 102*, 25-29. doi: 10.1016/j.paid.2016.06.029
- Gurven, M., Von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How universal is the Big Five? Testing the five-factor model of personality variation among forager–farmers in the Bolivian Amazon. *Journal of personality and social psychology, 104*(2), 354-370. doi: 10.1037/a0030841
- Haegerich, T. M., & Bottoms, B. L. (2000). Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. *Law and human behavior, 24*(4), 421-448. doi: 10.1023/A:1005592213294
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon.

## References

- Hanna, K., Davies, E., Crothers, C., & Henderson, E. (2012). Questioning child witnesses in New Zealand's criminal justice system: Is cross-examination fair?. *Psychiatry, Psychology and Law*, *19*(4), 530-546. doi: 10.1080/13218719.2011.615813
- Hannuschke, M., Gollwitzer, M., Geukes, K., Nestler, S., & Back, M. (2020). Neuroticism and interpersonal perception: Evidence for positive, but not negative, biases. *Journal of Personality*, *88*(2), 217-236. doi: 10.1111/jopy.12480
- Harter, S. (1999). *The construction of the self*. Guilford.
- Hassan, H., Asad, S., & Hoshino, Y. (2016). Determinants of leadership style in big five personality dimensions. *Universal Journal of Management*, *4*(4), 161-179. doi: 10.13189/ujm.2016.040402
- Havard, C. (2014). Are children less reliable at making visual identifications than adults? A review. *Psychology, Crime & Law*, *20*(4), 372–388. doi: 10.1080/1068316X.2013.793334
- Havard, C., & Memon, A. (2013). The mystery man can help reduce false identification for child witnesses: Evidence from video line-ups. *Applied Cognitive Psychology*, *27*(1), 50-59. doi: 10.1002/acp.2870
- Havard, C., Memon, A., Clifford, B., & Gabbert, F. (2010). A comparison of video and static photo lineups with child and adolescent witnesses. *Applied Cognitive Psychology*, *24*(9), 1209-1221. doi: 10.1002/acp.1645
- Hazan, C., Hazan, R., & Goodman, G. S. (1984). Translation of Varendonck, J. The testimony of children in a famous trial. *Journal of Social Issues*, *40*, 26–31. doi: 10.1111/j.1540-4560.1984.tb01091.x

## References

- Heaps, C., & Nash, M. (2001). Comparing recollective experience in true and false autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, 27(4), 920-930. doi: 10.1037/0278-7393.27.4.920
- Heiser, A. N., Turner, M. S., & Beidel, C. D. (2003). Relationship to social phobia and other psychiatric disorders. *Behavior Research and Therapy*, 41(2), 209-221. doi: 10.1016/S0005-7967(02)00003-7
- Henry, L., & Gudjonsson, G. (2004). The effects of memory trace strength on eyewitness recall in children with and without intellectual disabilities. *Journal Of Experimental Child Psychology*, 89(1), 53-71. doi: 10.1016/j.jecp.2004.05.002
- Hershkowitz, I., Lamb, M. E., Orbach, Y., Katz, C., & Horowitz, D. (2012). The development of communicative and narrative skills among preschoolers: Lessons from forensic interviews about child abuse. *Child Development*, 83(2), 611-622. doi: 10.1111/j.1467-8624.2011.01704.x
- Hershkowitz, I., & Terner, A. (2007). The effects of repeated interviewing on children's forensic statements of sexual abuse. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 21(9), 1131-1143. doi: 10.1002/acp.1319
- Hill, E. S., & Davies, G. M. (2013). Has the quality of investigative interviews with children improved with changes in guidance? An exploratory study. *Policing: A Journal of Policy and Practice*, 7(1), 63-71. doi: 10.1093/police/pas046
- Hinton, P. R., Brownlow, C., McMurray, I., & Cozens, B. (2004). *SPSS Explained*. Routledge. doi: 10.4324/9780203642597

## References

- Hirsh, J. B., DeYoung, C. G., Xu, X., & Peterson, J. B. (2010). Compassionate liberals and polite conservatives: Associations of agreeableness with political ideology and moral values. *Personality and Social Psychology Bulletin*, *36*(5), 655-664. doi: 10.1177/0146167210366854
- Hlavka, H. R., Olinger, S. D., & Lashley, J. L. (2010). The use of anatomical dolls as a demonstration aid in child sexual abuse interviews: A study of forensic interviewers' perceptions. *Journal of Child Sexual Abuse*, *19*(5), 519-553. doi: 10.1080/10538712.2010.511988
- Hofstee W.K.B. (1991) *The Concepts of Personality and Temperament*. In Strelau J., Angleitner A. (eds) *Explorations in Temperament. Perspectives on Individual Differences* (p. 177-188). Springer. doi: [10.1007/978-1-4899-0643-4\\_12](https://doi.org/10.1007/978-1-4899-0643-4_12)
- Holcomb, M. J., & Jacquin, K. M. (2007). Juror perceptions of child eyewitness testimony in a sexual abuse trial. *Journal of Child Sexual Abuse*, *16*(2), 79-95. doi: 10.1300/J070v16n02\_05
- Home Office. (2007). *Achieving best evidence in criminal proceedings: Guidance on interviewing victims and witnesses and using special measures*. London: TSO.
- Hopwood, A. R. (2016). A.R. *Hopwood's False Memory Archive Anthology: Volume 1*. Amazon.
- Howe, M., Courage, M., & Edison, S. (2003). When autobiographical memory begins. *Developmental Review*, *23*(4), 471-494. doi: 10.1016/j.dr.2003.09.001
- Howie, P., Sheehan, M., Mojarrad, T., & Wrzesinska, M. (2004). 'Undesirable' and 'desirable' shifts in children's responses to repeated questions: Age differences in the effect of providing a rationale for repetition. *Applied Cognitive Psychology: The Official*

## References

- Journal of the Society for Applied Research in Memory and Cognition*, 18(9), 1161-1180.  
doi: 10.1002/acp.1049
- Hritz, A. C., Royer, D. E., Helm, R. K., Burd, K. A., Ojeda, K., & Ceci, S. J. (2015). Children's suggestibility research. Things to know before interviewing a child. *Anuario De Psicología Jurídica*, 2, 3–12. doi: 10.1016/j.apj.2014.09.002
- Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First language*, 1(2), 149-160. doi: 10.1177/014272378000100205
- Humphries, J. E., Holliday, R. E., & Flowe, H. D. (2012). Faces in motion: Age-related changes in eyewitness identification performance in simultaneous, sequential, and elimination video lineups. *Applied Cognitive Psychology*, 26(1), 149-158. doi: 10.1002/acp.1808
- Hwang, J. Y. (2002). *Development of a temperament self-report measure for young children* (Unpublished doctoral thesis). University of Oregon.
- Hurlock, E. (1930). The Suggestibility of Children. *The Pedagogical Seminary And Journal Of Genetic Psychology*, 37(1), 59-74. doi: 10.1080/08856559.1930.10532223
- Imhoff, M. C., & Baker-Ward, L. (1999). Preschoolers' suggestibility: Effects of developmentally appropriate language and interviewer supportiveness. *Journal of Applied Developmental Psychology*, 20(3), 407-429. doi: 10.1016/S0193-3973(99)00022-2
- Innocence Project. Retrieved March 23, 2020, from <https://www.innocenceproject.org/about/#>
- Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). Heritability of facet-level traits in a cross-cultural twin sample: Support for a hierarchical model of

## References

- personality. *Journal of Personality and Social Psychology*, 74(6), 1556-1565. doi: [10.1037/0022-3514.74.6.1556](https://doi.org/10.1037/0022-3514.74.6.1556)
- Jeronimus, B. F., Kotov, R., Riese, H., & Ormel, J. (2016). Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: a meta-analysis on 59 longitudinal/prospective studies with 443 313 participants. *Psychological Medicine*, 46(14), 2883-2906. doi: [10.1017/S0033291716001653](https://doi.org/10.1017/S0033291716001653)
- Johnson, M., Hashtroudi, S., & Lindsay, D. (1993). Source monitoring. *Psychological Bulletin*, 114(1), 3-28. doi: [10.1037/0033-2909.114.1.3](https://doi.org/10.1037/0033-2909.114.1.3)
- Johnson, D. L., Wiebe, J. S., Gold, S. M., Andreasen, N. C., Hichwa, R. D., Watkins, G. L., & Boles Ponto, L. L. (1999). Cerebral blood flow and personality: a positron emission tomography study. *American Journal of Psychiatry*, 156(2), 252-257. doi: [10.1176/ajp.156.2.252](https://doi.org/10.1176/ajp.156.2.252)
- Johnston, V., Benedan, L., Brubacher, S., & Powell, M. (2021). The roles of child temperament and interviewer support on children's reports of adult wrongdoing. *Personality and Individual Differences*, 175, 110732. doi: [10.1016/j.paid.2021.110732](https://doi.org/10.1016/j.paid.2021.110732)
- Judge, T. A., Locke, E. A., Durham, C. C., & Kluger, A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology*, 83(1), 17-34. doi: [10.1037/0021-9010.83.1.17](https://doi.org/10.1037/0021-9010.83.1.17)
- Jung, C. G. (2014). *Psychological types*. Routledge.

## References

- Kagan, J. (1989). Temperamental contributions to social behavior. *American Psychologist*, 44(4), 668–674. doi: [10.1037/0003-066X.44.4.668](https://doi.org/10.1037/0003-066X.44.4.668)
- Kagan, J. (1997). Temperament and the reactions to unfamiliarity. *Child Development*, 68(1), 139-143. doi: [10.1111/j.1467-8624.1997.tb01931.x](https://doi.org/10.1111/j.1467-8624.1997.tb01931.x)
- Kagan, J., & Fox, N. A. (2006). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 167–225). Wiley.
- Kail, R. V. (2012). *Children and their development* (7<sup>th</sup> edition). Pearson.
- Kale, P., Inamdar, M., Shimpi, P., & Jha, P. (2020). Personality Assessment Tool using Artificial Intelligence: A Review. *International Journal of Progressive Research in Science and Engineering*, 1(3), 104-107.
- Kappe, R., & Van Der Flier, H. (2012). Predicting academic success in higher education: what's more important than being smart?. *European Journal of Psychology of Education*, 27(4), 605-619. doi: [10.1007/s10212-011-0099-9](https://doi.org/10.1007/s10212-011-0099-9)
- Karageorge, A., & Zajac, R. (2011). Exploring the effects of age and delay on children's person identifications: Verbal descriptions, lineup performance, and the influence of wildcards. *British Journal of Psychology*, 102(2), 161-183. doi: [10.1348/000712610X507902](https://doi.org/10.1348/000712610X507902)
- Karavasilis, L., Doyle, A. B., & Markiewicz, D. (2003). Associations between parenting style and attachment to mother in middle childhood and adolescence. *International Journal of Behavioral Development*, 27(2), 153-164. doi: [10.1080/0165025024400015](https://doi.org/10.1080/0165025024400015)

## References

- Karla, B., & Heath, W. P. (1997). Perceptions of a child witness: Effects of leading questions and the type of relationship between child and defendant. *Psychological Reports, 979-986*. doi: 10.2466/pr0.1997.80.3.979
- Karpinski, A., & Scullin, M. (2009). Suggestibility under pressure: Theory of mind, executive function, and suggestibility in preschoolers. *Journal Of Applied Developmental Psychology, 30(6)*, 749-763. doi: 10.1016/j.appdev.2009.05.004
- Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2016). Openness to experience and intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality, 84(2)*, 248-258. doi: 10.1111/jopy.12156
- Keast, A., Brewer, N., & Wells, G. L. (2007). Children's metacognitive judgments in an eyewitness identification task. *Journal of Experimental Child Psychology, 97(4)*, 286-314. doi: 10.1016/j.jecp.2007.01.007
- Kebbell, M. R., Evans, L., & Johnson, S. D. (2010). The influence of lawyers' questions on witness accuracy, confidence, and reaction times and on mock jurors' interpretation of witness accuracy. *Journal of Investigative Psychology and Offender Profiling, 7(3)*, 262-272. doi: 10.1002/jip.125
- Keller, A., Ford, L. H., & Meacham, J. A. (1978). Dimensions of self-concept in preschool children. *Developmental Psychology, 14(5)*, 483-489. doi: [10.1037/0012-1649.14.5.483](https://doi.org/10.1037/0012-1649.14.5.483)
- Keogh, B. K. (1982). Temperament: An individual difference of importance in intervention programs. *Topics in Early Childhood Special Education, 2(2)*, 25-31. doi: 10.1177/027112148200200207

## References

- Kertechian, S. K. (2018). Conscientiousness as a key to success for academic achievement among French university students enrolled in management studies. *The International Journal of Management Education*, 16(2), 154-165. doi: 10.1016/j.ijme.2018.02.003
- Kilianski, S. E. (2008). Who do you think I think I am? Accuracy in perceptions of others' self-esteem. *Journal of Research in Personality*, 42(2), 386-398. doi: 10.1016/j.jrp.2007.07.004
- King, L. A., Walker, L. M., & Broyles, S. J. (1996). Creativity and the five-factor model. *Journal of Research in Personality*, 30(2), 189-203. doi: 10.1006/jrpe.1996.0013
- Klemfuss, J. Z., & Olaguez, A. P. (2020). Individual differences in children's suggestibility: An updated review. *Journal of Child Sexual Abuse*, 29(2), 158-182. doi: [10.1080/10538712.2018.1508108](https://doi.org/10.1080/10538712.2018.1508108)
- Klemfuss, J., Rush, E., & Quas, J. (2016). Parental reminiscing style and children's suggestibility about an alleged transgression. *Cognitive Development*, 40, 33-45. doi: 10.1016/j.cogdev.2016.08.003
- Krähenbühl, S. J. (2007). *The effect of question repetition on young children's eyewitness testimony* (Unpublished doctoral dissertation). University of Sheffield.
- Krähenbühl, S., & Blades, M. (2006). The effect of question repetition within interviews on young children's eyewitness recall. *Journal Of Experimental Child Psychology*, 94(1), 57-67. doi: 10.1016/j.jecp.2005.12.002
- Krähenbühl, S. J., & Blades, M. (2009). Does the form of question repetition have an effect on children's recall accuracy and consistency?. *International Journal of Police Science & Management*, 11(4), 460-475. doi: 10.1350/ijps.2009.11.4.144

## References

- Kreitler, S., & Kreitler, H. (2013). *The cognitive foundations of personality traits*. Springer Science & Business Media.
- Kuehn, L. L. (1974). Looking down a gun barrel: Person perception and violent crime. *Perceptual and Motor Skills*, 39(3), 1159-1164. doi: 10.2466/pms.1974.39.3.1159
- Kwiatkowska, M. M., & Rogoza, R. (2019). Shy teens and their peers: Shyness in respect to basic personality traits and social relations. *Journal of Research in Personality*, 79, 130-142. doi: 10.1016/j.jrp.2019.03.005
- La Rooy, D. J., Earhart, B., & Nicol, A. (2013). Joint investigative interviews (JIIIs) conducted with children in Scotland: a comparison of the quality of interviews conducted before and after the introduction of the Scottish Executive (2011) guidelines. *Scots Law Times*, 31, 4-10.
- La Rooy, D., Pipe, M. E., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology*, 90(3), 235-254. doi: 10.1016/j.jecp.2004.11.002
- Lamb, M., & Brown, D. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal Of Developmental Psychology*, 24(1), 215-234. doi: 10.1348/026151005x57657
- Lamb, M. E., Brown, D. A., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Questioning children about abuse* (2<sup>nd</sup> edition). John Wiley & Sons
- Lamb, M., Malloy, L., & La Rooy, D. J. (2011). Setting realistic expectations: Developmental characteristics, capacities and limitations. In M. Lamb, J. La Rooy, L. Malloy, & C. Katz

## References

- (Eds.), *Children's testimony: a handbook of psychological research and forensic practice* (pp. 15-48). John Wiley & Sons.
- Lamb, M., Orbach, Y., Hershkowitz, I., Esplin, P., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31*(11-12), 1201-1231. doi: 10.1016/j.chiabu.2007.03.021
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Aldridge, J. A. N., Pearson, S., Stewart, H. L., ... & Bowler, L. (2009). Use of a structured investigative protocol enhances the quality of investigative interviews with alleged victims of child sexual abuse in Britain. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23*(4), 449-467. doi: 10.1002/acp
- Lamb, M., Sternberg, K., Orbach, Y., Esplin, P., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews. *Journal Of Consulting And Clinical Psychology, 71*(5), 926-934. doi: 10.1037/0022-006x.71.5.926
- Lander, K., & Poyarekar, S. (2015). Famous face recognition, face matching, and extraversion. *Quarterly Journal of Experimental Psychology, 68*(9), 1769-1776. doi: 10.1080/17470218.2014.988737
- Lanning, K. (1994). Dimensionality of observer ratings on the California Adult Q-set. *Journal of Personality and Social Psychology, 67*(1), 151-160. doi: [10.1037/0022-3514.67.1.151](https://doi.org/10.1037/0022-3514.67.1.151)
- Larsen, R. J. & Buss, D. M. (2001). *Personality Psychology: Domains of Knowledge about Human Nature*. (3<sup>rd</sup> edition). McGraw-Hill.

