‘It’s all in the detail’: Examining Verbal Differences between Children’s True and False Reports using Cognitive Lie Detection Techniques

HANNAH CASSIDY

This thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth.

23rd November 2016

Supervised by: Dr Lucy Akehurst

Dr Julie Cherryman

Professor Aldert Vrij
“A lie keeps growing and growing until it’s as plain as the nose on your face”

The Blue Fairy, Pinocchio (1940)
**General Abstract**

Police interviewers require a new investigative interviewing tool to facilitate the discrimination between children’s true and false reports. This thesis investigated whether cognitive lie detection techniques could fill this gap in current practice. Chapter 1 introduces the cognitive lie detection paradigm, highlighting the lack of research within the child deception literature and the paradigm’s potential as a means for detecting deceit in children. Chapter 2 explores imposing cognitive load through the use of gaze maintenance to exaggerate differences between child truth-tellers and child lie-tellers. In Experiment 1, maintaining gaze (either with the interviewer’s face or a teddy bear’s face) resulted in truth-tellers providing significantly more detailed reports than lie-tellers. This finding was not apparent for the control condition where children were given no gaze instruction. In Experiment 2, this exaggerated difference between the accounts of the truth- and lie-tellers facilitated deception detection when the children were instructed to look at interviewer’s face, but not at the teddy bear’s face. Poor discrimination for the latter group was discussed with regard to the gaze behaviour of the children being regarded as ‘fishy’ by the evaluators. Chapter 3 investigates whether playing an example of a detailed free recall provided by a peer (referred to as another child’s model statement, AMS) elicits longer statements that contain more cues to deceit in an eyewitness context than when no model statement is used. Both child truth-tellers and child lie-tellers provided more details and more new information following AMS. However, truth-teller accuracy decreased. In Chapter 4, interview clips from Chapter 3 were judged by adult evaluators who found it difficult to differentiate between children’s true and false reports. This could be a consequence of quantity of detail not being a reliable indicator of veracity for this sample of interviews. Chapter 5 tests the use of children’s practice interviews as their own model statements (OMS) compared to AMS and having no model statement (NMS). Only AMS
encouraged children to include more details and more new information in their post-model statement true and false reports. Further research is required to understand the socio-cognitive mechanisms that create this behavioural difference. Chapter 6 describes a field study that presented the cognitive lie detection techniques investigated in the previous chapters to police officers who interview child witnesses regularly. Of all the techniques, OMS was considered to be the most viable option, although police officers suggested that all of the interview techniques would require adaptation for use in the real world. The practitioners provided an insightful look at the current child-interviewing context in the UK, which provides a basic framework that could be considered when designing child deception detection strategies in the future. Finally, Chapter 7 summarises the main findings of this doctoral thesis, discusses their theoretical and practical implications, and puts forward ideas for future research.
Contents

Declaration ......................................................................................................................... xii
List of tables ..................................................................................................................... xiii
List of figures .................................................................................................................... xiii
Acknowledgements ....................................................................................................... xiv
Dissemination ................................................................................................................... xv

Chapter 1: General Introduction to Thesis ................................................................. 1
1.1 Introduction ............................................................................................................... 2
  1.1.1 Cognitive lie detection ...................................................................................... 4
  1.1.2 A gap in the literature ..................................................................................... 4
  1.1.3 Child deception .................................................................................................. 7
    1.1.3.1 Type of lie .................................................................................................. 7
    1.1.3.2 Age of children ......................................................................................... 9
  1.2 Thesis outline .......................................................................................................... 12
  1.3 References ........................................................................................................... 17

Chapter 2: ‘Look this way’: Using gaze maintenance to facilitate the detection of children’s false reports ................................................................. 26
2.1 Abstract ................................................................................................................... 27
2.2 Introduction ............................................................................................................. 28
  2.2.1 Imposing cognitive load ................................................................................... 29
  2.2.2 Gaze maintenance ......................................................................................... 31
2.3 Experiment 1 ......................................................................................................... 34
  2.3.1 Method ........................................................................................................... 34
    2.3.1.1 Participants ............................................................................................... 34
    2.3.1.2 Procedure ................................................................................................ 35
    2.3.1.3 Coding for detail ..................................................................................... 37
    2.3.1.4 Coding for gaze maintenance ................................................................. 37
  2.3.2 Results ............................................................................................................ 38
    2.3.2.1 Manipulation checks ................................................................................ 38
    2.3.2.2 Hypotheses-testing .................................................................................. 42
Chapter 3: ‘Truly crime-relevant’: Using a model statement in an eyewitness context to differentiate between children’s true and false reports ........................................ 71