## References

- Larson, B. A., & Brodsky, S. L. (2014). Assertive women as expert witnesses: A study of assertive and defensive responses in male and female experts. *Behavioral Sciences & the Law*, 32(2), 149-163. doi: 10.1002/bsl.2116
- Larsson, H., Andershed, H., & Lichtenstein, P. (2006). A genetic factor explains most of the variation in the psychopathic personality. *Journal of Abnormal Psychology*, 115(2), 221–230. doi: [10.1037/0021-843X.115.2.221](https://doi.org/10.1037/0021-843X.115.2.221)
- Leander, L., Christianson, S. Å., & Granhag, P. A. (2007). A sexual abuse case study: Children's memories and reports. *Psychiatry, Psychology and Law*, 14(1), 120-129. doi: 10.1375/pplt.14.1.120
- Leander, L., Granhag, P. A., & Christianson, S. Å. (2005). Children exposed to obscene phone calls: What they remember and tell. *Child Abuse & Neglect*, 29(8), 871-888. doi: 10.1016/j.chiabu.2004.12.012
- Lee, K. (2004). Age, neuropsychological, and social cognitive measures as predictors of individual differences in susceptibility to the misinformation effect. *Applied Cognitive Psychology*, 18(8), 997-1019. doi: 10.1002/acp.1075
- Lee, K., & Ashton, M. C. (2008). The HEXACO personality factors in the indigenous personality lexicons of English and 11 other languages. *Journal of personality*, 76(5), 1001-1054. doi: 10.1111/j.1467-6494.2008.00512.x
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, 31(4), 568-578. doi: 10.1037/0012-1649.31.5.758

## References

- Le Pine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*(3), 563-593. doi: 10.1111/j.1744-6570.2000.tb00214.x
- Li, J., Tian, M., Fang, H., Xu, M., Li, H., & Liu, J. (2010). Extraversion predicts individual differences in face recognition. *Communicative & Integrative Biology, 3*(4), 295-298. doi: 10.4161/cib.3.4.12093
- Lindsay, R. C., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions. *Law and Human Behavior, 18*(5), 527-541. doi: 10.1007/BF01499172
- Lindsay, R. C., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior, 21*(4), 391-404. doi: 10.1023/A:1024807202926
- Linder, D. (2007). The McMartin Preschool Abuse Trial. *Available at SSRN 1030559*.
- Lippert, T., Cross, T. P., Jones, L., & Walsh, W. (2009). Telling interviewers about sexual abuse: Predictors of child disclosure at forensic interviews. *Child Maltreatment, 14*(1), 100-113. doi: 10.1177/1077559508318398
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive psychology, 7*(4), 560-572. doi: 10.1016/0010-0285(75)90023-7
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & memory, 12*(4), 361-366. doi: 10.1101/lm.94705
- Loftus, E., & Davies, G. (1984). Distortions in the Memory of Children. *Journal Of Social Issues, 40*(2), 51-67. doi: 10.1111/j.1540-4560.1984.tb01093.x

## References

- Loftus, E., Miller, D., & Burns, H. (1978). Semantic integration of verbal information into a visual memory. *Journal Of Experimental Psychology: Human Learning And Memory*, 4(1), 19-31. doi: 10.1037/0278-7393.4.1.19
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, 13(5), 585–589. doi: [10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3)
- Loftus, E., & Pickrell, J. (1995). The Formation of False Memories. *Psychiatric Annals*, 25(12), 720-725. doi: 10.3928/0048-5713-19951201-07
- Loggia, M. L., Mogil, J. S., & Bushnell, M. C. (2008). Empathy hurts: compassion for another increases both sensory and affective components of pain perception. *Pain*, 136(1-2), 168-176. doi: 10.1016/j.pain.2007.07.017
- London, K., Bruck, M., Wright, D., & Ceci, S. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological issues, and implications for forensic interviewers. *Memory*, 16(1), 29-47. doi: 10.1080/09658210701725732
- London, K., Henry, L. A., Conradt, T., & Corser, R. (2013). Suggestibility and individual differences in typically developing and intellectually disabled children. *Suggestibility in legal contexts: Psychological research and forensic implications*, 129-148. doi: 10.1002/9781118432907
- Luby, J. L., Svrakic, D. M., McCallum, K., Przybeck, T. R., & Cloninger, C. R. (1999). The Junior Temperament and Character Inventory: Preliminary validation of a child self-report measure. *Psychological Reports*, 84, 1127–1138. doi: 10.2466/pr0.1999.84.3c.1127

## References

- Luther, K., Snook, B., Barron, T., & Lamb, M. E. (2015). Child interviewing practices in Canada: A box score from field observations. *Journal of Police and Criminal Psychology, 30*(3), 204-212. doi: 10.1007/s11896-014-9149-y
- Luus, C. A. E., Wells, G. L., & Turtle, J. W. (1995). Child eyewitnesses: Seeing is believing. *Journal of Applied Psychology, 80*, 317-326. doi: 10.1037/0021-9010.80.2.317
- Lyon, T. (2014). Interviewing Children. *Annual Review Of Law And Social Science, 10*(1), 73-89. doi: 10.1146/annurev-lawsocsci-110413-030913
- MacDonald, D. A. (2000). Spirituality: Description, measurement, and relation to the five factor model of personality. *Journal of Personality, 68*(1), 153-197. doi: 10.1111/1467-6494.t01-1-00094
- Machida, S., Taylor, A. R., & Kim, J. (2002). The role of maternal beliefs in predicting home learning activities in Head Start families. *Family Relations, 51*(2), 176-184. doi: 10.1111/j.1741-3729.2002.00176.x
- Magnusson, M., Ernberg, E., & Landström, S. (2017). Preschoolers' disclosures of child sexual abuse: Examining corroborated cases from Swedish courts. *Child Abuse & Neglect, 70*, 199-209. doi: 10.1016/j.chiabu.2017.05.018
- MacLennan v. HM Advocate (2015).
- Majdandžić, M., van den Boom, D. C., & Heesbeen, D. G. (2008). Peas in a pod: biases in the measurement of sibling temperament?. *Developmental Psychology, 44*(5), 1354-1368. doi: 10.1037/a0013064
- Malloy, L., Lyon, T., & Quas, J. (2007). Filial dependency and recantation of child sexual abuse allegations. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(2), 162-170. doi: 10.1097/01.chi.0000246067.77953.f7

## References

- Mann, J. J., Arango, V., & Underwood, M. D. (1990). Serotonin and suicidal behavior. *Annals of the New York Academy of Sciences*, 600, 476–485. doi: [10.1111/j.1749-6632.1990.tb16903.x](https://doi.org/10.1111/j.1749-6632.1990.tb16903.x)
- Marcus, D. K., Lyons, P. M., & Guyton, M. R. (2000). Studying perceptions of juror influence in vivo: A social relations analysis. *Law and Human Behavior*, 24(2), 173-186. doi: 10.1023/A:1005406902505
- Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human behavior*, 3(4), 295-306. doi: 10.1007/BF01039808
- Mark, I. L. V. D., IJzendoorn, M. H. V., & Bakermans-Kranenburg, M. J. (2002). Development of empathy in girls during the second year of life: Associations with parenting, attachment, and temperament. *Social development*, 11(4), 451-468. doi: 10.1111/1467-9507.00210
- Markey, P., Markey, C., & Tinsley, B. (2004). Children's behavioral manifestations of the Five-Factor Model of personality. *Personality and Social Psychology Bulletin*, 30, 423–432. doi: 10.1177/0146167203261886
- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach. *Journal of Personality and Social Psychology*, 88(1), 139–157. doi: [10.1037/0022-3514.88.1.139](https://doi.org/10.1037/0022-3514.88.1.139)
- Marshall, G. N., Wortman, C. B., Vickers, R. R., Kusulas, J. W., & Hervig, L. K. (1994). The five-factor model of personality as a framework for personality-health research. *Journal of Personality and Social Psychology*, 67(2), 278–286. doi: [10.1037/0022-3514.67.2.278](https://doi.org/10.1037/0022-3514.67.2.278)
- Martin, R. (1988). *The temperament assessment battery for children*. Clinical Psychology Publishing Company. doi: [10.1037/t16139-000](https://doi.org/10.1037/t16139-000)

## References

- Martin, R. P., & Bridger, R. C. (1999). *The temperament assessment battery for children-revised*. Department of Educational Psychology, University of Georgia.
- Martin, R. P., Wisenbaker, J., Baker, J., & Huttunen, M. O. (1997). Gender differences in temperament at six months and five years. *Infant Behavior and Development, 20*(3), 339-347. doi: 10.1016/S0163-6383(97)90005-9
- Mathiesen, K. S., & Tambs, K. (1999). The EAS Temperament Questionnaire-Factor structure, age trends, reliability, and stability in a Norwegian sample. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(3), 431-439. doi: [10.1111/1469-7610.00460](https://doi.org/10.1111/1469-7610.00460)
- Matz, S. C., & Gladstone, J. J. (2020). Nice guys finish last: When and why agreeableness is associated with economic hardship. *Journal of Personality and Social Psychology, 118*(3), 545–561. doi: [10.1037/pspp0000220](https://doi.org/10.1037/pspp0000220)
- Maziade, M., Boudreault, M., Thivierge, J., Capéraà, P., & Côté, R. (1984). Infant temperament: SES and gender differences and reliability of measurement in a large Quebec sample. *Journal of the American Academy of Child Psychiatry, 23*(5), 213-226. doi: 10.1016/S0002-7138(09)60351-4
- Maziade, M., Côté, R., Boudreault, M., Thivierge, J., & Capéraà, P. (1984). The New York Longitudinal Studies model of temperament: Gender differences and demographic correlates of a French-speaking population. *Journal of the American Academy of Child Psychiatry, 23*(5), 582–587. doi: 10.1016/S0002-7138(09)60351-4
- McCarron, A. L., Ridgway, S., & Williams, A. (2004). The truth and lie story: Developing a tool for assessing child witnesses' ability to differentiate between truth and lies. *Child*

## References

- Abuse Review: Journal of the British Association for the Study and Prevention of Child Abuse and Neglect*, 13(1), 42-50. doi: 10.1002/car.832
- McCauley, M. R. & Parker, J. F. (2001). When will a child be believed? The impact of the victim's age and juror's gender on children's credibility and verdict in a sexual abuse case. *Child Abuse & Neglect*, 25, 523-539. doi: 10.1016/S0145-2134(01)00224-1
- McConnell, T. R., Jr. (1963). Suggestibility in children as a function of chronological age. *The Journal of Abnormal and Social Psychology*, 67(3), 286–289. doi: [10.1037/h0042170](https://doi.org/10.1037/h0042170)
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 52(6), 1258–1265. doi: [10.1037/0022-3514.52.6.1258](https://doi.org/10.1037/0022-3514.52.6.1258)
- McCrae, R. R., & Costa, P. T. (1985a). Updating Norman's "adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49(3), 710-721. doi: 10.1037/0022-3514.49.3.710
- McCrae, R. R., & Costa Jr, P. T. (1985b). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences*, 6(5), 587-597. doi: 10.1016/0191-8869(85)90008-X
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81–90. doi: [10.1037/0022-3514.52.1.81](https://doi.org/10.1037/0022-3514.52.1.81)
- McCrae, R. R., & Costa, P. T., Jr. (1997). Conceptions and correlates of openness to experience. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (p. 825–847). Academic Press. doi: 10.1016/B978-012134645-4/50032-9

## References

- McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: data from 50 cultures. *Journal of personality and social psychology, 88*(3), 547-561. doi: 10.1037/0022-3514.88.3.547
- McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3-7 year old children. *Journal of Child Psychology and Psychiatry, 19*(3), 245-253. doi: [10.1111/j.1469-7610.1978.tb00467.x](https://doi.org/10.1111/j.1469-7610.1978.tb00467.x)
- McDevitt, S. C., & Carey, W. B. (1996). *Manual for the behavioral style questionnaire*. Behavioural-Developmental Initiatives.
- McDonald, R., Jouriles, E., Ramisetty-Mikler, S., Caetano, R., & Green, C. (2006). Estimating the number of American children living in partner-violent families. *Journal Of Family Psychology, 20*(1), 137-142. doi: 10.1037/0893-3200.20.1.137
- McDonald, N. M., & Messinger, D. S. (2011). The development of empathy: How, when, and why. *Moral Behavior and Free Will: A Neurobiological and Philosophical Approach, 341-368*.
- McFarlane, F., Powell, M., & Dudgeon, P. (2002). An examination of the degree to which IQ, memory performance, socio-economic status and gender predict young children's suggestibility. *Legal And Criminological Psychology, 7*(2), 227-239. doi: 10.1348/135532502760274729
- McGowen, R., & King, G. D. (1982). Effects of authoritarian, anti-authoritarian, and egalitarian legal attitudes on mock juror and jury decisions. *Psychological Reports, 51*(3\_suppl), 1067-1074. doi: 10.2466/pr0.1982.51.3f.1067

## References

- McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29*(1), 96-109. doi: [10.1037/0012-1649.29.1.96](https://doi.org/10.1037/0012-1649.29.1.96)
- McWilliams, K., Narr, R., Goodman, G. S., Ruiz, S., & Mendoza, M. (2013). Children's memory for their mother's murder: Accuracy, suggestibility, and resistance to suggestion. *Memory, 21*(5), 591-598. doi: [10.1080/09658211.2013.763983](https://doi.org/10.1080/09658211.2013.763983)
- Meares, J. (2016). *Dangerous women and reasonable men: Gender and eyewitness testimony at the turn of the 20th century*. (Unpublished doctoral dissertation). University of Canterbury. Retrieved from [https://ir.canterbury.ac.nz/bitstream/handle/10092/12983/Meares per cent20percent20MA per cent20Thesis.pdf?sequence=1&isAllowed=y](https://ir.canterbury.ac.nz/bitstream/handle/10092/12983/Meares%20percent20MA%20Thesis.pdf?sequence=1&isAllowed=y).
- Measelle, J. R., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (1998). Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child development, 69*(6), 1556-1576. doi: [10.1111/j.1467-8624.1998.tb06177.x](https://doi.org/10.1111/j.1467-8624.1998.tb06177.x)
- Measelle, J. R., John, O. P., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (2005). Can Children Provide Coherent, Stable, and Valid Self-Reports on the Big Five Dimensions? A Longitudinal Study From Ages 5 to 7. *Journal of Personality and Social Psychology, 89*(1), 90-106. doi: [10.1037/0022-3514.89.1.90](https://doi.org/10.1037/0022-3514.89.1.90)
- Mervielde, I., De Clercq, B., De Fruyt, F., & Van Leeuwen, K. (2005). Temperament, personality, and developmental psychopathology as childhood antecedents of personality disorders. *Journal of personality disorders, 19*(2), 171-201. doi: [10.1521/pedi.19.2.171.62627](https://doi.org/10.1521/pedi.19.2.171.62627)

## References

- Mestre Escrivá, M. V., Samper García, P., Frías Navarro, M. D., & Tur Porcar, A. M. (2009). Are women more empathetic than men? A longitudinal study in adolescence. *The Spanish Journal of Psychology*, 12(1), 76-83.
- Melinder, A., Endestad, T., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal Of Psychology*, 47(6), 485-495. doi: 10.1111/j.1467-9450.2006.00542.x
- Melnyk, L. (2002). *The influence of imagery, timing, and individual differences on the accuracy of children's recall*. (Unpublished doctoral dissertation). McGill University.
- Memon, A., Havard, C., Clifford, B., Gabbert, F., & Watt, M. (2011). A field evaluation of the VIPER system: A new technique for eliciting eyewitness identification evidence. *Psychology, Crime & Law*, 17(8), 711-729. doi: 10.1080/10683160903524333
- Memon, A., Holley, A., Wark, L., Bull, R., & Koehnken, G. (1996). Reducing suggestibility in child witness interviews. *Applied Cognitive Psychology*, 10(6), 503-518. doi: [10.1002/\(SICI\)1099-0720\(199612\)10:6<503::AID-ACP416>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-0720(199612)10:6<503::AID-ACP416>3.0.CO;2-R)
- Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology*, 87(3), 403-415. doi: 10.1111/j.2044-8295.1996.tb02598.x
- Messerschmidt, R. (1933). The suggestibility of boys and girls between the ages of six and sixteen years. *The Pedagogical Seminary and Journal of Genetic Psychology*, 43, 422-437. doi: [10.1080/08856559.1933.10532468](https://doi.org/10.1080/08856559.1933.10532468)
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

## References

- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*(4), 371 – 378. doi: 10.1037/h0040525
- Miller, J. D., Lynam, D., Zimmerman, R. S., Logan, T. K., Leukefeld, C., & Clayton, R. (2004). The utility of the Five Factor Model in understanding risky sexual behavior. *Personality and Individual Differences, 36*(7), 1611-1626. doi: 10.1016/j.paid.2003.06.009
- Milne, R., Clare, I. C., & Bull, R. (2002). Interrogative suggestibility among witnesses with mild intellectual disabilities: The use of an adaptation of the GSS. *Journal of Applied Research in Intellectual Disabilities, 15*(1), 8-17. doi: 10.1046/j.1360-2322.2001.00096.x
- Miyamoto, Y. R., Wang, S., & Smith, M. A. (2020). Implicit adaptation compensates for erratic explicit strategy in human motor learning. *Nature Neuroscience, 23*(3), 443-455.
- Mondak, J. J., & Halperin, K. D. (2008). A Framework for the Study of Personality and Political Behaviour. *British Journal of Political Science, 23*, 443-455. doi: 10.1038/s41593-020-0600-3
- Montoya, E.R., Terburg, D., Bos, P.A., & van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and Emotion, 36*(1), 65-73. doi: 10.1007/s11031-011-9264-3
- Moran, G., & Comfort, J. C. (1986). Neither "tentative" nor "fragmentary": Verdict preference of impaneled felony jurors as a function of attitude toward capital punishment. *Journal of Applied Psychology, 71*(1), 146-155. doi: [10.1037/0021-9010.71.1.146](https://doi.org/10.1037/0021-9010.71.1.146)
- Morgan III, C. A., Southwick, S., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful

## References

- events. *International Journal of Law and Psychiatry*, 36(1), 11-17. doi: 10.1016/j.ijlp.2012.11.002
- Morrison, J., Forrester-Jones, R., Bradshaw, J., & Murphy, G. (2019). Communication and cross-examination in court for children and adults with intellectual disabilities: A systematic review. *The International Journal of Evidence & Proof*, 23(4), 366-398. doi: 10.1177/1365712719851134
- Mossiere, A., & Dalby, J. T. (2008). The influence of gender and age in mock juror decision-making. *Europe's Journal of Psychology*, 4(4), 1-11.
- Mumford, M. D., Baughman, W. A., Threlfall, K. V., Uhlman, C. E., & Costanza, D. P. (1993). Personality, adaptability, and performance: Performance on well-defined problem solving tasks. *Human Performance*, 6(3), 241-285. doi: [10.1207/s15327043hup0603\\_4](https://doi.org/10.1207/s15327043hup0603_4)
- Muir-Broadbudd, J., King, T., Downey, D., & Petersen, M. (1998). Conservation as a predictor of individual differences in children's susceptibility to leading questions. *Psychonomic Bulletin & Review*, 5(3), 454-458. doi: [10.3758/BF03208821](https://doi.org/10.3758/BF03208821)
- Muris, P., & Ollendick, T. H. (2005). The role of temperament in the etiology of child psychopathology. *Clinical Child and Family Psychology Review*, 8(4), 271-289. doi: 10.1007/s10567-005-8809-y
- Myklebust, T., & Alison, L. (2000). The current state of police interviews with children in Norway: How discrepant are they from models based on current issues in memory and communication?. *Psychology, Crime and Law*, 6(4), 331-351. doi: 10.1080/10683160008409810

## References

- Nagoshi, J. L., Adams, K. A., Terrell, H. K., Hill, E. D., Brzuzy, S., & Nagoshi, C. T. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles, 59*, 521-531. doi: 10.1007/s11199-008-9458-7
- Narby, D. J., Cutler, B. L., & Moran, G. (1993). A meta-analysis of the association between authoritarianism and jurors' perceptions of defendant culpability. *Journal of Applied Psychology, 78*(1), 34-42. doi: 10.1037/0021-9010.78.1.34
- National Records of Scotland (Mid-2012) population estimates. (2012, June 30). Retrieved March 23, 2020, from <http://marine.gov.scot/information/national-records-scotland-mid-2012-population-estimates>
- National Registry of Exonerations. Retrieved March 23, 2020, from <http://www.law.umich.edu/special/exoneration/Pages/about.aspx>
- Neal, T. M., & Brodsky, S. L. (2008). Expert witness credibility as a function of eye contact behavior and gender. *Criminal Justice and Behavior, 35*(12), 1515-1526. doi: 10.1177/0093854808325405
- Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological review, 111*(2), 486-511. doi: 10.1037/0033-295X.111.2.486
- Newcombe, P. A., & Bransgrove, J. (2007). Perceptions of witness credibility: Variations across age. *Journal of Applied Developmental Psychology, 28*(4), 318-331. doi: 10.1016/j.appdev.2007.04.003
- Nicholson, A. S., Yarbrough, A. M., & Penrod, S. D. (2014). Jury decision making and eyewitness testimony. *Encyclopedia of criminology and criminal justice, 2727-2735*.

## References

- Nicolas, S., Collins, T., Gounden, Y., & Roediger, H. L. III. (2011). Natural suggestibility in children. *Consciousness and Cognition: An International Journal*, 20(2), 394–398. doi: [10.1016/j.concog.2010.11.001](https://doi.org/10.1016/j.concog.2010.11.001)
- Nietzel, M. T., McCarthy, D. M., & Kern, M. J. (1999). Juries: the current state of the empirical literature. In R. Roesch, S. D. Hart, & J. R. P. Ogloff (Eds.), *Psychology and law: The state of the discipline* (pp. 23–42). Kluwer/Plenum.
- Nightingale, N. N. (1993). Juror reactions to child victim witnesses. *Law and Human Behavior*, 17(6), 679-694. doi: 10.1007/BF01044689
- Nikonova, O., & Ogloff, J. R. (2005). Mock Jurors' Perceptions of Child Witnesses: The Impact of Judicial Warning. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 37(1), 1-19. doi: 10.1080/1068316X.2018.1543421
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6), 574-583. doi: 10.1037/h0040291
- Nunez, N., Kehn, A., & Wright, D. (2010). When children are witnesses: The effects of context, age and gender on adults' perceptions of cognitive ability and honesty. *Applied Cognitive Psychology*, 25(3), 460-468. doi: 10.1002/acp.1713
- Nystrom, B., & Bengtsson, H. (2017). A psychometric evaluation of the Temperament in Middle Childhood Questionnaire (TMCQ) in a Swedish sample. *Scandinavian journal of psychology*, 58(6), 477-484. doi: [10.1111/sjop.12393](https://doi.org/10.1111/sjop.12393)
- Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medicine, Science and the Law*, 31(1), 4-10.

## References

- O'Donohue, W., Cummings, C., & Willis, B. (2018). The frequency of false allegations of child sexual abuse: A critical review. *Journal of Child Sexual Abuse, 27*(5), 459-475. doi: 10.1080/10538712.2018.1477224
- O'Donohue, W. T., & Fanetti, M. (2016). *Forensic interviews regarding child sexual abuse*. Springer.
- Office for National Statistics. (2020, October 28). *Crime in England and Wales*. Retrieved 2 January 2021, from <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingiune2020>.
- Olaguez, A. P., & Klemfuss, J. Z. (2020). Differential effects of direct and cross examination on mock jurors' perceptions and memory in cases of child sexual abuse. *Psychiatry, Psychology and Law, 27*(5), 778-796. doi: 10.1080/13218719.2020.1742239
- Onraet, E., Van Hiel, A., Roets, A., & Cornelis, I. (2011). The closed mind: 'Experience' and 'cognition' aspects of openness to experience and need for closure as psychological bases for right-wing attitudes. *European Journal of Personality, 25*(3), 184-197. doi: 10.1002/per.775
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child abuse & neglect, 24*(6), 733-752. doi: 10.1016/S0145-2134(00)00137-X
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged

## References

- child abuse victims. *Child Abuse & Neglect*, 24(6), 733-752. doi: 10.1016/S0145-2134(00)00137-X
- Orbach, Y., & Lamb, M. (2007). Young Children's References to Temporal Attributes of Allegedly Experienced Events in the Course of Forensic Interviews. *Child Development*, 78(4), 1100-1120. doi: 10.1111/j.1467-8624.2007.01055.x
- Otgaar, H., Candel, I., Merckelbach, H., & Wade, K. (2009). Abducted by a UFO: prevalence information affects young children's false memories for an implausible event. *Applied Cognitive Psychology*, 23(1), 115-125. doi: 10.1002/acp.1445
- Otgaar, H., Candel, I., Smeets, T., & Merckelbach, H. (2010). 'You didn't take Lucy's skirt off': The effect of misleading information on omissions and commissions in children's memory reports. *Legal and Criminological Psychology*, 15(2), 229-241. doi: 10.1348/135532509X471951
- Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General*, 145(1), 31-55. doi: 10.1037/xge0000127
- Otgaar, H., Howe, M. L., Peters, M., Smeets, T., & Moritz, S. (2014). The production of spontaneous false memories across childhood. *Journal of Experimental Child Psychology*, 121, 28-41. doi: 10.1016/j.jecp.2013.11.019.
- Ornstein, P. A., Shapiro, L. B., Clubb, P. A., Follmer, A., & Baker-Ward, L. (1997). The influence of prior knowledge on children's memory for salient medical experiences. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 83-112). Erlbaum.

## References

- Pantell, R. (2017). The Child Witness in the Courtroom. *Pediatrics*, *139*(3), e20164008. doi: 10.1542/peds.2016-4008
- Pardini, D. A., Raine, A., Erickson, K., & Loeber, R. (2014). Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence. *Biological Psychiatry*, *75*(1), 73-80. doi: 10.1016/j.biopsych.2013.04.003
- Parker, G., & Brotchie, H. (2010). Gender differences in depression. *International Review of Psychiatry*, *22*(5), 429-436. doi: 10.3109/09540261.2010.492391
- Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior*, *13*(2), 133-149. doi: 10.1007/BF01055920
- Parkes, K. R., & Razavi, T. D. (2004). Personality and attitudinal variables as predictors of voluntary union membership. *Personality and Individual Differences*, *37*(2), 333-347. doi: 10.1016/j.paid.2003.09.007
- Patihis, L., Frenda, S. J., LePort, A. K., Petersen, N., Nichols, R. M., Stark, C. E., ... & Loftus, E. F. (2013). False memories in highly superior autobiographical memory individuals. *Proceedings of the National Academy of Sciences*, *110*(52), 20947-20952. doi: 10.1073/pnas.1314373110
- Paulhus, D. L. (1991). Measurement control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes: Vol. I: Measures of personality and social psychological attitudes* (p. 17-59). Academic Press.  
[doi: 10.1016/B978-0-12-590241-0.50006-X](https://doi.org/10.1016/B978-0-12-590241-0.50006-X)

## References

- Paulhus, D. L., & Trapnell, P. D. (1998). Typological measures of shyness: Additive, interactive, and categorical. *Journal of Research in Personality, 32*(2), 183-201. doi: 10.1006/jrpe.1998.2215
- Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. *Journal of personality and social psychology, 84*(2), 411-424. doi: 10.1037/0022-3514.84.2.411
- Paunonen, S. V., Haddock, G., Forsterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European Journal of Personality, 17*(6), 413-433. doi: 10.1002/per.496
- Paz-Alonso, P. M., & Goodman, G. S. (2016). Developmental differences across middle childhood in memory and suggestibility for negative and positive events. *Behavioral Sciences & the Law, 34*(1), 30-54. doi: 10.1002/bsl.2239
- Pearse, S. L., Powell, M. B., & Thomson, D. M. (2003). The effect of contextual cues on children's ability to remember an occurrence of a repeated event. *Legal and Criminological Psychology, 8*(1), 39-50. doi: 10.1348/135532503762871228
- Pellegrini, A. D., & Bartini, M. (2000). A longitudinal study of bullying, victimization, and peer affiliation during the transition from primary school to middle school. *American Educational Research Journal, 37*(3), 699-725. doi: 10.3102/00028312037003699
- Pellegrini, A. D., Galda, L., Flor, D., Bartini, M., & Charak, D. (1997). Close relationships, individual differences, and early literacy learning. *Journal of Experimental Child Psychology, 67*(3), 409-422. doi: 10.1006/jecp.1997.2415
- Pendergrast, M. (2017). *The Repressed Memory Epidemic*. Springer. doi: 10.1007/978-3-319-63375-6\_5

## References

- Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and wh- questions. *Law and Human Behavior, 23*(5), 539-555. doi: 10.1023/a:1022396112719
- Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 15*(7), S7-S24. doi: 10.1002/acp.832
- Pezdek, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology, 17*(9), 1033-1045. doi: 10.1002/acp.984
- Pezdek, K., & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology, 60*(1), 116-128. doi: [10.1006/jecp.1995.1034](https://doi.org/10.1006/jecp.1995.1034)
- Pfeffer, J., & Martin, R. P. (1983). Comparison of mothers' and fathers' temperament ratings of referred and nonreferred preschool children. *Journal of Clinical Psychology, 39*(6), 1013–1020. doi: [10.1002/1097-4679\(198311\)39:6<1013::AID-JCLP2270390634>3.0.CO;2-5](https://doi.org/10.1002/1097-4679(198311)39:6<1013::AID-JCLP2270390634>3.0.CO;2-5)
- Phares, E. J., & Wilson, K. G. (1972). Responsibility attribution: Role of outcome severity, situational ambiguity, and internal-external control. *Journal of Personality, 40*(3), 392-406. doi: [10.1111/j.1467-6494.1972.tb00069.x](https://doi.org/10.1111/j.1467-6494.1972.tb00069.x)
- Pierce, H. C. (2019). *Perception versus reality: Investigating the impact of talkativeness on children's credibility and reliability*. (Unpublished master's thesis). Victoria University of Wellington.

## References

- Platt, L. F., & Szoka, S. L. (2019). Endorsement of feminist beliefs, openness, and mindful acceptance as predictors of decreased transphobia. *Journal of Homosexuality*, 68(2), 1-18. doi: 10.1080/00918369.2019.1651109
- Poole, D. A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review*, 32(3), 165-180. doi: 10.1016/j.dr.2012.06.007
- Poole, D. A., Dickinson, J. J., Brubacher, S. P., Liberty, A. E., & Kaake, A. M. (2014). Deficient cognitive control fuels children's exuberant false allegations. *Journal of Experimental Child Psychology*, 118, 101– 109. doi:10.1016/j.jecp.2013.08.013
- Poole, D., & Lindsay, D. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal Of Experimental Psychology: Applied*, 7(1), 27-50. doi: 10.1037/1076-898x.7.1.27
- Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology*, 27(6), 975-986. doi: 10.1037/0012-1649.27.6.975
- Porter, S., Yuille, J., & Lehman, D. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law And Human Behavior*, 23(5), 517-537. doi: 10.1023/a:1022344128649
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12(3), 427-441. doi: 10.1017/S0954579400003096

## References

- Powell, M. B., & Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: Effects of age, repetition, and retention interval across three question types. *Child Development, 67*(5), 1988-2004. doi: 10.1111/j.1467-8624.1996.tb01839.x
- Powers, C. J., & Bierman, K. L. (2013). The multifaceted impact of peer relations on aggressive–disruptive behavior in early elementary school. *Developmental Psychology, 49*(6), 1174-1186. doi: 10.1037/a0028400
- Pozzulo, J. D. (2017). *The young eyewitness: How well do children and adolescents describe and identify perpetrators?*. American Psychological Association. doi: 10.1037/14956-000
- Pozzulo, J. D., & Balfour, J. (2006). Children's and adults' eyewitness identification accuracy when a culprit changes his appearance: Comparing simultaneous and elimination lineup procedures. *Legal and Criminological Psychology, 11*(1), 25-34. doi: 10.1348/135532505X52626
- Pozzulo, J., & Dempsey, J. (2006). Biased Lineup Instructions: Examining the Effect of Pressure on Children's and Adults' Eyewitness Identification Accuracy1. *Journal Of Applied Social Psychology, 36*(6), 1381-1394. doi: 10.1111/j.0021-9029.2006.00064.x
- Pozzulo, J. D., & Dempsey, J. L. (2009). Witness factors and their influence on jurors' perceptions and verdicts. *Criminal Justice and Behavior, 36*(9), 923-934. doi: 10.1177/0093854809338450
- Pozzulo, J. D., Dempsey, J., Bruer, K., & Sheahan, C. (2012). The culprit in target-absent lineups: Understanding young children's false positive responding. *Journal of Police and Criminal Psychology, 27*(1), 55-62. doi: 10.1007/s11896-011-9089-8

## References

- Pozzulo, J. D., Dempsey, J., & Crescini, C. (2009). Preschoolers' person description and identification accuracy: A comparison of the simultaneous and elimination lineup procedures. *Journal of Applied Developmental Psychology, 30*(6), 667-676. doi: 10.1016/j.appdev.2009.01.004
- Pozzulo, J. D., Forth, A. E., & Bennell, C. (2018). *Forensic psychology* (4<sup>th</sup> edition). Pearson Canada.
- Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *Journal of Applied Psychology, 88*(2), 315-323. doi: 10.1037/0021-9010.88.2.315
- Pratt, C. (1990). On asking children-and adults-bizarre questions. *First Language, 10*(29), 167-175. doi: 10.1177/014272379001002905
- Presley, R., & Martin, R. P. (1994). Toward a structure of preschool temperament: Factor structure of the Temperament Assessment Battery for Children. *Journal of Personality, 62*(3), 415-448. doi: 10.1111/j.1467-6494.1994.tb00304.x
- Price, H., Connolly, D., & Gordon, H. (2016). Children who experienced a repeated event only appear less accurate in a second interview than those who experienced a unique event. *Law And Human Behavior, 40*(4), 362-373. doi: 10.1037/lhb0000194
- Principe, G. F., Greenhoot, A. F., & Ceci, S. (2013). Young Children's Eyewitness Memory. In T. J. Perfect., & D. S. Lindsay (Ed.). *The SAGE handbook of applied memory*. (pp. 363-653). SAGE.
- Purdy, T. (2001). *The role of the interviewer and type of interview on children's susceptibility to false memory* (Unpublished master's thesis). Emporia State University.

## References

- Pynoos, R. S., & Nader, K. (1989). Prevention of psychiatric morbidity in children after disaster. *Prevention of Mental Disorders, Alcohol, and Other Drug Use in Children and Adolescents. Office for Substance Abuse Prevention Monograph, 2*, 225-271.
- Quas, J. A., Bottoms, B. L., Haegerich, T. M., & Nysse-Carris, K. L. (2002). Effects of victim, defendant, and juror gender on decisions in child sexual assault cases 1. *Journal of Applied Social Psychology, 32*(10), 1993-2021. doi: 10.1111/j.1559-1816.2002.tb02061.x
- Quas, J.A., Goodman, G.S., Bidrose, S. et al. (1999). Emotion and memory: Children's long-term remembering, forgetting, and suggestibility. *Journal of Experimental Child Psychology, 72*, 235–270. doi: 10.1006/jecp.1999.2491
- Quas, J. A., Goodman, G. S., Ghetti, S., & Redlich, A. D. (2000). Questioning the child witness: What can we conclude from the research thus far?. *Trauma, Violence, & Abuse, 1*(3), 223-249. doi: 10.1177/1524838000001003002
- Quas, J. A., Malloy, L. C., Melinder, A., Goodman, G. S., D'Mello, M., & Schaaf, J. (2007). Developmental differences in the effects of repeated interviews and interviewer bias on young children's event memory and false reports. *Developmental Psychology, 43*(4), 823-837. doi: 10.1037/0012-1649.43.4.823
- R v. Russell Bishop (2018).
- Radford, L., Corral, S., Bradley, C., Fisher, H., Bassett, C., Howat, N., & Collishaw, S. (2011). Child abuse and neglect in the UK today.
- Raine, A. (2014). *The anatomy of violence: The biological roots of crime*. Vintage.

## References

- Ran, G., Zhang, Q., & Huang, H. (2018). Behavioral inhibition system and self-esteem as mediators between shyness and social anxiety. *Psychiatry Research, 270*, 568-573. doi: 10.1016/j.psychres.2018.10.017
- Rasmussen, A., & Berntsen, D. (2010). Personality traits and autobiographical memory: Openness is positively related to the experience and usage of recollections. *Memory, 18*(17), 774-786. doi: 10.1080/09658211.2010.514270
- Realmuto, G. M., & Wescoe, S. (1992). Agreement among professionals about a child's sexual abuse status: Interviews with sexually anatomically correct dolls as indicators of abuse. *Child abuse & neglect, 16*(5), 719-725. doi: 10.1016/0145-2134(92)90108-4
- Riemann, R., Grubich, C., Hempel, S., Mergl, S., & Richter, M. (1993). Personality and attitudes towards current political topics. *Personality and Individual Differences, 15*(3), 313-321. doi: 10.1016/0191-8869(93)90222-O
- Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological bulletin, 138*(1), 146-174. doi: 10.1037/a0025750
- Richards, M., & Eves, F. F. (1991). Personality, temperament and the cardiac defense response. *Personality and Individual Differences, 12*(10), 999-1007. doi: 10.1016/0191-8869(91)90030-F
- Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self-and peer report NEO-FFI scales. *Journal of Personality, 65*(3), 449-475. doi: 10.1111/j.1467-6494.1997.tb00324.x

## References

- Roberts, K., & Cameron, S. (2015). Observations From Canadian Practitioners About the Investigation and Prosecution of Crimes Involving Child and Adult Witnesses. *Journal Of Forensic Psychology Practice, 15*(1), 33-57. doi: 10.1080/15228932.2015.997611
- Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology, 18*(2), 189-202. doi: [10.1002/acp.957](https://doi.org/10.1002/acp.957)
- Roberts, K. P., & Powell, M. B. (2001). Describing individual incidents of sexual abuse: A review of research on the effects of multiple sources of information on children's reports. *Child Abuse & Neglect, 25*(12), 1643-1659. doi: [10.1016/S0145-2134\(01\)00290-3](https://doi.org/10.1016/S0145-2134(01)00290-3)
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological bulletin, 132*(1), 1-25. doi: 10.1037/0033-2909.132.1.1
- Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology, 82*(3), 416-425. doi: [10.1037/0021-9010.82.3.416](https://doi.org/10.1037/0021-9010.82.3.416)
- Robinson, M. D., Moeller, S. K., & Ode, S. (2010). Extraversion and reward-related processing: Probing incentive motivation in affective priming tasks. *Emotion, 10*(5), 615-626. doi: 10.1037/a0019173
- Roebbers, C. M., & Schneider, W. (2001). Individual differences in children's eyewitness recall: The influence of intelligence and shyness. *Applied Developmental Science, 5*(1), 9-20. doi.org/[10.1207/S1532480XADS0501\\_2](https://doi.org/10.1207/S1532480XADS0501_2)

## References

- Roediger, H., & McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *21*(4), 803-814. doi: 10.1037/0278-7393.21.4.803
- Roth, J., Dadds, M., & McAloon, J. (2004). Evaluation of a Puppet Interview to Measure Young Children's Self-Reports of Temperament. *Behavior Change*, *21*(1), 37-56. doi: 10.1375/bech.21.1.37.35973
- Rothbart, M. K. (1991). Temperament: A developmental framework. In J. Strelau & A. Angleitner (Eds.), *Perspectives on individual differences. Explorations in temperament: International perspectives on theory and measurement* (p. 61–74). Plenum Press. doi: [10.1007/978-1-4899-0643-4\\_5](https://doi.org/10.1007/978-1-4899-0643-4_5)
- Rothbart, M. K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science*, *16*(4), 207-212. doi: 10.1111/j.1467-8721.2007.00505.x
- Rothbart, M. K. (2012). Advances in temperament: History, concepts, and measures. In M. Zentner & R. L. Shiner (Eds.), *Handbook of temperament* (p. 3-20). Guilford Press.
- Rothbart, M. K., & Ahadi, S. A. (1994). Temperament and the development of personality. *Journal of Abnormal Psychology*, *103*(1), 55–66. doi: [10.1037/0021-843X.103.1.55](https://doi.org/10.1037/0021-843X.103.1.55)
- Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: origins and outcomes. *Journal of Personality and Social Psychology*, *78*(1), 122-135. doi: 10.1037/0022-3514.78.1.122
- Rothbart, M. K., Ahadi, S. A., & Hershey, K. L. (1994). Temperament and social behavior in childhood. *Merrill-Palmer Quarterly* (1982-), *21*-39.