3.1 Abstract ..................................................................................... 72
3.2 Introduction ................................................................................ 73
  3.2.1 Using a model statement ......................................................... 73
  3.2.2 Interviewee productivity .......................................................... 74
  3.2.3 Forensic usefulness ................................................................. 76
  3.2.4 Between-statement consistency .............................................. 77
  3.2.5 Memory accuracy ................................................................. 78
  3.2.6 Aims .................................................................................... 79
3.3 Method ...................................................................................... 80
  3.3.1 Participants ........................................................................... 80
  3.3.2 Model statement .................................................................. 80
  3.3.3 Procedure ............................................................................ 81
  3.3.4 Coding schemes .................................................................. 83
    3.3.4.1 Interviewee productivity ................................................. 83
Chapter 5: ‘Own vs. other’: The effects of modelling on verbal cues to children’s false reports ………………………………………………………………………123

5.1 Abstract ……………………………………………………………………………………………124

5.2 Introduction ………………………………………………………………………………………125
  5.2.1 Interviewee productivity ………………………………………………………………………126
  5.2.2 Between-statement consistency ……………………………………………………………130
  5.2.3 Memory accuracy ……………………………………………………………………………132

5.3 Method …………………………………………………………………………………………………133
  5.3.1 Participants …………………………………………………………………………………………………133
  5.3.2 Materials …………………………………………………………………………………………………134
    5.3.2.1 Adapted practice interview ……………………………………………………………………………134
    5.3.2.2 Activities …………………………………………………………………………………………………134
      5.3.2.2.1 Lego activity ……………………………………………………………………………………………135
      5.3.2.2.2 Dice activity ……………………………………………………………………………………………135
    5.3.2.3 Another child’s model statement (AMS) ………………………………………………………136
    5.3.2.4 Own model statement (OMS) ………………………………………………………………………136
    5.3.2.5 Social/self-comparison ……………………………………………………………………………137
    5.3.2.6 Perceived self-efficacy ……………………………………………………………………………137
  5.3.3 Procedure ……………………………………………………………………………………………138
    5.3.3.1 Testing session 1 ……………………………………………………………………………………………138
    5.3.3.2 Testing session 2 ……………………………………………………………………………………………139
  5.3.4 Coding schemes ……………………………………………………………………………………………141
    5.3.4.1 Interviewee productivity ……………………………………………………………………………141
    5.3.4.2 Between-statement consistency ………………………………………………………………………142
    5.3.4.3 Memory accuracy …………………………………………………………………………………………142

5.4 Results ………………………………………………………………………………………………………143
  5.4.1 Interviewee productivity ……………………………………………………………………………143
  5.4.2 Social/self-comparison, perceived self-efficacy and change in interviewee productivity …………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………
Chapter 6: Elicit, evaluate, empower and end product: A basic framework for designing interview strategies to uncover children’s false reports

6.1 Abstract ................................................................. 163
6.2 Introduction ............................................................. 164
6.3 Method ..................................................................... 167
   6.3.1 Participants .......................................................... 167
   6.3.2 Procedure ............................................................ 167
6.4 Results .................................................................. 168
   6.4.1 Analytic strategy ................................................... 168
6.4.2 Evaluation of cognitive lie detection techniques .......... 157
   6.4.2.1 Advantages ....................................................... 157
   6.4.2.1.1 Child’s own model statement ......................... 158
   6.4.2.2 Disadvantages .................................................. 158
   6.4.2.2.1 Maintaining gaze ........................................ 158
   6.4.2.2.2 Another child’s model statement .................... 161
   6.4.2.3 Child’s own model statement ............................ 163
   6.4.2.3 Suggestions for use .......................................... 165
   6.4.2.3.1 Maintaining gaze ........................................ 165
   6.4.2.3.2 Child’s own model statement ....................... 166
6.4.3 Broader child interviewing context ......................... 167
   6.4.3.1 Eliciting information ....................................... 167
   6.4.3.1.1 Techniques ................................................. 167
   6.4.3.1.2 Internal barriers ......................................... 169
   6.4.3.2 Evaluating information .................................... 173
Appendix E: Favourable Ethical Opinion (Chapter 6) .......................................................... 230
Appendix F: UPR 16 form .................................................................................................... 231
Appendix G: Journal article published from Chapter 2 ..................................................... 233
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Hannah Cassidy

WORD COUNT = 51,709 words

(Original word count prior to examiners’ amendment = 47,447 words)
List of tables