## References

- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child development, 72*(5), 1394-1408. doi: [10.1111/1467-8624.00355](https://doi.org/10.1111/1467-8624.00355)
- Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (p. 99–166). John Wiley & Sons, Inc.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *APA science Vols. Temperament: Individual differences at the interface of biology and behavior* (p. 83–116). American Psychological Association. doi: [10.1037/10149-003](https://doi.org/10.1037/10149-003)
- Rossi, L., Benatti, F., Pesce, S., Oppo, A., & Avato, F. M. (2011). Suggestionabilità interrogativa: versione italiana del Bonn Test of Statement Suggestibility (BTSS). *Suggestionabilità interrogativa: versione italiana del Bonn Test of Statement Suggestibility (BTSS)*, 684-696. doi: [10.1400/214628](https://doi.org/10.1400/214628)
- Ross, D. F., Dunning, D., Tolia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury. *Law and Human Behavior, 14*(1), 5-23. doi: [10.1007/BF01055786](https://doi.org/10.1007/BF01055786)
- Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Replications and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology, 33*(2), 418-431. doi: [10.1111/j.1559-1816.2003.tb01903.x](https://doi.org/10.1111/j.1559-1816.2003.tb01903.x)
- Ross, D. F., Read, J. D., & Tolia, M. P. (1994). Adult eyewitness testimony: Current trends and developments. Cambridge University Press, doi: [10.1017/CBO9780511759192](https://doi.org/10.1017/CBO9780511759192)
- Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment, 41*(2), 150-156. doi: [10.1207/s15327752jpa4102\\_5](https://doi.org/10.1207/s15327752jpa4102_5)

## References

- Royal College of General Practitioners and National Society for the Prevention of Cruelty to Children. (2014) *Safeguarding children and young people: the RCGP/ NSPCC safeguarding children toolkit for general practice*. RCGP/NSPCC. Retrieved February 24<sup>th</sup>, 2021 at <https://www.rcgp.org.uk/clinical-and-research/clinical-resources/~media/Files/CIRC/Safeguarding-Children-Toolkit-2014/RCGP-NSPCC-Safeguarding-Children-Toolkit.ashx>.
- Rudasill, K. M., Prokasky, A., Tu, X., Frohn, S., Sirota, K., & Molfese, V. J. (2014). Parent vs. teacher ratings of children's shyness as predictors of language and attention skills. *Learning and Individual Differences*, 34, 57-62. [doi: 10.1016/j.lindif.2014.05.008](https://doi.org/10.1016/j.lindif.2014.05.008)
- Ruf, H. T., Schmidt, N. L., Lemery-Chalfant, K., & Goldsmith, H. H. (2008). Components of childhood impulsivity and inattention: Child, family, and genetic correlates. *European Journal of Developmental Science*, 2(1-2), 52–76. doi: 10.3233/DEV-2008-21205
- Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, 10(1-2), 117-125. doi: 10.1111/1468-2389.00198
- Salmon, K., & Pipe, M. E. (1997). Props and children's event reports: The impact of a 1-year delay. *Journal of Experimental Child Psychology*, 65(3), 261-292. doi: 10.1006/jecp.1996.2362
- Sandelands, L. E., Brockner, J., & Glynn, M. A. (1988). If at first you don't succeed, try, try again: Effects of persistence-performance contingencies, ego involvement, and self-esteem on task persistence. *Journal of Applied Psychology*, 73(2), 208. [doi: 10.1037/0021-9010.73.2.208](https://doi.org/10.1037/0021-9010.73.2.208)

## References

- Saucier, G., Georgiades, S., Tsaousis, I., & Goldberg, L. R. (2005). The Factor Structure of Greek Personality Adjectives. *Journal of Personality and Social Psychology, 88*(5), 856–875. doi: 10.1037/0022-3514.88.5.856
- Saucier, G., & Goldberg, L. R. (2003). The structure of personality attributes. In M. Barrick & A. M. Ryan (Eds.), *Personality and work* (pp. 1–29). Jossey-Bass-Pfeiffer. [HYPERLINK  
http://pages.uoregon.edu/gsaucier/Saucier.Golberg.Structure.Personality.Attributes.Personality-Work.2003.pdf](http://pages.uoregon.edu/gsaucier/Saucier.Golberg.Structure.Personality.Attributes.Personality-Work.2003.pdf)
- Saywitz, K. J., & Camparo, L. B. (2014). *Evidence-based child forensic interviewing: The developmental narrative elaboration interview*. Oxford University Press
- Saywitz, K., Camparo, L. B., & Romanoff, A. (2010). Interviewing children in custody cases: Implications of research and policy for practice. *Behavioral Sciences and the Law, 28*(4), 542-562. doi: 10.1002/bsl.945
- Saywitz, K. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology, 59*(5), 682-691. doi: 10.1037/0022-006X.59.5.682
- Saywitz, K. J., Larson, R. P., Hobbs, S. D., & Wells, C. R. (2015). Developing rapport with children in forensic interviews: Systematic review of experimental research. *Behavioral Sciences & the Law, 33*(4), 372-389. doi: 10.1002/bsl.2186
- Scanlon, E. M., & Ollendick, T. H. (1986). Children's assertive behavior: The reliability and validity of three self-report measures. *Child & Family Behavior Therapy, 7*(3), 9-22. doi: [10.1300/J019v07n03\\_02](https://doi.org/10.1300/J019v07n03_02)
- Schaaf, J. M., Alexander, K. W., & Goodman, G. S. (2008). Children's false memory and true

## References

- disclosure in the face of repeated questions. *Journal of Experimental Child Psychology*, 100 (3), 157–185. doi: 10.1016/j.jecp.2007.09.002
- Schacter, D. (2012). Adaptive constructive processes and the future of memory. *American Psychologist*, 67(8), 603-613. doi: 10.1037/a0029869
- Schacter, D. L., Kagan, J., & Leichtman, M. D. (1995). True and false memories in children and adults: A cognitive neuroscience perspective. *Psychology, Public Policy, and Law*, 1(2), 411. doi: [10.1037/1076-8971.1.2.411](https://doi.org/10.1037/1076-8971.1.2.411)
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum.
- Schmidt, S. (2004). Autobiographical memories for the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition*, 32(3), 443-454. doi: 10.3758/bf03195837
- Schneider, W., & Pressley, M. (1989). *Springer series in cognitive development. Memory development between 2 and 20*. Springer-Verlag Publishing. doi: [10.1007/978-1-4613-9717-5](https://doi.org/10.1007/978-1-4613-9717-5)
- Schreiber, N., Bellah, L. D., Martinez, Y., McLaurin, K. A., Strok, R., Garven, S., & Wood, J. M. (2006). Suggestive interviewing in the McMartin Preschool and Kelly Michaels daycare abuse cases: A case study. *Social influence*, 1(1), 16-47. doi: 10.1080/15534510500361739
- Schultz, D. P. & Schultz, S. E. (2005). *Theories of personality*. (8<sup>th</sup> edition). Thomson/Wadsworth.

## References

- Schutte, J. W., & Hosch, H. M. (1997). Gender differences in sexual assault verdicts: A meta-analysis. *Journal of Social Behavior and Personality*, 12(3), 759-772.
- Schwaba, T., Luhmann, M., Denissen, J. J. A., Chung, J. M., & Bleidorn, W. (2018). Openness to experience and culture-openness transactions across the lifespan. *Journal of Personality and Social Psychology*, 115(1), 118–136. doi: [10.1037/pspp0000150](https://doi.org/10.1037/pspp0000150)
- Scottish Government. (2014, May 14). *National Guidance for Child Protection in Scotland*. Retrieved 1 January 2021, from <http://www.scotland.gov.uk/Publications/2010/12/09134441/21>
- Scullin, M. H. (1997). *Development of a suggestibility scale for children* (Unpublished master's thesis). Cornell University.
- Shapiro, L. R. (2006). The effects of question type and eyewitness temperament on accuracy and quantity of recall for a simulated misdemeanor crime. *Emporia State Research Studies*, 43(1), 1-7.
- Shapiro, L. R., Blackford, C., & Chen, C. F. (2005). Eyewitness memory for a simulated misdemeanor crime: The role of age and temperament in suggestibility. *Applied Cognitive Psychology*, 19(3), 267-289. doi.org/[10.1002/acp.1089](https://doi.org/10.1002/acp.1089)
- Shaw, J. S. III. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology: Applied*, 2(2), 126-146. doi: [10.1037/1076-898X.2.2.126](https://doi.org/10.1037/1076-898X.2.2.126)
- Shiner, R. L. (1998). How shall we speak of children's personalities in middle childhood? A preliminary taxonomy. *Psychological Bulletin*, 124(3), 308-332. doi: [10.1037/0033-2909.124.3.308](https://doi.org/10.1037/0033-2909.124.3.308)

## References

- Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: Measurement, development, and consequences. *Journal of child psychology and psychiatry*, 44(1), 2-32. doi: 10.1111/1469-7610.00101
- Shiner, R. L., & DeYoung, C. G. (2013). *The structure of temperament and personality traits: A developmental perspective*. In P. Zelazo (Ed.), *Oxford handbook of developmental psychology*. Oxford University Press.
- Shute, V., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. Mit Press.
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, 12(3), 248-279. doi: 10.1177/1088868308319226
- Sibley, C. G., Osborne, D., & Duckitt, J. (2012). Personality and political orientation: Meta-analysis and test of a Threat-Constraint Model. *Journal of Research in Personality*, 46(6), 664-677. doi: 10.1016/j.jrp.2012.08.002
- Silverthorne, C. (2001). Leadership effectiveness and personality: A cross cultural evaluation. *Personality and Individual Differences*, 30(2), 303-309. doi: 10.1016/S0191-8869(00)00047-7
- Silvia, P. J., Nusbaum, E. C., Berg, C., Martin, C., & O'Connor, A. (2009). Openness to experience, plasticity, and creativity: Exploring lower-order, high-order, and interactive effects. *Journal of Research in Personality*, 43(6), 1087-1090. doi: 10.1016/j.jrp.2009.04.015
- Simonds, J. (2006). *The role of reward sensitivity and response execution in childhood extraversion* (Unpublished doctoral thesis). University of Oregon.

## References

- Singh, K., & Gudjonsson, G. (1992). Interrogative suggestibility among adolescent boys and its relationship with intelligence, memory, and cognitive set. *Journal Of Adolescence, 15*(2), 155-161. doi: 10.1016/0140-1971(92)90044-6
- Sjöberg, R. L., & Lindblad, F. (2002). Limited disclosure of sexual abuse in children whose experiences were documented by videotape. *American Journal of Psychiatry, 159*(2), 312-314. doi: 10.1176/appi.ajp.159.2.312
- Snow, P. C., Powell, M. B., & Murfett, R. (2009). Getting the story from child witnesses: Exploring the application of a story grammar framework. *Psychology, Crime & Law, 15*(6), 555-568. doi: 10.1080/10683160802409347
- Snyder, H. N. (2000, July). Sexual assault of young children as reported to law enforcement: Victim, incident, and offender characteristics (Bureau of Justice Statistics report, NCJ 182990). Washington, DC: U.S. Government Printing Office.
- Sommer, S., Reynolds, J. J., & Kehn, A. (2016). Mock juror perceptions of rape victims: Impact of case characteristics and individual differences. *Journal of Interpersonal Violence, 31*(17), 2847-2866. doi: 10.1177/0886260515581907
- Sosis, R. H. (1974). Internal-external control and the perception of responsibility of another for an accident. *Journal of Personality and Social Psychology, 30*(3), 393-399. doi: [10.1037/h0036893](https://doi.org/10.1037/h0036893)
- Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of research in personality, 43*(1), 84-90. doi: 10.1016/j.jrp.2008.10.002
- Spanos, N. P., Myers, B., Dubreuil, S. C., & Pawlak, A. E. (1992). The effects of polygraph evidence and eyewitness testimony on the beliefs and decisions of mock

## References

- jurors. *Imagination, Cognition and Personality*, 12(2), 103-113. doi: 10.2190/1AB2-3WLX-BFY1-8YCP
- Spooner, A. L., Evans, M. A., & Santos, R. (2005). Hidden shyness in children: Discrepancies between self-perceptions and the perceptions of parents and teachers. *Merrill-Palmer Quarterly (1982-)*, 437-466. doi: [10.1353/mpq.2005.0028](https://doi.org/10.1353/mpq.2005.0028)
- State of Utah v. Guard (2015).
- Stelmack, R. M. (1990). Biological bases of extraversion psychophysiological evidence. *Journal of Personality*, 58(1), 293-311. doi: 10.1111/j.1467-6494.1990.tb00917.x
- Stein, T., End, A., & Sterzer, P. (2014). Own-race and own-age biases facilitate visual awareness of faces under interocular suppression. *Frontiers In Human Neuroscience*, 8. doi: 10.3389/fnhum.2014.00582
- Stern, W. (1939). The psychology of testimony. *The Journal of Abnormal and Social Psychology*, 34(1), 3–20. doi: [10.1037/h0054144](https://doi.org/10.1037/h0054144)
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., & Hovav, M. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse & Neglect*, 21(11), 1133-1146. doi: 10.1016/S0145-2134(97)00071-9
- Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of applied psychology*, 86(5), 997-1005. doi: 10.1037/0021-9010.86.5.997

## References

- Steward, M. S., Steward, D. S., Farquhar, L., Myers, J. E., Reinhart, M., Welker, J., ... & Bruck, M. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development*, i-232. doi: 10.2307/1166205
- Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behaviour problems. *Journal of Clinical Child Psychology*, 22(1), 107-116. doi: [10.1207/s15374424jccp2201\\_11](https://doi.org/10.1207/s15374424jccp2201_11)
- Sullivan, C., St. George, S., & Stolzenberg, S. (2020). Misunderstanding children's reports of the mechanics of abuse, American-Psychology Law-Society, New Orleans, LA, 2020, March.
- Sutherland, R., & Hayne, H. (2001). Age-related changes in the misinformation effect. *Journal of experimental child psychology*, 79(4), 388-404. doi: 10.1006/jecp.2000.2610
- Swerdlow-Freed, D. H (2018, July 29). *The importance of questioning children in a developmentally-sensitive fashion*. Retrieved 7 January 2020, from <https://www.drswerdlow-freed.com/the-importance-of-questioning-children-in-a-developmentally-sensitive-fashion/>
- Talwar, V., & Crossman, A. M. (2012). Children's lies and their detection: Implications for child witness testimony. *Developmental Review*, 32(4), 337-359. doi: 10.1016/j.dr.2012.06.004
- Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development*, 26(5), 436-444. doi: 10.1080/01650250143000373

## References

- Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child development, 79*(4), 866-881.
- Tarter, R. E. (2002). Etiology of adolescent substance abuse: A developmental perspective. *American Journal on Addictions, 11*(3), 171-191. doi: 10.1080/10550490290087965
- Telford, R. M., Telford, R. D., Cochrane, T., Cunningham, R. B., Olive, L. S., & Davey, R. (2016). The influence of sport club participation on physical activity, fitness and body fat during childhood and adolescence: the LOOK longitudinal study. *Journal of Science and Medicine in Sport, 19*(5), 400-406. doi: 10.1016/j.jsams.2015.04.008
- Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K. J., Segal, N. L., & Rich, S. (1988). Personality similarity in twins reared apart and together. *Journal of Personality and Social Psychology, 54*(6), 1031–1039. doi: [10.1037/0022-3514.54.6.1031](https://doi.org/10.1037/0022-3514.54.6.1031)
- Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology, 47*(4), 1065-1077. doi: [10.1037/a0023273](https://doi.org/10.1037/a0023273)
- Thomas, A., & Chess, S. (1977). *Temperament and development*. Brunner/Mazel.
- Thomas, A., & Chess, S. (1986). The New York Longitudinal Study: From infancy to early adult life. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuities, and challenges* (pp. 39- 52). Erlbaum.
- Thomas, R. I. (2018, July 6). *Submission to the Commission for Justice in Wales on the provision of Registered Intermediary assistance for Vulnerable Witnesses*. Welsh Government. <https://gov.wales/sites/default/files/publications/2018-07/Submission-from-registered-intermediary-assistance-for-vulnerable-witnesses.pdf>

## References

- Tobey, A. E., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse & Neglect, 16*(6), 779-796. doi: 10.1016/0145-2134(92)90081-2
- Townsend, C., & Rheingold, A. A. (2013). Estimating a child sexual abuse prevalence rate for practitioners: A review of child sexual abuse prevalence studies. *Charleston, SC: Darkness to Light.*
- Trobst, K. K., Herbst, J. H., Masters III, H. L., & Costa Jr, P. T. (2002). Personality pathways to unsafe sex: Personality, condom use, and HIV risk behaviors. *Journal of Research in Personality, 36*(2), 117-133. doi: 10.1006/jrpe.2001.2334
- Tubb, A. V., Wood, J. M., & Hosch, H. M. (1999). Effects of suggestive interviewing and indirect evidence on child credibility in a sexual abuse case. *Journal of Applied Social Psychology, 29*(6), 1111-1127. doi: 10.1111/j.1559-1816.1999.tb02031.x
- Tucker, J. S., Elliott, M. N., & Klein, D. J. (2006). Social control of health behavior: Associations with conscientiousness and neuroticism. *Personality and Social Psychology Bulletin, 32*(9), 1143-1152. doi: 10.1177/0146167206289728
- Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review Of Psychology, 53*(1), 1-25. doi: 10.1146/annurev.psych.53.100901.135114
- Turan, B., Guo, J., Boggiano, M. M., & Bedgood, D. (2014). Dominant, cold, avoidant, and lonely: Basal testosterone as a biological marker for an interpersonal style. *Journal of Research in Personality, 50*, 84-89. doi: 10.1016/j.jrp.2014.03.008
- Ullman, S. (2007). A 10-Year Update of "Review and Critique of Empirical Studies of Rape Avoidance". *Criminal Justice And Behavior, 34*(3), 411-429. doi: 10.1177/0093854806297117

## References

- U.S. v. Mitchell (2010).
- Valentine, T., Darling, S., & Memon, A. (2007). Do strict rules and moving images increase the reliability of sequential identification procedures?. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 21(7), 933-949. doi: 10.1002/acp.1306
- Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, 44(4), 671-703. doi: 10.1080/14640749208401305
- Vallano, J. P., & Schreiber Compo, N. (2015). Rapport-building with cooperative witnesses and criminal suspects: A theoretical and empirical review. *Psychology, Public Policy, and Law*, 21(1), 85-99. doi: 10.1037/law0000035
- VanVoorhis, C. W., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, 3(2), 43-50.
- Ventura, M., Shute, V., & Zhao, W. (2013). The relationship between video game use and a performance-based measure of persistence. *Computers & Education*, 60(1), 52-58. doi: 10.1016/j.compedu.2012.07.003
- Vernon, P. A., Villani, V. C., Vickers, L. C., & Harris, J. A. (2008). A behavioral genetic investigation of the Dark Triad and the Big 5. *Personality and Individual Differences*, 44(2), 445-452. doi: 10.1016/j.paid.2007.09.007
- Virkkunen, M., Goldman, D., Nielsen, D. A., & Linnoila, M. (1995). Low brain serotonin turnover rate (low CSF 5-HIAA) and impulsive violence. *Journal of Psychiatry and Neuroscience*, 20(4), 271-275.

## References

- Vize, C. E., Lynam, D. R., Lamkin, J., Miller, J. D., & Pardini, D. (2016). Identifying essential features of juvenile psychopathy in the prediction of later antisocial behavior: Is there an additive, synergistic, or curvilinear role for fearless dominance?. *Clinical Psychological Science, 4*(3), 572-590. doi: 10.1177/2167702615622384
- Volbrecht, M. M., Lemery-Chalfant, K., Aksan, N., Zahn-Waxler, C., & Goldsmith, H. H. (2007). Examining the familial link between positive affect and empathy development in the second year. *The Journal of Genetic Psychology, 168*(2), 105-130. doi: 10.3200/GNTP.168.2.105-130
- Volpini, L., Melis, M., Petralia, S. & Rosenberg, M.D. (2016). Measuring children's suggestibility in forensic interviews. *Journal of Forensic Sciences, 61*(1), 104–108. doi: 10.1111/1556-4029.12987
- Vu, M., Murrie, D., Gonzalez, V., & Jobe, J. (2006). Listening to girls and boys talk about girls' physical activity behaviors. *Health Education and Behavior, 33*(1), 81-96. doi: 10.1177/1090198105282443
- Wade, K., Garry, M., Don Read, J., & Lindsay, D. (2002). A picture is worth a thousand lies: Using false photographs to create false childhood memories. *Psychonomic Bulletin & Review, 9*(3), 597-603. doi: 10.3758/bf03196318
- Walker, S., Berthelsen, D. C., & Irving, K. A. (2001). Temperament and peer acceptance in early childhood: Sex and social status differences. *Child Study Journal, 31*(3), 177-192.
- Walsh, W. A., Jones, L. M., Cross, T. P., & Lippert, T. (2010). Prosecuting child sexual abuse: The importance of evidence type. *Crime & Delinquency, 56*(3), 436-454. doi: 10.1177/0011128708320484

## References

- Walton, K. E., & Roberts, B. W. (2004). On the relationship between substance use and personality traits: Abstainers are not maladjusted. *Journal of Research in Personality, 38*(6), 515-535. doi: 10.1016/j.jrp.2004.01.002
- Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science, 23*(2), 169-177. doi: 10.1177/0956797611420575
- Ward, R., & Loftus, E. (1985). Eyewitness Performance in Different Psychological Types. *The Journal Of General Psychology, 112*(2), 191-200. doi: 10.1080/00221309.1985.9711003
- Warren, A., Hulse-Trotter, K., & Tubbs, E. C. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior, 15*(3), 273-285. [doi: 10.1007/BF01061713](https://doi.org/10.1007/BF01061713)
- Waterhouse, G. (2016). *Investigating the forensic interviewing of children: Multiple interviews and social support* (Unpublished doctoral dissertation). London South Bank University.
- Waterhouse, G. F., Ridley, A. M., Bull, R., La Rooy, D. J., & Wilcock, R. (2019). Mapping repeated interviews. *Journal of Police and Criminal Psychology, 34*(4), 392-409. [doi: 10.1007/s11896-018-9288-7](https://doi.org/10.1007/s11896-018-9288-7)
- Waterman, A. H., & Blades, M. (2011). Helping children correctly say "I don't know" to unanswerable questions. *Journal of Experimental Psychology: Applied, 17*(4), 396–405. [doi: 10.1037/a0026150](https://doi.org/10.1037/a0026150)
- Waterman, A., Blades, M., & Spencer, C. (2000). Do children try to answer nonsensical questions?. *British Journal Of Developmental Psychology, 18*(2), 211-225. doi: 10.1348/026151000165652

## References

- Waterman, A. H., Blades, M., & Spencer, C. (2001). Interviewing children and adults: The effect of question format on the tendency to speculate. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 15(5), 521-531. doi: 10.1002/acp.741
- Waterman, A. H., Blades, M., & Spencer, C. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's 'don't know' responses. *British Journal of Developmental Psychology*, 22(3), 335-348. doi: 10.1348/0261510041552710
- Waterman, C., & Foss-Goodman, D. (1984). Child molesting: Variables relating to attribution of fault to victims, offenders, and 112 nonparticipating parents. *The Journal of Sex Research*, 20(4), 329–349. doi: 10.1080/00224498409551231
- Weisberg, Y. J., C. G. De Young, and J. B. Hirsch. (2011). Gender Differences in Personality Across the Ten Aspects of the Big Five. *Frontiers in Psychology*, 33, 1–11. doi: 10.3389/fpsyg.2011.00178
- Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54(1), 277-295. doi: 10.1146/annurev.psych.54.101601.145028
- Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological psychology*, 89(1), 137-147. doi: 10.1016/j.biopsycho.2011.10.002
- Wilcock, R., Crane, L., Hobson, Z., Nash, G., Kirke-Smith, M., & Henry, L. A. (2018). Supporting child witnesses during identification lineups: Exploring the effectiveness of registered intermediaries. *Applied Cognitive Psychology*, 32(3), 367-375. doi: [10.1002/acp.3412](https://doi.org/10.1002/acp.3412)

## References

- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103-128. doi: 10.1016/0010-0277(83)90004-5
- Wixted, J. T., Mickes, L., Clark, S. E., Gronlund, S. D., & Roediger, H. L. III. (2015). Initial eyewitness confidence reliably predicts eyewitness identification accuracy. *American Psychologist*, *70*(6), 515–526. doi: 10.1037/a0039510
- Wheatcroft, J. M. (2002). *Some influences of courtroom questioning styles on eyewitness confidence and accuracy*. (Unpublished Thesis). University of Liverpool.
- Wheatcroft, J. M., Wagstaff, G. F., & Kebbell, M. R. (2004). The influence of courtroom questioning style on actual and perceived eyewitness confidence and accuracy. *Legal and Criminological Psychology*, *9*(1), 83-101. doi: 10.1348/135532504322776870
- Wheatcroft, J. M., & Woods, S. (2010). Effectiveness of witness preparation and cross-examination non-directive and directive leading question styles on witness accuracy and confidence. *The International Journal of Evidence & Proof*, *14*(3), 187-207. doi: 10.1350/ijep.2010.14.3.353
- Wolfman, M., Brown, D., & Jose, P. (2018). The use of visual aids in forensic interviews with children. *Journal of Applied Research in Memory and Cognition*, *7*(4), 587-596. doi: 10.1016/j.jarmac.2018.06.004
- Woodruff, G., & Premack, D. (1978). Does the chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, *4*, 515-526. doi: 10.1017/S0140525X00076512
- Wright, D. B., Hanoteau, F., Parkinson, C., & Tatham, A. (2010). Perceptions about memory reliability and honesty for children of 3 to 18 years old. *Legal and Criminological Psychology*, *15*(2), 195-207. doi: 10.1348/135532508X400347

## References

- Xu, X., Mar, R. A., & Peterson, J. B. (2013). Does cultural exposure partially explain the association between personality and political orientation?. *Personality and Social Psychology Bulletin*, 39(11), 1497-1517. doi: 10.1177/0146167213499235
- Xu, X., Plaks, J. E., & Peterson, J. B. (2016). From dispositions to goals to ideology: Toward a synthesis of personality and social psychological approaches to political orientation. *Social and Personality Psychology Compass*, 10(5), 267-280. doi: 10.1111/spc3.12248
- Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K., Ostendorf, F., Angleitner, A., Riemann, R., Spinath, F. M., Livesley, W. J., & Jang, K. L. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, 90(6), 987-998. doi: [10.1037/0022-3514.90.6.987](https://doi.org/10.1037/0022-3514.90.6.987)
- Yang, J., Cao, Z., Xu, X., & Chen, G. (2012). The amygdala is involved in affective priming effect for fearful faces. *Brain and cognition*, 80(1), 15-22. doi: 10.1016/j.bandc.2012.04.005
- Yang, Y., Raine, A., Colletti, P., Toga, A. W., & Narr, K. L. (2010). Morphological alterations in the prefrontal cortex and the amygdala in unsuccessful psychopaths. *Journal of Abnormal Psychology*, 119(3), 546-554. doi: 10.1037/a0019611
- Yang, Y., Raine, A., Narr, K. L., Colletti, P., & Toga, A. W. (2009). Localization of deformations within the amygdala in individuals with psychopathy. *Archives of general psychiatry*, 66(9), 986-994. doi: 10.1001/archgenpsychiatry.2009.110

## References

- Yap, M. B. H., Pilkington, P. D., Ryan, S. M., & Jorm, A. F. (2014). Parental factors associated with depression and anxiety in young people: A systematic review and meta-analysis. *Journal of Affective Disorders, 156*, 8-23. doi: 10.1016/j.jad.2013.11.007
- Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of psychosomatic research, 65*(5), 445-451. doi: 10.1016/j.jpsychores.2008.03.013
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459-482.
- Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review, 16*(2), 116-142. doi: 10.1177/1088868311418987
- Young, K., Powell, M. & Dudgeon, P. (2003). Individual differences in children's suggestibility: A comparison between intellectually disabled and mainstream samples. *Personality & Individual Differences, 35*, 31-49. doi: 10.1016/S0191-8869(02)00138-1
- Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of consulting and clinical psychology, 68*(6), 1038. doi: [10.1037/0022-006X.68.6.1038](https://doi.org/10.1037/0022-006X.68.6.1038)
- Zajac, R., & Karageorge, A. (2009). The wildcard: A simple technique for improving children's target-absent lineup performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23*(3), 358-368. doi: 10.1002/acp.1511

## References

- Zajac, R., Westera, N., & Kaladelfos, A. (2018). The “good old days” of courtroom questioning: Changes in the format of child cross-examination questions over 60 years. *Child Maltreatment, 23*(2), 186-195. doi: 10.1177/1077559517733815
- Zhiyan, T., & Singer, J. L. (1997). Daydreaming styles, emotionality and the big five personality dimensions. *Imagination, Cognition and Personality, 16*(4), 399-414. doi: 10.2190/ATEH-96EV-EXYX-2ADB
- Ziegler, M., Cengia, A., Mussel, P., & Gerstorf, D. (2015). Openness as a buffer against cognitive decline: The Openness-Fluid-Crystallized-Intelligence (OFCI) model applied to late adulthood. *Psychology and Aging, 30*(3), 573–588. doi: [10.1037/a0039493](https://doi.org/10.1037/a0039493)
- Zietsch, B. P., Verweij, K. J. H., Bailey, J. M., Wright, M. J., & Martin, N. G. (2010). Genetic and environmental influences on risky sexual behaviour and its relationship with personality. *Behavior Genetics, 40*(1), 12-21. doi: 0.1007/s10519-009-9300-1
- Zhou, Q., Lengua, L. J., & Wang, Y. (2009). The relations of temperament reactivity and effortful control to children’s adjustment problems in China and the United States. *Developmental Psychology, 45*(3), 724–739. doi: [10.1037/a0013776](https://doi.org/10.1037/a0013776)

## Appendix A: Parent Letters

---

### INFORMATION SHEET FOR PARENTS/CARERS

#### *How interviewing techniques and temperament affect children as eyewitnesses and the decision making of jurors. Study Two*

**Researcher:** Ben Cotterill

*(Supervisors: Dr Kathy Charles, Dr Faye Skelton and Dr Rory MacLean).*

Your child is being invited to take part in a research study. Before you decide whether your child would like to take part, it is important for you to understand why the research is being done and what it will involve. Please ask if there is anything that is not clear or if you would like more information.

#### **What is the purpose of the study?**

This project is investigating how interviewing techniques and temperament affect children (4 to 8-years-old) as eyewitnesses.

Faulty eyewitness accounts are the leading cause of wrongful arrests, and evidence shows that children make poorer eyewitnesses than adults, both remembering less and performing worse at facial recognition (for a review, see Ceci & Bruck, 1995). Clearly, it must be carefully considered how much weight should be placed on the importance of eyewitness testimony. If testimonies are believed without question, it could result in wrongful arrests, but if they are not believed at all, then guilty and dangerous offenders could walk free. This is especially a point of concern because of the large number of children who become involved in the legal system every year. In Scotland, children were asked to give testimony in a criminal court 4, 297 times during 2017 (Crown Office and Procurator Fiscal Service, 2018).

In Scotland, investigators are trained to use the Stepwise Protocol when interviewing children. All children are interviewed with the same system variables and there is little flexibility to the process. Unfortunately, some children are more suggestible than others, meaning that some testimonies are contaminated because of suggestive questioning (e.g. Luther, Snook, Barron, & Lamb, 2014; Roberts & Cameron, 2015), and cases can get dismissed because the testimony is no longer suitable evidence for court.

Ornstein, Shapiro, Clubb, Follmer, and Baker-Ward (1997) theorised that particular elements of temperament affect the perception of eyewitnesses as they witness events take place (activity level, emotionality and persistence), while other elements (adaptability, shyness and distractibility) impact on their performance during forensic interviews. Further research has investigated more closely how some of these characteristics affect memory recall (e.g. Bruck & Melnyk, 2005; Chae & Ceci, 2005; Pozzulo, Coplan & Wilson, 2005; Roebbers & Schneider, 2001; Shapiro, Blackford & Chen, 2005; Shapiro, 2006), but little have focused specifically on how these may influence the eyewitness performance of children aged between 4 and 8-years-old. Having a better understanding of how temperament may impact eyewitness performance will provide some direction regarding what system variables are most likely to increase description and identification accuracy.

The project expects to find that the negative effects (i.e. poorer accuracy and confidence) associated with leading questioning, closed-ended questions and repeated questioning will be significantly heightened when they are applied to children with certain temperament qualities (e.g. shyness, distractibility). This would demonstrate the importance of taking a child's personality into account when undertaking forensic interviews with children.

**Why has my child been chosen?**

As part of this research we are looking for about 2000 children between the ages of 4 and 8 to take part in the study. Your child has been chosen to take part because they fit these criteria.

**Does my child have to take part?**

No. This is an entirely voluntary project. If you choose not to participate it will not affect you or your child in any way. If you give your consent for your child to participate, you will be asked to sign a consent form. Even if you give consent, it will be made clear to your child that it is their decision to take part and they can withdraw at any time and without giving a reason.

**What will my child be asked to do if we agree to take part?**

We will take every care to reduce to a minimum disruption to the school routine. First, children will be asked to answer questions about their temperament traits. Second, the children will be asked to watch a short video that depicts a non-violent bike theft. Third, the children will be asked questions about what they remember from the video. The questions will follow the instructions of Scotland's Stepwise Protocol, by initially asking for free recall and then using follow-up questions. Children will receive follow-up questions that are in one of four conditions (open-ended, closed-ended, leading or repeated).

Every effort will be made to ensure that the research sessions are as enjoyable and relaxed as possible for the children. It is also hoped that the children will learn something about memory and police procedures. The total testing time should not exceed 20 minutes per child.

**Who will run the research sessions?**

The researcher, Ben Cotterill, is a PhD student at Edinburgh Napier University and has PVG clearance for working with children, and will meet with and test the children taking part. He has just completed a MSc at the University of York studying Forensic Psychology, during which he conducted similar research to this, and has experience both with working with children and as an interviewer. Please note that the research has permission from both Edinburgh Napier University's Ethics Committee and the local council.

**Will all my child's details and the assessment results be kept confidential?**

Yes. All the information about participants in this study will be kept confidential and data will be anonymous and stored securely.

We are not at liberty to provide personal results or individualised feedback on any measures taken in the study.

Appendices

Please note, in the unlikely event that the child makes any disclosures of being at risk to harm, then the researcher will have a legal duty to report this to yourself or a teacher.

Contact:

If you require any further information or have any questions about this study, please do not hesitate to contact the researcher, Ben Cotterill ([ben.cotterill@napier.ac.uk](mailto:ben.cotterill@napier.ac.uk); [REDACTED]) or the research's supervisor, Kathy Charles ([k.charles@napier.ac.uk](mailto:k.charles@napier.ac.uk)).

Alternatively, if you would like to contact an independent advisor, please feel free to contact Cedric English ([C.English@napier.ac.uk](mailto:C.English@napier.ac.uk); [REDACTED]).

**PARENT CONSENT FORM**

**How interviewing techniques and temperament affect children as eyewitnesses and the decision making of jurors**

Researcher's name: Ben Cotterill

*Please circle either YES or NO*

- 1. Have you read the parents' information sheet? YES / NO
- 2. Do you understand that your child is free to withdraw from the study:
  - At any time? YES / NO
  - Without having to give a reason?
- 3. Do you understand the time requirements of the study? YES / NO
- 4. Do you agree to let your child take part in this study? YES / NO
- 5. Do you agree to let your child be audio recorded during the study? (*Your child's audio will not be identifiable*) YES / NO

---

**6. PARENT**

Name of Child (BLOCK LETTERS): .....

Signature of Parent: ..... Date: .....

Name of Parent (BLOCK LETTERS): .....

---

## Appendix B: TATC Item Examples

---

1. I'm quiet with people I don't know (**Shyness**).
2. When I start a puzzle, I don't stop until it's finished, even if it takes a long time (**Persistence**).
3. I run and jump when I play (**Activity**).
4. I adjust quickly to new rules at home or in the classroom (**Adaptability**).
5. I cry or scream when I'm upset (**Emotionality**).
6. I daydream when I should be listening to someone (**Distractibility**).
7. I don't like talking to children I don't know (**Shyness**).
8. I don't move on to a new task before completing the old one even if it's a challenge (**Persistence**).
9. I don't like being asked to perform in front of people (**Shyness**).
10. I get along easily with new routines (**Adaptability**).
11. I cry, yell, or stomp my foot when I'm angry (**Emotionality**).
12. I get distracted when being read a book or told a story (**Distractibility**).
13. I usually run everywhere, rather than walk (**Activity**).
14. If the shops don't have my favourite sweet, I'm quite happy to just get something else instead (**Adaptability**).
15. I prefer outdoor, running around games than indoor (**Activity**).
16. I don't mind getting my hair cut and nails clipped (**Adaptability**).
17. When I'm upset, it's easy to cheer me up (**Emotionality**).
18. When I'm promised something by an adult, I constantly remind them (**Persistence**).
19. When I'm at the park, I don't like playing with kids I don't know (**Shyness**).
20. I prefer friends to visit my house, rather than visiting their house (**Adaptability**).
21. I get out of bed as soon as I wake up (**Activity**).
22. I cry or scream when I get in trouble (**Emotionality**).
23. I get distracted from a task if there's noises outside the window (**Distractibility**).
24. When I'm talking about my favourite things, I don't like the subject being changed (**Persistence**).
25. I get more upset than other children when sharing toys (**Emotionality**).
26. I feel shy when people sing Happy Birthday to me (**Shyness**).

## Appendices

27. When I cry, I don't cry for a very long time (**Emotionality**).
28. I'm easily distracted by siblings or friends when I should be doing chores or homework (**Distractibility**).
29. I always have lots of energy (**Activity**).
30. I doesn't feel shy when the head teacher speaks to me (**Shyness**).
31. I don't find it hard to pay attention (**Distractibility**).
32. When a toy or game is too difficult, I quickly turn to another activity (**Persistence**).
33. I feel uncomfortable when joining a new class (**Adaptability**).
34. I sit quietly throughout long movies, without getting distracted (**Distractibility**).
35. I like sitting still and playing quietly, instead of running around (**Activity**).
36. When told to wear certain clothes, I wear them without protest even if I don't like them (**Emotionality**).
37. When I'm having fun with a toy or enjoying a game, it's easy getting me to stop (**Persistence**).
38. I do not adjust well when plans are cancelled (**Adaptability**).
39. When I sit, I usually sit still, rather than swing my arms or legs (**Activity**).
40. I enjoy speaking in front of the class (**Shyness**).
41. During class, I am not side-tracked easily (**Distractibility**).
42. I don't like playing games where you run around a lot (**Activity**).
43. I talk and laugh with people I don't know (**Shyness**).
44. I find change to be difficult (**Adaptability**).
45. When told to stop playing, I don't protest very strongly (**Emotionality**).
46. If an adult is busy and I want their attention, I will go away instead of keeping after the adult (**Persistence**).
47. I don't find it hard to concentrate when someone is telling a story (**Distractibility**).
48. If I want something from an adult, it's easy to get me to stop talking about it (**Persistence**).

## Appendix C: Parent Form

### Temperament Assessment Tool for Children – Parent Form

Thank you for agreeing to your child taking part in this study. If you could please complete the following questions about your child, it would be greatly appreciated.

The results will be compared to the answers your child says during the study so the researcher can check both your responses and their responses are in agreement.

Questions	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. My child is quiet with adults he/she doesn't know.					
2. When my child starts a puzzle, they don't stop until it's finished, even if it takes a long time.					
3. My child runs and jumps when they play.					
4. My child adjusts quickly to new rules at home or in the classroom.					
5. My child cries or screams when they're upset.					
6. My child daydreams when he/she should be listening to someone.					
7. My child doesn't like talking to children they don't know.					
8. My child doesn't move on to a new task before completing the old one even if it's a challenge.					
9. My child doesn't like being asked to perform in front of people.					
10. My child goes along easily with new routines.					
11. My child cries, yells, or stomps their foot when they're angry.					
12. My child gets distracted when					

being read a book or told a story.					
13. My child usually runs everywhere, rather than walk.					
14. If the shops don't have my child's favourite sweet, he/she is quite happy to just get something else instead.					
15. My child prefers outdoor, running around games.					
16. My child doesn't mind getting their hair cut or their nails clipped.					
17. When my child is upset, it's easy to cheer them up.					
18. When my child is promised something by an adult, he/she constantly reminds them.					
19. When my child's at the park, he/she doesn't like playing with kids he/she doesn't know.					
20. My child prefers friends to visit our house, rather than visiting their house.					
21. My child gets out of bed as soon as they wake up.					
22. My child cries or screams when they get in trouble.					
23. My child gets distracted from their task if there's noises outside the window.					
24. When my child is talking about their favourite things, they don't like the subject being changed.					
25. My child gets more upset than other children when sharing toys.					

Appendices

26. My child feels shy when people sing Happy Birthday to him/her.					
27. When my child cries, they don't cry for a very long time.					
28. My child is easily distracted by siblings or friends when he/she should be doing chores or homework.					
29. My child always has lots of energy.					
30. My child doesn't feel shy when the head teacher speaks to them.					
31. My child doesn't find it hard to pay attention.					
32. When a toy or game is too difficult, my child quickly turns to another activity.					
33. My child feels uncomfortable when joining a new class.					
34. My child sits quietly throughout long movies, without getting distracted.					
35. My child likes sitting still and playing quietly, instead of running around.					
36. When told to wear certain clothes, my child wears them without protest even if they don't like them.					
37. When my child is having fun with a toy or enjoying a game, it's easy getting them to stop.					
38. My child does not adjust well when plans are cancelled.					
39. When my child sits, they usually sit still, rather than swing arms or legs.					

Appendices

<p><b>40. My child enjoys speaking in front of the class.</b></p>					
<p><b>41. When playing with a friend, my child's friend gets bored of the game before my child.</b></p>					
<p><b>42. My child doesn't like to play games where you run around a lot.</b></p>					
<p><b>43. My child talks and laughs with people that he/she doesn't know.</b></p>					
<p><b>44. My child finds change to be difficult.</b></p>					
<p><b>45. When told to stop playing, my child doesn't protest very strongly.</b></p>					
<p><b>46. If an adult is busy and my child wants their attention, they will go away instead of keeping after the adult.</b></p>					
<p><b>47. My child doesn't find it hard to concentrate when someone is telling them a story.</b></p>					
<p><b>48. When my child wants something from an adult, it's easy to get them to stop talking about it.</b></p>					

## Appendix D: TATC Correlation Matrix

---

**Table 8.1**

*TATC correlation matrix.*

	1	2	3	4	5	6
1. Activity	--					
2. Adaptability	-.14*	--				
3. Distractibility	-.08	.04	--			
4. Emotionality	-.01	.46**	-.06	--		
5. Persistence	-.20*	.02	.04	-.02	--	
6. Shyness	.04	-.20*	-.04	.13	-.12	--

*Note.* \*  $p > .05$ .

\*\*  $p > .001$ .

## Appendix E: Open-Ended Interview Format

---

### Open-ended Interview Format

#### Introduction

When certain things happen, people sometimes need to call the police for help. It is the job of the police to find out information about what happened. To do this, they might go about asking if anyone saw anything. If people did see anything, they are supposed to tell the police what they saw. You just watched a movie and I was told that something happened to a child's bike in the movie. If you had seen this, the police might ask you about it and you would have to tell them everything you saw. My job is important because I want to find out how much children can remember about events that they see. I'm going to write down everything you say so try not to talk too fast. Okay, are you ready? (*Proceed only if the child indicates they're ready.*)

I don't know what happened in the movie because I didn't watch it, so I want you to tell me everything you REALLY, REALLY remember about what happened to the bike. I will be asking you a lot of questions. If you don't understand a question, just say "I don't understand." It's really important that I get told only the truth. So, before we begin, I want to make sure that you understand how important it is to tell the truth.

If I say that my shoes are red (*or green*), is that true or not true? (*Wait for an answer.*)

It would be not true because my shoes are really [blue/back/etc.]. And if I say I am sitting down right now, would that be true or not true? (*Wait for an answer.*)

It would be true, because you can see I really am sitting down. I see that you understand what telling the truth means. It is important that you only tell me the truth today. You should only tell me about things that really happened. Also, if I ask question and you don't remember or you're not sure about the answer, just tell me, "I don't know." So, if I ask you, "What is my dog's name?", what would you say? (*Wait for an answer.*)

You don't know, do you? And if I say things that are wrong, you should tell me. So if I said that you are a 2-year-old girl, what would you say? (*Wait for an answer.*)

Exactly, don't be afraid to tell me if I say something that's wrong.

### **Rapport Building**

Now I want to get to know you better. Tell me about something you like to do. (*Wait for child to respond.*)

Tell me more about [activity].

### **Practice Interview**

Tell me everything that happened on [summer holiday/Christmas/activity mentioned earlier/yesterday]. (*Wait for an answer.*)

Think hard about [activity or event] and tell me what happened on that day from the time you got up that morning until [some portion of the event mentioned by the child in response to the previous question]. (*Wait for an answer.*)

And then what happened? (*Wait for an answer.*)

Tell me everything that happened after [some portion of the event mentioned by the child] until you went to bed that night. (*Wait for an answer.*)

### **Free Narrative**

Now that I know you a little better, I want to talk about why you're here today. I understand that you saw something happen to the bike. Tell me everything that happened from the beginning to the end. (*Let the child list ALL the features before you go back through the list to ask for elaboration.*)

What else happened with the bike? (*Repeat the question until the child's list seems exhausted.*)

Was there anything else that happened to the bike? (*Wait for an answer.*)

You said [mention EACH feature mentioned by the child, one at a time. The goal is to have the child elaborate as much as possible without leading the child at all or without using closed-ended questions]. Tell me more about that. (*Wait for an answer.*)

**Prompts (if necessary)**

Then what happened? (*Wait for an answer.*)

Think back and tell me everything that happened from [some preceding event mentioned by the child] until the bike was taken. (*Wait for an answer.*)

Tell me more about [crime feature/person feature mentioned by the child]. (*Wait for an answer.*)

*(When the child has told you all that she or he can, proceed to Specific Questions and ask about those items not already mentioned.)*

**Specific Questions**

Thank you for your help so far. I have some more questions for you. I want you to think about what happened with the bike again. For these questions, I need you to tell me only what you REALLY, REALLY remember. If you don't remember or you are not sure about your answers, just tell me, "I don't know". (*If, any time during the interview, the child responds with, "I think" or "Maybe", then remind them it's really important they only report what they REALLY, REALLY remember.*)

1). Tell me who the bike belonged to.

2). Tell me the colour of the bike.

3). Tell me what type of bike it was.

4). Tell me about the handlebars of the bike.

5a). I need to know a little more about what happened between the girl and the boy. Tell me what the girl was doing when they boy first came up to her.

>> *If the child tells you sitting, swinging or standing, but doesn't mention singing skip to #6.*

5b). What song was someone singing?

6a). Tell me what the boy did when he first saw the bike.

>> *If the child says took it, ask if the boy did anything before that. If the child still mentions nothing about touching it skip to #7.*

>> *If the child says touched it skip to #6b.*

6b). Where did the boy touch it?

7). Tell me if the boy and girl were arguing about anything.

8). Tell me if the boy touched the girl.

9). Tell me what the girl did when the boy first tried to walk off with the bike.

>> *If the child gives answer indicating they struggled and the girl moved the bike the other side of the bench, skip to #10.*

>> *If the child provides answer for only one part, then ask the other part.*

A. Tell me if there was a struggle over the bike.

B. Tell me if the girl tried to move the bike.

10a). Tell me what the girl did after she moved the bike from the boy.

>> *If the child tells you sitting, swinging or standing, but doesn't mention singing skip to #11.*

>> *If the child mentions singing skip to #10b.*

10b). What song did she sing?

>> *If the child claims not to know the song mentioned, just say "That's okay".*

11). Tell me what the boy did when the girl wouldn't let him use the bike?

>> *If the child gives answer indicating the boy slit his throat with his finger and came back to ride away with the bike skip to #12.*

>> *If the child provides answer for only one part, then ask the other part.*

A. Tell me if the boy made any gesture.

B. Tell me how the boy took the bike.

12). Tell me, did the boy call the girl a name?

*Appendices*

13). Tell me, did the girl do anything when the boy rode away on the bike?

>> *If the child gives answer indicating the girl was angry and stomped her foot, skip to #14.*

>> *If the child provides answer for only one part, then ask the other part.*

A. Tell me how the girl reacted.

B. Tell me if the girl did any actions.

14). Tell if anyone came up to the girl when she was upset.

15). Tell me what the girl's father/mother did when they saw the girl was upset?

*For this question, refer to #14 above. Use "mother" if the child indicated the mother comforted the victim. If the child indicated the father comforted the victim, use the word "father".*

16). Tell me where the crime took place.

17). I need to know a little more about the boy. Tell me the boy's name.

18). Tell me, what colour was the boy's hair?

19). Tell me, what style of hair did the boy have?

>> *If the child gives answer indicating short, skip to #19.*

>> *If the child responds medium or long then ask second question ONLY.*

>> *If the child responds IDK or doesn't respond, then ask BOTH sets:*

A. Tell me about the length of the hair.

B. Tell me how the boy wore the hair.

20). Tell me the girl's name.

21). Tell me, what colour was the girl's hair?

22). Tell me, how long was the girl's hair and what style was it?

23). Tell me which of the children was taller.

24). Tell me which of the children was older.

25). Tell me, what colour hair did the father (mother) have?

26). Now I need to know about the clothes that the children were wearing. Tell me everything the boy was wearing.

>> *Do NOT ask about an item that the child has already answered, either correctly or incorrectly.*

- A. Tell me about the kind of shirt he was wearing.
- B. Tell me about what pants he was wearing.
- C. Tell me what kind of shoes he was wearing.
- D. Tell me what colour of shoes he was wearing.

27). Tell me everything the girl was wearing.

>> *Do NOT ask about an item that the child has already answered, either correctly or incorrectly.*

- A. Tell me about the colour of her shirt.
- B. Tell me about her pants.
- C. Tell me about what kind of shoes she was wearing.
- D. Tell me about what colour of shoes she was wearing.

28). Tell me everything the father (mother) was wearing.

>> *Do NOT ask about an item that the child has already answered, either correctly or incorrectly.*

- A. Tell me about the kind of shirt they were wearing.
- B. Tell me what colour of shirt they were wearing.
- C. Tell me about their pants.
- D. Tell me about what colour of shoes she was wearing.

29a). Tell me, was anyone wearing a watch?

>> If the child indicates who was wearing a watch even the wrong one, skip to #29b.

29b). Tell me what kind of watch it was.

Thank you for helping me.

## Appendix F: Closed-Ended Interview Format

---

### Closed-ended Interview Format

#### Introduction

When certain things happen, people sometimes need to call the police for help. It is the job of the police to find out information about what happened. To do this, they might go about asking if anyone saw anything. If people did see anything, they are supposed to tell the police what they saw. You just watched a movie and I was told that something happened to a child's bike in the movie. If you had seen this, the police might ask you about it and you would have to tell them everything you saw. My job is important because I want to find out how much children can remember about events that they see. I'm going to write down everything you say so try not to talk too fast. Okay, are you ready? (*Proceed only if the child indicates they're ready.*)

I don't know what happened in the movie because I didn't watch it, so I want you to tell me everything you REALLY, REALLY remember about what happened to the bike. I will be asking you a lot of questions. If you don't understand a question, just say "I don't understand." It's really important that I get told only the truth. So, before we begin, I want to make sure that you understand how important it is to tell the truth.

If I say that my shoes are red (*or green*), is that true or not true? (*Wait for an answer.*)

It would be not true because my shoes are really [blue/back/etc.]. And if I say I am sitting down right now, would that be true or not true? (*Wait for an answer.*)

It would be true, because you can see I really am sitting down. I see that you understand what telling the truth means. It is important that you only tell me the truth today. You should only tell me about things that really happened. Also, if I ask question and you don't remember or you're not sure about the answer, just tell me, "I don't know." So, if I ask you, "What is my dog's name?", what would you say? (*Wait for an answer.*)

You don't know, do you? And if I say things that are wrong, you should tell me. So if I said that you are a 2-year-old girl, what would you say? (*Wait for an answer.*)

Exactly, don't be afraid to tell me if I say something that's wrong.

### **Rapport Building**

Now I want to get to know you better. Tell me about something you like to do. (*Wait for child to respond.*)

Tell me more about [activity].

### **Practice Interview**

Tell me everything that happened on [summer holiday/Christmas/activity mentioned earlier/yesterday]. (*Wait for an answer.*)

Think hard about [activity or event] and tell me what happened on that day from the time you got up that morning until [some portion of the event mentioned by the child in response to the previous question]. (*Wait for an answer.*)

And then what happened? (*Wait for an answer.*)

Tell me everything that happened after [some portion of the event mentioned by the child] until you went to bed that night. (*Wait for an answer.*)

### **Free Narrative**

Now that I know you a little better, I want to talk about why you're here today. I understand that you saw something happen to the bike. Tell me everything that happened from the beginning to the end. (*Let the child list ALL the features before you go back through the list to ask for elaboration.*)

What else happened with the bike? (*Repeat the question until the child's list seems exhausted.*)

Was there anything else that happened to the bike? (*Wait for an answer.*)

You said [mention EACH feature mentioned by the child, one at a time. The goal is to have the child elaborate as much as possible without leading the child at all or without using closed-ended questions]. Tell me more about that. (*Wait for an answer.*)

**Prompts (if necessary)**

Then what happened? (*Wait for an answer.*)

Think back and tell me everything that happened from [some preceding event mentioned by the child] until the bike was taken. (*Wait for an answer.*)

Tell me more about [crime feature/person feature mentioned by the child]. (*Wait for an answer.*)

*(When the child has told you all that she or he can, proceed to Specific Questions and ask about those items not already mentioned.)*

**Specific Questions**

Thank you for your help so far. I have some more questions for you. I want you to think about what happened with the bike again. For these questions, I need you to tell me only what you REALLY, REALLY remember. If you don't remember or you are not sure about your answers, just tell me, "I don't know". (*If, any time during the interview, the child responds with, "I think" or "Maybe", then remind them it's really important they only report what they REALLY, REALLY remember.*)

1). Did the bike belong to the girl or the boy?

2). Was the bike red or green?

3). Was the bike a mountain bike or a road bike?

4). Were the handlebars curved or curved?

5a). I need to know a little more about what happened between the girl and the boy when the boy first came up to the girl. Was she sitting on a bench, or was she swinging on a swing?

>> *If the child tells you sitting, swinging or standing, but doesn't mention singing skip to #6.*

>> *If the child mentions singing skip to #5b.*

5b). Did she sing "Itsy Bitsy Spider" or "Bingo"?

6a). PLQ: When the boy first touched the bike, did the boy touch it with his hand or did he kick it with his foot?

6b). Did he touch the tyres, or did he touch the seat and handlebars?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

>> *If the child claims not to know where the bike was touched, just say "That's okay".*

7). I want to know if the boy and girl were arguing about anything. Did the girl want the boy to leave the bike alone, or did the girl want the boy to sit somewhere else?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

8). Did the boy pat the girl's head, or did the boy punch the girl's arm?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

9a). I want to know about what the girl did when the boy first tried to walk off with the bike. Did the girl struggle over the bike, or did the girl kick over the bench?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

(b). Did the girl push the bike under the bench, or did the girl move the bike to the other side of the bench?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

10a). After the girl moved the bike, did she sit back down or did she keep standing?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

10b). Did she sing "Row, Row Your Boat", or did she sing "Mary Had a Little Lamb"?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

>> *If the child claims not to know the song mentioned, just say "That's okay".*

11a). Did the boy pretend to slit his throat with his finger, or did the boy stick out his tongue at the girl?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

*Appendices*

11b). Did the boy knock the bike down and walk away, or did the boy grab the bike and ride away?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

12). Did the boy call her "a stupid jerk", or did he call her "a dumb baby"?"?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

13a). Did the girl get sad after the boy stole the bike, or did the girl get angry?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

13b). Did the girl stomp her foot, or did the girl begin to cry?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

14). Did her mother come up to her when the girl was upset, or did her father come up to her?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

15). Tell me what the girl's father/mother did when they saw the girl was upset? Did s/he put a hand on the girl's shoulder, or did s/he go running after the boy?

*For this question, refer to #14 above. Use "mother" if the child indicated the mother comforted the victim. If the child indicated the father comforted the victim, use the word "father".*

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

16). Did the crime happen at the park, or did it happen at the zoo?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

17). Was the boy's name Frankie, or was the boy's name Ashley?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

18). Was the boy's hair light blonde, or was the boy's hair dark brown?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ*

19a). Did the boy have long hair, or did the boy have short hair?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

19b). Did he wear it in a pony-tail, or did he wear it down?

*Appendices*

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

20). Was the girl's name Frankie, or was the girl's name Ashley?

*Be sure to get clarification if the child responds Yes to both.*

21). Was the girl's hair light blonde, or was her hair dark brown?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

22a). Was the girl's hair below her shoulder, or was it above her shoulders?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

22b). Was her hair straight, or was her hair curly?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

23). Was the girl taller than the boy, or was the boy taller?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

24). Was the boy older than the girl, or was the girl older?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

25). Was the father's (mother's) hair light blonde, or was the father's (mother's) hair dark brown?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

26a). Now I need to know about the clothes that the children were wearing. Was the boy wearing a t-shirt, or was he wearing a collared shirt?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

26b). Was he wearing shorts, or was he wearing jeans?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

27b). Was he wearing hiking boots, or was he wearing trainers?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

27c). Was he wearing blue shoes, or was he wearing brown shoes?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

27a). Was the girl wearing a white shirt, or was she wearing a pink shirt??

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

*Appendices*

27b). Was she wearing shorts, or was she wearing jeans?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

27c). Was she wearing trainers, or was she wearing sandals?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

27d). Was she wearing blue shoes, or was she wearing white shoes?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

28a). Was the father (mother) wearing a collared shirt, or was s/he wearing a t-shirt??

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

28b). Was s/he wearing a white shirt, or was s/he wearing a brown shirt?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

28c). Was s/he wearing jeans, or was s/he wearing shorts?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

28d). Was she wearing white shoes, or was she wearing brown shoes?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

29a). Was the boy wearing a watch, or was the girl wearing a watch?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

>>*If the child responds IDK or doesn't respond, or answer "Yes" to PLQ ,ask set B:*

29b). Was it a small gold watch, or was it a big black watch?

*Be sure to get clarification if the child responds Yes to both PLQ and NLQ.*

Thank you for helping me.

---

---

## Appendix G: Misleading Interview Format

---

### Misleading Interview Format

#### Introduction

When certain things happen, people sometimes need to call the police for help. It is the job of the police to find out information about what happened. To do this, they might go about asking if anyone saw anything. If people did see anything, they are supposed to tell the police what they saw. You just watched a movie and I was told that something happened to a child's bike in the movie. If you had seen this, the police might ask you about it and you would have to tell them everything you saw. My job is important because I want to find out how much children can remember about events that they see. I'm going to write down everything you say so try not to talk too fast. Okay, are you ready? *(Proceed only if the child indicates they're ready.)*

I don't know what happened in the movie because I didn't watch it, so I want you to tell me everything you REALLY, REALLY remember about what happened to the bike. I will be asking you a lot of questions. If you don't understand a question, just say "I don't understand." It's really important that I get told only the truth. So, before we begin, I want to make sure that you understand how important it is to tell the truth.

If I say that my shoes are red *(or green)*, is that true or not true? *(Wait for an answer.)*

It would be not true because my shoes are really [blue/back/etc.]. And if I say I am sitting down right now, would that be true or not true? *(Wait for an answer.)*

It would be true, because you can see I really am sitting down. I see that you understand what telling the truth means. It is important that you only tell me the truth today. You should only tell me about things that really happened. Also, if I ask question and you don't remember or you're not sure about the answer, just tell me, "I don't know." So, if I ask you, "What is my dog's name?", what would you say? *(Wait for an answer.)*

You don't know, do you? And if I say things that are wrong, you should tell me. So if I said that you are a 2-year-old girl, what would you say? *(Wait for an answer.)*

Exactly, don't be afraid to tell me if I say something that's wrong.

### **Rapport Building**

Now I want to get to know you better. Tell me about something you like to do. (*Wait for child to respond.*)

Tell me more about [activity].

### **Practice Interview**

Tell me everything that happened on [summer holiday/Christmas/activity mentioned earlier/yesterday]. (*Wait for an answer.*)

Think hard about [activity or event] and tell me what happened on that day from the time you got up that morning until [some portion of the event mentioned by the child in response to the previous question]. (*Wait for an answer.*)

And then what happened? (*Wait for an answer.*)

Tell me everything that happened after [some portion of the event mentioned by the child] until you went to bed that night. (*Wait for an answer.*)

### **Free Narrative**

Now that I know you a little better, I want to talk about why you're here today. I understand that you saw something happen to the bike. Tell me everything that happened from the beginning to the end. (*Let the child list ALL the features before you go back through the list to ask for elaboration.*)

What else happened with the bike? (*Repeat the question until the child's list seems exhausted.*)

Was there anything else that happened to the bike? (*Wait for an answer.*)

You said [mention EACH feature mentioned by the child, one at a time. The goal is to have the child elaborate as much as possible without leading the child at all or without using closed-ended questions]. Tell me more about that. (*Wait for an answer.*)

**Prompts (if necessary)**

Then what happened? (*Wait for an answer.*)

Think back and tell me everything that happened from [some preceding event mentioned by the child] until the bike was taken. (*Wait for an answer.*)

Tell me more about [crime feature/person feature mentioned by the child]. (*Wait for an answer.*)

*(When the child has told you all that she or he can, proceed to Specific Questions and ask about those items not already mentioned.)*

**Specific Questions**

Thank you for your help so far. I have some more questions for you. I want you to think about what happened with the bike again. For these questions, I need you to tell me only what you REALLY, REALLY remember. If you don't remember or you are not sure about your answers, just tell me, "I don't know". (*If, any time during the interview, the child responds with, "I think" or "Maybe", then remind them it's really important they only report what they REALLY, REALLY remember.*)

1). I need to know more about the bike that was taken. The bike belonged to the boy, didn't it?

2). Was the colour of the bike red?

3). The bike was a road bike, wasn't it?

4). The handlebars on the bike were curved, weren't they?

5a). I need to know a little more about what happened between the girl and the boy. Think about what the girl was doing when the boy first came up to her. The girl was swinging on a swing, wasn't she?

5b). Think about the song she sang. It was "Itsy Bitsy Spider", wasn't it?

6a). Now I want you to think about when the boy first saw the bike. Did the boy kick it with his foot?

6b). He touched the tyres, didn't he?

7). The boy and girl were arguing about something. Was it that the girl wanted the boy to sit somewhere else?

8). The boy patted the girl's head, didn't he?

9a). I want you to think about when the boy first tried to walk off with the bike. Did the girl kick over the bench?

9b). The girl pushed the bike under the bench, didn't she?

10a). Did she then keep standing?

10b). Did she then start singing "Mary Had a Little Lamb"?

11a). Think about when the girl wouldn't let the boy use the bike. Did the boy stick out his tongue at the girl?

11b). The boy knocked the bike down and walked away, didn't he?

12). The boy called the girl "a dumb baby", didn't he?

13a). Now think about when the boy rode away on the bike. The girl got sad, didn't she?

13b). Did the girl begin to cry?

14). Who came up to the girl when she was upset -it was her mother, wasn't it?

15). When the mother saw the girl was upset, did she go running after the boy?

16). The crime took place at the park, didn't it?

*Appendices*

- 17). I need to know a little more about the boy. The boy's name was Ashley, wasn't it?
- 18). Was the boy's hair light blond?
- 19a). Think about the boy's hair. The boy's hair was past his shoulders, wasn't it?
- 19b). The boy's long hair was in a ponytail, wasn't it?
- 20). Now I need to know about the girl. Was the girl's name Frankie?
- 21). Her hair was dark brown, wasn't it?
- 22a). It was below her shoulders, wasn't it?
- 22b). She had curly hair, right?
- 23). The girl was taller than the boy, wasn't she?
- 24). Was the girl a few years older than the boy?
- 25). Now let's talk about the mother. Was the mother's hair light blonde?
- 26a). Now I need to know about the clothes that the children were wearing. Tell me everything the boy was wearing. First, think about what the boy was wearing. Was he wearing a collared shirt?
- 26b) He was wearing shorts, wasn't he?
- 26c). He was wearing trainers, wasn't he?
- 26d). They were blue, weren't they?
- 27a). Now think about what the girl was wearing. The girl was wearing a pink shirt, wasn't she?

*Appendices*

27b). She was wearing shorts, wasn't she?

27c). Was the girl wearing sandals?

27d). Were her shoes blue?

28a). Let's talk about what the father (mother) was wearing. S/he was wearing a t-shirt, wasn't s/he?

28b). It was a brown shirt, wasn't it?

28c). He was also wearing shorts, wasn't he?

28d). His shoes were white, weren't they?

29). The girl was wearing a watch, wasn't she?

>> *If the child responds Yes, IDK, or doesn't respond, ask:*

Was it a small gold watch?

Thank you for helping me.

---

## Appendix H: Children's Debrief

---

### DEBRIEF FOR PARTICIPANTS

*How interviewing techniques and temperament affect children as eyewitnesses and the decision making of jurors.*

**Researcher:** Ben Cotterill

*(Supervisors: Dr Kathy Charles, Dr Faye Skelton and Dr Rory MacLean).*

Thank you for taking part.

The aim of this study was to understand how personality may be measured in young children.

I asked you some questions that I hope will tell me a little about your personality. I then showed you a video, and asked you some questions on what you remembered from the video. I did this because I wanted to know how you may respond to the questions based on your personality.

If you require any further information or have any questions about this study, please do not hesitate to contact the researcher, Ben Cotterill ([ben.cotterill@napier.ac.uk](mailto:ben.cotterill@napier.ac.uk); [REDACTED]) or the research's supervisor, Kathy Charles ([k.charles@napier.ac.uk](mailto:k.charles@napier.ac.uk)). You may also ask your teacher, head teacher, or parents to contact me on your behalf. They have all my contact information, and then I could answer your questions through them.

Alternatively, if you would like to contact an independent advisor, please feel free to contact Cedric English ([C.English@napier.ac.uk](mailto:C.English@napier.ac.uk); [REDACTED])

---

## Appendix I: Coding Scheme

---

### Coding Scheme

#### Instructions:

Values reflect the completeness of the answer. Total points will be added for both correct and incorrect responses. For example, children can be correct about the type of shoes an actor is wearing, but be wrong about the colour of the shoes. In this instance, the children will receive both correct and incorrect points. The coding for correct/incorrect point values will be assigned as follows: Elaborated credit (3 points) will be given when the children gives correct/incorrect information with details (e.g., curly, brown hair) and/or dialogue, Complete (2 points) will be given when the children give correct/incorrect information alone, and Partial (1 point) will be given when they give some correct/incorrect information (e.g., detail or correct dialogue). Credit for elaboration does not have to be given at the time the correct/incorrect response is given. For example, children can provide this information during free recall, or in response to the follow-up questions. Please note that if certain details are covered during free recall then certain follow-up questions might not be asked later in the interview. The order of the coding scheme, therefore, is only a guide, and the order of reported details may differ between interviews. The point value of 0 will be assigned when the children do not respond with an answer, or if they indicate they do not know the answer.

If children initially give the wrong answer but later, during the interview, correct themselves, it is considered a spontaneous correction and will be coded as if the wrong answer had not been given.

Both the gist dialogue and the verbatim statement will be scored as elaboration whether it is given as a direct quote or given indirectly. For example, it is not necessary for the children to remember the exactly wording of 'Get back here, stop, stop, that's my bike, come back, come back.' It would be acceptable for children to state that the girl yelled to come back, or to state 'She said, come back with my bike.'

#### **Bike Owner (Central Crime)**

**Correct 2 points** – Complete: Girl's bike.

**Error 2 points** – Anyone else's bike (e.g., boy or dad).

**Colour of the Bike (Central Crime)**

**Correct 3 points** – Elaboration: Black bike with Trek (or silver writing) written on it.

**Correct 2 points** – Complete: Black.

**Correct 1 point** – Partial: Dark or blackish.

**Error 3 points** – Elaboration: Incorrect colour and incorrect detail.

**Error 2 points** – Complete: Incorrect colour.

**Error 1 point** – Partial: Incorrect detail.

**Model and Handlebars of the Bike (Central Crime)**

**Correct 3 points** – Elaboration: Mountain bike with straight handlebars and/or water bottle holder.

**Correct 2 points** – Complete: Mountain bike, or straight handlebars.

**Correct 1 point** – Partial: For both girls and boys or for bigger children (no stabilizers).

**Error 3 points** – Elaboration: Incorrect features.

**Error 2 points** – Complete: Curved handlebars or road bike.

**Error 1 points** – Partial: Incorrect feature.

**Girl Prior to Boy's Arrival (Peripheral Crime)**

**Correct 3 points** – Elaboration: Sitting on a bench.

**Correct 2 points** – Complete: Sitting.

**Correct 1 point** – Partial: In a picnic area or the bike was next to her.

**Error 2 points** – Complete: Swinging, or any other answer.

**First Song (Peripheral Crime)**

**Correct 3 points** – Elaboration: Song singing (i.e., Bingo) and clapping hands.

**Correct 2 points** – Complete: Singing and clapping.

**Correct 1 point** – Partial: Singing or clapping.

**Error 3 points** – Elaborate: Incorrect song.

**Error 2 points** – Complete: Incorrect action.

**Boy First Saw the Bike (Peripheral Crime)**

**Correct 3 points** – Elaboration: Touched it on seat and handlebars and looked at tyres.

## *Appendices*

**Correct 2 points** – Touched it, grabbed it, wheeled it, or tried to take it away.

**Error 3 points** – Elaboration: Wrong action (e.g., kicked the bike).

**Error 2 points** – Complete: Touched in wrong place (i.e., not seat and handlebars).

### **What They Were Arguing About (Central Crime)**

**Correct 3 points** – Elaboration: Use of bike and dialogue on why she said no (e.g., the boy wanted to use the bike but the girl said no because her dad will get mad).

**Correct 2 points** – Complete: The use of bike or an implication of wanting to take it.

**Correct 1 point** – Partial: The bike.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue.

**Error 2 points** – Incorrect answer (e.g., she wanted the boy to sit somewhere or argued about something else).

**Error 1 point** – Incorrect dialogue.

### **Boy Touched the Girl (Peripheral Crime)**

**Correct 3 points** – Elaboration: Punched her in the left arm or with right hand.

**Correct 2 points** – Complete: Punched her in the arm.

**Correct 1 point** – Partial: Punched, slugged, touched, or hit her.

**Error 2 points** – Complete: Hit her anywhere else other than the arm.

### **Girl's Initial Response (Central Crime)**

**Correct 3 points** – Elaboration: Struggled and some form of dialogue (e.g., boy said nothing bad will happen).

**Correct 2 points** – Complete: Struggled, wrestled, grabbed bike away, tried to take the bike.

**Correct 1 point** – Partial: Pulled on bike, took it back.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue.

**Error 2 points** – Complete: Anything that does not include a struggle.

**Error 1 point** – Partial: Incorrect dialogue.

### **Moved Bike (Peripheral Crime)**

**Correct 3 points** – Elaboration: Moved the bike to the right side of the bench.

**Correct 2 points** – Complete: Moved bike to the other side of her (bench).

**Correct 1 point** – Partial: Moved bike.

**Error 2 points** – Complete: Anything that does not include moving the bike.

**Second Song (Peripheral Crime)**

**Correct 3 points** – Elaboration: Song singing (i.e., Row Row Row Your Boat).

**Correct 2 points** – Complete: Singing.

**Error 3 points** – Elaboration: Incorrect song.

**Error 2 points** – Complete: Incorrect action.

**Boy's Response (Peripheral Crime)**

**Correct 2 points** – Complete: Slit throat while walking away.

**Error 2 points** – Complete: Anything that does not include that specific action.

**Took Bike (Central Crime)**

**Correct 3 points** – Elaboration: Rode off to the right side of the screen, gave dialogue (e.g., you should have let me use it before) and snuck up from behind.

**Correct 2 points** – Complete: Grabbed bike and rode away.

**Correct 1 point** – Partial: Used bike or borrowed bike or sneaked up.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue.

**Error 2 points** – Complete: Gave the bike back or other incorrect information.

**Error 1 point** – Partial: Incorrect dialogue.

**Boy Called Girl a Name (Peripheral Crime)**

**Correct 3 points** – Elaboration: Stupid jerk.

**Correct 2 points** – Complete: Stupid or jerk.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue.

**Error 2 points** – Complete: Any other name.

**Error 1 point** – Partial: Incorrect dialogue.

**Girl's Emotional Response (Peripheral Crime)**

**Correct 3 points** – Elaboration: Angry and give dialogue (e.g., the girl said she hates the boy) and/or action (i.e., the girl stomped her foot).

**Correct 2 points** – Complete: Angry, mad.

**Correct 1 point** – Partial: Upset, or gives dialogue.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue and/or incorrect action (e.g., the girl cried).

## *Appendices*

**Error 2 points** – Complete: Sad or anything that does not imply anger.

**Error 1 point** – Partial: Incorrect dialogue.

### **Who Came Up to The Girl (Peripheral Crime)**

**Correct 3 points** – Elaboration: Father and gives dialogue (e.g., dad said don't worry, we'll get it back), or indicates father put his hand on shoulder (or around her) or they went to look for the bike.

**Correct 2 points** – Complete: Father.

**Correct 1 point** – Partial: A man.

**Error 3 points** – Elaboration: Incorrect answer and incorrect dialogue.

**Error 2 points** – Complete: Any other person.

**Error 1 point** – Partial: Incorrect dialogue.

### **Where Crime Takes Place (Central Crime)**

**Correct 2 points** – Complete: Zoo.

**Error 2 points** – Complete: Any other place.

### **Boy's Name (Central Person)**

**Correct 2 points** – Complete: Frankie, Frank.

**Error 2 points** – Complete: Any other name.

### **Boy's Hair Colour (Central Person)**

**Correct 3 points** – Elaboration: Dark brown.

**Correct 2 points** – Complete: Brown or black.

**Correct 1 point** – Partial: Dark.

**Error 2 points** – Complete: Blonde or light.

### **Boy's Hair Length (Central Person)**

**Correct 3 points** – Elaboration: Gives length (i.e., to ears).

**Correct 2 points** – Complete: Short.

**Error 3 points** – Elaboration: Long and in a ponytail.

**Error 2 points** – Complete: Any length past the chin.

### **Girl's Name (Peripheral Person)**

## *Appendices*

**Correct 2 points** – Complete: Sport, kid, didn't say the name.

**Error 2 points** – Complete: Ashley, Frankie or any other name.

### **Girl's Hair (Peripheral Person)**

**Correct 3 points** – Elaboration: Blonde and give length (i.e., to chin or above shoulders), wavy.

**Correct 2 points** – Complete: Blonde, light blonde, yellow or length.

**Correct 1 point** – Partial: Light or short.

**Error 3 points** – Elaboration: Incorrect colour and incorrect detail.

**Error 2 points** – Complete: Incorrect colour.

**Error 1 point** – Partial: Incorrect detail (e.g. curly, past her shoulders).

### **Which Child Was Taller (Central Person)**

**Correct 3 points** – Elaboration: Boy and specify by 10 inches or 5.3 feet.

**Correct 2 points** – Complete: Boy.

**Error 3 points** – Elaboration: Girl and incorrect specification.

**Error 2 points** – Complete: Girl.

### **Which Child Was Older (Central Person)**

**Correct 3 points** – Elaboration: Boy and specify age range for boy 13-15 or girl 8-10.

**Correct 2 points** – Complete: Boy.

**Error 3 points** – Elaboration: Girl and incorrect specification.

**Error 2 points** – Complete: Girl.

### **Father's Hair Colour (Peripheral Person)**

**Correct 3 points** – Elaboration: Correct colour and receding hairline, moustache, short hair, glasses.

**Correct 2 points** – Complete: Black, dark brown, brown.

**Correct 1 point** – Partial: Dark or any one correct feature.

**Error 3 points** – Elaboration: Incorrect colour and incorrect feature.

**Error 2 points** – Complete: Incorrect colour.

**Error 1 point** – Incorrect feature.

### **Boy's Clothing (Central Person)**

**Correct 3 points** – Elaboration: Black with white lettering and wore jeans.

## *Appendices*

**Correct 2 points** – Complete: Black shirt.

**Correct 1 point** – Partial: Dark, wore jeans or white letters on shirt, but no colour of shirt.

**Error 3 points** – Elaboration: Two or more incorrect details.

**Error 2 points** – Complete: Incorrect colour.

**Error 1 point** – Partial: One incorrect item.

### **Boy's Shoes (Central Person)**

**Correct 3 points** – Elaboration: Hiking boots and brown.

**Correct 2 points** – Complete: Boots, hiking boots.

**Correct 1 point** – Partial: Brown.

**Error 3 points** – Elaboration: Incorrect type and colour of shoe.

**Error 2 points** – Complete: Incorrect type.

**Error 1 point** – Partial: Incorrect colour.

### **Girl's Clothing (Peripheral Person)**

**Correct 3 points** – Elaboration: Wore jeans and white t-shirt.

**Correct 2 points** – Complete: Jeans, blue jeans.

**Correct 1 point** – Partial: Pants or white t-shirt.

**Error 3 points** – Elaboration: Incorrect colour and incorrect item.

**Error 2 points** – Complete: Wore shorts.

**Error 1 point** – Partial: Incorrect colour or item.

### **Girl's Shoes (Peripheral Person)**

**Correct 3 points** – Elaboration: Trainers and white.

**Correct 2 points** – Complete: Trainers.

**Correct 1 point** – Partial: White.

**Error 3 points** – Elaboration: Incorrect colour and incorrect type.

**Error 2 points** – Complete: Incorrect type.

**Error 1 point** – Partial: Incorrect colour.

### **Dad's Clothing (Peripheral Person)**

**Correct 3 points** – Elaboration: White shirt with stripes or pocket and jeans.

**Correct 2 points** – Complete: White shirt.

*Appendices*

**Correct 1 point** – Partial: Light, wore jeans or stripes/pocket on shirt, but no colour of shirt.

**Error 3 points** – Elaboration: Two or more incorrect details.

**Error 2 points** – Complete: Incorrect colour.

**Error 1 point** – Partial: One incorrect item.

**Dad's Shoes (Peripheral Person)**

**Correct 3 points** – Elaboration: Shoes and black.

**Correct 2 points** – Complete: Shoes (i.e., not trainers).

**Correct 1 point** – Partial: Black.

**Error 3 points** – Elaboration: Incorrect type and colour.

**Error 2 points** – Complete: Incorrect type.

**Error 1 point** – Partial: Incorrect colour.

**Watch (Peripheral Person)**

**Correct 3 points** – Elaboration: Boy's and big and black

**Correct 2 points** – Complete: Boy's and big or black.

**Correct 1 point** – Partial: Boy's or big or black.

**Error 3 points** – Elaboration: Incorrect person and incorrect colour.

**Error 2 points** – Complete: Incorrect person.

**Error 1 point** – Partial: Incorrect colour.

**Total Correct Points** – 92

**Total Error Points** – 87

## Appendix J: Recall Scoring Sheet

### Recall Scoring Sheet

Item	Correct	Error
Bike Owner		
Bike Colour		
Bike Model and Handlebars		
Girl Prior to Boy's Arrival		
First Song		
Boy First Saw the Bike		
What They Were Arguing About		
Boy Touched the Girl		
Girl's Initial Response		
Moved Bike		
Second Song		
Boy's Response		
Took Bike		
Boy Called Girl a Name		
Girl's Emotional Response		
Who Came Up to The Girl		
Where Crime Takes Place		
Boy's Name		
Boy's Hair Colour		
Boy's Hair Length		
Girl's Name		
Girl's Hair		

*Appendices*

Which Child Was Taller		
Which Child Was Older		
Father's Hair Colour		
Boy's Clothing		
Boy's Shoes		
Girl's Clothing		
Girl's Shoes		
Dad's Clothing		
Dad's Shoes		
Watch		
Total		

---

## Appendix K: Study two correlation matrixes

**Table 8.2**

*Open-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	.09	.28*	.05	.11	-.07	.08
Errors	.16	.12	.26	-.05	-.09	-.08
Accuracy	-.15	-.03	-.20	-.07	.07	.09

Note. \*  $p > .05$ .

**Table 8.3**

*Closed-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.12	.003	.08	.05	-.13	.03
Errors	.06	-.08	.53**	-.03	.01	.07
Accuracy	-.13	.04	-.46**	.06	-.04	-.08

Note. \*  $p > .05$ .

\*\*  $p > .001$ .

**Table 8.4**

*Misleading condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.07	.37*	-.16	-.22	.48**	-.04
Errors	-.14	-.25*	.56**	.06	.10	.17
Accuracy	.12	.45**	-.55**	-.18	.13	-.20

Note. \*  $p > .05$ .

\*\*  $p > .001$ .

**Table 8.5**

*Young children: Open-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	.22	.30	.22	.10	-.25	.12
Errors	.11	.35*	.17	.03	-.27	-.05
Accuracy	-.04	-.24	-.03	-.04	.17	.07

Note. \*  $p > .05$ .

**Table 8.6**

*Young children: Closed-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.05	-.04	.22	.13	-.20	-.09
Errors	-.03	-.33	.40*	.30	.03	.19
Accuracy	-.05	.29	-.30	-.21	.11	-.25

Note. \*  $p > .05$ .

**Table 8.7**

*Young children: Misleading condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.08	.39*	-.02	-.34	.64**	-.10
Errors	-.24	-.28	.63**	.03	.34	.24
Accuracy	.22	.51*	-.62**	-.20	-.02	-.30

Note. \*  $p > .05$ .

\*\*  $p > .001$ .

**Table 8.8**

*Older children: Open-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	.10	.09	-.12	.25	.07	.06
Errors	-.22	.14	.38*	-.22	.10	-.20
Accuracy	.29	-.10	-.41*	.29	-.05	.23

\*\*  $p > .001$ .

**Table 8.9**

*Older children: Closed-ended condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.39*	.03	.09	-.15	-.02	.16
Errors	.29	.15	.61**	-.23	-.03	-.03
Accuracy	-.36*	-.16	-.52*	.17	.05	.05

Note. \*  $p > .05$ .

\*\*  $p > .001$ .

**Table 8.10**

*Older children: Misleading condition correlation matrix.*

	Activity	Adaptability	Distractibility	Emotionality	Persistence	Shyness
Correct	-.13	.17	-.36*	.04	.30	.06
Errors	-.03	-.22	.50*	.07	-.05	.10
Accuracy	.02	.30	-.55*	-.09	.14	-.10

*Note.* \*  $p > .05$ .

## **Appendix L: Differences between video mock-jurors and transcript-mock-jurors**

---

When using only the data from the mock-jurors who saw the videos ( $n = 15$ ), there were no significant differences on reliability, confidence, compelling-nature, bias, or shyness based upon interview condition or shyness levels ( $ps > .05$ ). There were also no significant interaction effects ( $ps > .05$ ).

There were no significant differences between those who saw the videos and those who read the transcripts on any of the dependent variables ( $ps > .05$ ).

## **Appendix M: Participant information sheet and consent form (Study Three)**

---

### **7. INFORMATION SHEET FOR PARTICIPANTS**

*How interviewing techniques and temperament affect children as eyewitnesses and the decision making of juror: Study three.*

**Researcher:** Ben Cotterill

*(Supervisors: Faye Skelton, Rory MacLean, and Kathy Charles,).*

You are being invited to take part in a research study. Before you decide whether you would like to take part, it is important for you to understand why the research is being done and what it will involve. Please ask if there is anything that is not clear or if you would like more information.

#### **What is the purpose of the study?**

This project is investigating how child temperament and interviewing techniques affect jurors' perceptions of witness credibility.

Eyewitness testimonies provide jurors with information that would not otherwise be available and thus are incredibly important to criminal trials. That said, memory is fallible and eyewitness testimony is the leading cause of wrongful arrests (Innocence Project, 2018).

If a child is the only eyewitness, it is especially vital to understand how their testimony is perceived by jurors. The perception of jurors over the credibility of a child witness may be affected by a number of factors. For example, Golding (2003) found that a crying child witness led to more guilty verdicts from jurors than a child who did not cry while giving testimony.

In some cases, child witnesses have been viewed as incredibly honest and as having no reason to lie (e.g., Nunez, Kehn, & Wright, 2011). In other cases, they have been perceived as having a poor memory and thus to be less reliable than adult witnesses (e.g., Bottoms & Goodman, 1994). The decision making of jurors is ultimately influenced by a number of factors. The objective of this study is to examine how the interviewing technique used and the temperament of the child affect the perceptions and decision making of mock jurors.

#### **Why have I been chosen?**

As part of this research we are looking for 60+ native English speakers over the age of 16 to take part in the study. You have been chosen to take part because you fit this criteria.

#### **Do I have to take part?**

No. This is an entirely voluntary project. If you give your consent to participate, you will be asked to sign a consent form. Even if you give consent, you can withdraw at any time and without giving a reason.

#### **What will I be asked to do if I agree to take part?**

## *Appendices*

If you agree to take part, you will be asked to read transcripts from three witness testimony videos from three separate witnesses. After each transcript, you will be asked to complete a questionnaire on how accurate, detailed, and confident you found the witness in the video to be. The questionnaire will be in the form of a 9-point Likert scale and you will respond based upon the degree to which you found the child to be accurate, reliable, or confident (e.g. (1) extremely inaccurate to (9) extremely accurate). The witnesses in the transcripts are 4 to 8-year-old children who are being interviewed over what they remember about a non-violent bike theft. Each of the interviews lasts for about seven minutes. The total testing time should not exceed 30 minutes. At the end of the session, you will be fully debriefed.

### **Who will run the research sessions?**

The researcher, Ben Cotterill, is a PhD student and Associate Lecturer at Edinburgh Napier University. He has completed a MSc in Applied Forensic Psychology at the University of York, during which he conducted similar research to this. He has experience both as an interviewer and with working with children. Please note that the research has permission from both Edinburgh Napier University's Ethics Committee and the local council.

### **Will all my details and the assessment results be kept confidential?**

Yes. All the information about participants in this study will be kept confidential and data will be anonymous and stored securely.

We are not at liberty to provide personal results or individualised feedback on any measures taken in the study.

### **Contact:**

If you require any further information or have any questions about this study, please do not hesitate to contact the researcher, Ben Cotterill ([ben.cotterill@napier.ac.uk](mailto:ben.cotterill@napier.ac.uk); XXXXXXXXXX) or the research's supervisors: Faky Skelton ([f.skelton@napier.ac.uk](mailto:f.skelton@napier.ac.uk)), Rory MacLean ([r.maclean@napier.ac.uk](mailto:r.maclean@napier.ac.uk)), Kathy Charles ([kathy.charles@ntu.ac.uk](mailto:kathy.charles@ntu.ac.uk)).

Alternatively, if you would like to contact an independent advisor, please feel free to contact Cedric English ([C.English@napier.ac.uk](mailto:C.English@napier.ac.uk)).

**PARTICIPANT CONSENT FORM**

**How interviewing techniques and temperament affect children  
as eyewitnesses and the decision making of jurors**

Researcher's name: Ben Cotterill

*Please circle  
either YES or NO*

8. Have you read the information sheet? YES / NO
9. Do you understand that you are free to withdraw from the study:
- At any time? YES / NO
  - Without having to give a reason?
10. Do you understand the time requirements of the study? YES / NO
11. Do you agree to take part in this study? YES / NO

---

**12. PARTICIPANT**

Name of PARTICIPANT (BLOCK LETTERS):

.....

Signature of Participant: ..... Date: .....

---

## Appendix N: Mock-juror questionnaire

---

### JUROR PERCEPTION QUESTIONNAIRE

Subject number: .....

Witness number: .....

**How reliable did you find the witness to be overall?**

Not reliable at all

Extremely reliable



1      2      3      4      5      6      7      8  
9

**How detailed did you find the witness to be overall?**

Not detailed at all

Extremely detailed



1      2      3      4      5      6      7      8  
9

**How confident did you find the witness to be overall?**

Not confident at all

Extremely confident



1      2      3      4      5      6      7      8  
9

**How compelling of a witness do you think the eyewitness would make in a courtroom?**

Not compelling at all

Extremely compelling



1            2            3            4            5            6            7            8  
9

**How bias do you think the questioner was?**

Not bias at all

Extremely bias



1            2            3            4            5            6            7            8  
9

**How focused did you find the witness to be overall?**

Not focused at all

Extremely focused



1            2            3            4            5            6            7            8  
9

**How shy did you find the witness to be overall?**

Not shy at all

Extremely shy



1            2            3            4            5            6            7            8  
9

*Appendices*

**How articulate did you find the witness to be overall (i.e., how was their vocabulary/language skills)?**

**Not articulate at all**

**Extremely articulate**



**1**

**2**

**3**

**4**

**5**

**6**

**7**

**8**

**9**

**Please write any comments below in relation to the witness you just watched that may explain your decision making above:**

---

---

## **Appendix O: Participant debrief (Study Three)**

---

### **13. DEBRIEF FOR PARTICIPANTS**

*How interviewing techniques and temperament affect children as eyewitnesses and the decision making of jurors: Study three.*

**Researcher:** Ben Cotterill

*(Supervisors: Faye Skelton, Rory MacLean, and Kathy Charles).*

Thank you for taking part.

The aim of this study was to understand how child temperament and interviewing techniques affect the decision making of jurors.

You just read transcripts from a set of videos during which child witnesses were interviewed via one of three interview styles: (1) open-ended, (2) closed-ended, or (3) misleading. You were allocated to one of these conditions at random, and each child witness you read about received the same format. For each interview style, there were three different types of witnesses: (1) children who reported low shyness during a previous study, (2) children who reported an average level of shyness, (3) and children who reported high shyness.

#### **Effects of interviewing technique**

Suggestive interviewing techniques that have detrimental effects on eyewitness accuracy may be used to discredit their testimony (e.g., Goodman Quas, Bukley, & Shapiro, 1999). When such leading questions are used with children in courtrooms, it can be difficult for jurors to determine whether the child is answering in the way they do because it is the correct answer or because they are simply complying with the lawyer's suggestion (e.g., Kebbell, Evans, & Johnson, 2010; Tubb, Wood, & Hosch, 1999).

#### **Effects of witness confidence and shyness**

Additionally, some research has found that leading questioning can decrease witness confidence (Wheatcroft, Wagstaff & Kebbell, 2004). This is a point of concern because eyewitness confidence remains one of the most persuading factors for jurors (Nicholson, Yarbrough and Penrod, 2014). It is commonly thought that the more confident the witness, the more accurate they are, even though confidence and accuracy do not always positively correlate (Wheatcroft et al., 2004). Since shyer individuals tend to have a lack of confidence in their everyday life and are more likely to question themselves, as well as rely on others for information (Crozier, 2000), it would make sense if they were perceived as less confident and therefore less credible to potential jurors based upon the findings of Wheatcroft et al. (2004).

#### **Research objective**

The objective of the study was to examine how the interviewing techniques and shyness of a child witness may affect the perceptions and decision making of jurors. It was hypothesised that: (1) children who underwent closed-ended questions and misleading questions would be perceived as less credible than those who underwent

## *Appendices*

open-ended questions, (2) and that shy individuals would be perceived as being less confident in their responses and therefore as also being less credible.

### **Further information**

If you require any further information or have any questions about this study, please do not hesitate to contact the researcher, Ben Cotterill ([ben.cotterill@napier.ac.uk](mailto:ben.cotterill@napier.ac.uk); [REDACTED]) or the research's supervisors: Faye Skelton ([f.skelton@napier.ac.uk](mailto:f.skelton@napier.ac.uk)), Rory MacLean ([r.maclea@napier.ac.uk](mailto:r.maclea@napier.ac.uk)), Kathy Charles ([kathy.charles@ntu.ac.uk](mailto:kathy.charles@ntu.ac.uk)).

Alternatively, if you would like to contact an independent advisor, please feel free to contact Cedric English ([C.English@napier.ac.uk](mailto:C.English@napier.ac.uk)).

---

## **Appendix P: Big Five Aspect Scale items**

---

### Neuroticism

#### Volatility

+ keyed      Get angry easily.

Get upset easily.

Change my mood a lot.

Am a person whose moods go up and down easily.

Get easily agitated.

Can be stirred up easily.

– keyed      Rarely get irritated.

Keep my emotions under control.

Rarely lose my composure.

Am not easily annoyed.

#### Withdrawal

+ keyed      Am filled with doubts about things.

Feel threatened easily.

Worry about things.

Am easily discouraged.

Become overwhelmed by events.

Am afraid of many things.

– keyed      Seldom feel blue.

## *Appendices*

Feel comfortable with myself.

Rarely feel depressed.

Am not embarrassed easily.

## Agreeableness

### Compassion

+ keyed      Feel others' emotions.

Inquire about others' well-being.

Sympathize with others' feelings.

Take an interest in other people's lives.

Like to do things for others.

– keyed      Am not interested in other people's problems.

Can't be bothered with other's needs.

Am indifferent to the feelings of others.

Take no time for others.

Don't have a soft side.

## Politeness

+ keyed      Respect authority.

Hate to seem pushy.

Avoid imposing my will on others.

Rarely put people under pressure.

## *Appendices*

– keyed      Insult people.

Believe that I am better than others.

Take advantage of others.

Seek conflict.

Love a good fight.

Am out for my own personal gain.

## Conscientiousness

### Industriousness

+ keyed      Carry out my plans.

Finish what I start.

Get things done quickly.

Always know what I am doing.

– keyed      Waste my time.

Find it difficult to get down to work.

Mess things up.

Don't put my mind on the task at hand.

Postpone decisions.

Am easily distracted.

## Orderliness

+ keyed      Like order.

Keep things tidy.

## *Appendices*

Follow a schedule.

Want everything to be “just right.”

See that rules are observed.

Want every detail taken care of.

– keyed      Leave my belongings around.

Am not bothered by messy people.

Am not bothered by disorder.

Dislike routine.

## Extraversion

### Enthusiasm

+ keyed      Make friends easily.

Warm up quickly to others.

Show my feelings when I’m happy.

Have a lot of fun.

Laugh a lot.

– keyed      Am hard to get to know.

Keep others at a distance.

Reveal little about myself.

Rarely get caught up in the excitement.

Am not a very enthusiastic person.

## *Appendices*

### Assertiveness

+ keyed      Take charge.

Have a strong personality.

Know how to captivate people.

See myself as a good leader.

Can talk others into doing things.

Am the first to act.

– keyed      Do not have an assertive personality.

Lack the talent for influencing people.

Wait for others to lead the way.

Hold back my opinions.

### Openness/Intellect

#### Intellect

+ keyed      Am quick to understand things.

Can handle a lot of information.

Like to solve complex problems.

Have a rich vocabulary.

Think quickly.

Formulate ideas clearly.

– keyed      Have difficulty understanding abstract ideas.

Avoid philosophical discussions.

## *Appendices*

Avoid difficult reading material.

Learn things slowly.

### Openness

+ keyed      Enjoy the beauty of nature.

Believe in the importance of art.

Love to reflect on things.

Get deeply immersed in music.

See beauty in things that others might not notice.

Need a creative outlet.

– keyed      Do not like poetry.

Seldom get lost in thought.

Seldom daydream.

Seldom notice the emotional aspects of paintings and pictures.

## Appendix Q: Study three correlation matrix

**Table 8.11**

*Correlation matrix: Juror personality and questionnaire responses.*

	Reliability	Confidence	Compelling	Shy
Openness	.04	.13	.28	.25
Openness	.05	.08	.18	.25
Intellect	.01	.14	.31	.19
Conscientiousness	-.40	-.35	-.32	-.06
Orderliness	-.34	-.26	-.30	-.05
Industriousness	-.34	-.34	-.23	-.05
Agreeableness	-.05	.01	-.06	.49*
Compassion	-.04	.01	-.02	.59*
Politeness	-.06	.01	-.09	.33
Extraversion	-.10	-.22	-.19	-.20
Assertiveness	-.07	-.14	-.12	-.51*
Enthusiasm	-.24	.21	-.18	.24
Neuroticism	.29	.32	.26	.04
Withdrawal	.47*	.55*	.49*	.24
Volatility	-.10	-.05	-.09	-.21

*Note.* \*  $p > .05$ .