2.1 Mean and Standard Deviations for Total Number of Details as a Function of Veracity and Gaze Instruction…………………………………………………………………………………………………….44

2.2 Discrimination Accuracy (d') and Response Bias (β) as a Function of Gaze Instruction, Guidance Provision and Modality of Presentation……………………………………………………57

List of figures

2.1 Average percentage of time spent gazing at each face stimulus as a function of veracity………………………………………………………………………………………………………………………39

2.2 Average percentage of time spent gazing at each face stimulus as a function of gaze instruction……………………………………………………………………………………………………………………….40

4.1 Decision-making process when evaluating ‘pre-model statement only’ recall attempts…………………………………………………………………………………………………………………………112

4.2 Decision-making process when evaluating ‘post-model statement only’ recall attempts……………………………………………………………………………………………………………………………………113

4.3 Decision-making process when evaluating ‘pre- and post-model statement’ recall attempts……………………………………………………………………………………………………………………………114

6.1 Summary of themes highlighting participant contribution per theme……………155

6.2 Representation of superordinate themes by each of the interview techniques…………156

6.3 Superordinate and subordinate themes within the broad child-interviewing context……………………………………………………………………………………………………………………………………157
Acknowledgements

So many people have supported me over the course of my PhD journey, and I would like to thank all of them collectively for keeping me sane and putting up with my occasional crises of confidence!

I would like to say a very special thank you to my supervisor, Dr Lucy Akehurst. Thank you for always believing in me and for giving me great advice! You are an excellent role model and I have the utmost respect for you. I would also like to thank my other two supervisors, Dr Julie Cherryman and Professor Aldert Vrij, for all your feedback and guidance going forward. I consider myself very lucky to have had such a supportive and positive supervisory team. Additionally, I would like to thank Dr Amy-May Leach, who acted as an external advisor on my first study. I really appreciate all your help and positive thinking.

Next, I would like to thank all the people that have made my research possible at a practical level. Thank you to my research assistants and my interviewers for being so reliable. Thank you to the schools and police forces for being so cooperative. Thank you to all the participants, both adults and children, for taking part. Without all of your support, I couldn’t have achieved so much!

I would also like to thank my fellow PhD students within the Psychology department at the University of Portsmouth – you know who you are! We have ridden the emotional rollercoaster together and will all see the light at the end of the tunnel.

Finally, I would like to say a massive thank you to my family and friends outside the Psychology department. Words cannot describe how grateful I am for your love and faith in me. I am especially appreciative of my ‘better half’ who has been the mistress to my PhD for the past three years. Thank you for your understanding and patience.
Dissemination

Conference presentations


Seminar presentations


Publications

Chapter 1: General Introduction to Thesis
1.1 Introduction

Since the Youth Justice and Criminal Evidence Act (1999), young children and other vulnerable witnesses have received more support, via special measures, prior to and during their investigative interviews and, subsequently, during any resultant court proceedings. These special measures include video-recorded interviews, which can be played in court and live link facilities whereby vulnerable witnesses can give their evidence in a room away from the main courtroom.

In the UK, all professionals who interview child witnesses as part of a police investigation are trained and advised to follow guidance entitled, Achieving Best Evidence in Criminal Proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures (ABE, Ministry of Justice, 2011). The main objective of this guidance is to enable vulnerable and intimidated witnesses to achieve the highest possible standard of testimony (i.e. to give their best evidence). This is in line with the primary aim of any investigative interview, which is to elicit reliable information (Vrij, Hope & Fisher, 2014). The credibility of the information elicited is also covered within the guidance. Originally conceived to replace the oath taken by adult witnesses in court, ABE tackles the issue of child credibility by recommending that interviewers perform a competency test to evaluate a children’s conceptual understanding of truths and lies. It also suggests that interviewers should remind child witnesses of the importance of telling the truth. However, in a recent joint inspection of child sexual abuse cases, conducted by Her Majesty’s Inspector of Constabularies and Her Majesty’s Crown Prosecution Service Inspectorate (Criminal Justice Joint Inspection, 2014), it was reported that the competency test was used in less than 60% of interviews, with only 36.2% of those interviews using the correct protocol. Krähenbühl, Blades and Cherryman (2015) also found a lack of consistency in the implementation of the ‘truths and lies’ discussion in a separate sample of
child witness transcripts. At face value, poor compliance appears to be a serious issue. However, empirical findings show that performing a competency test does not necessarily deter children from deception (Lyon, Malloy, Quas & Talwar, 2008). That is, children’s knowledge of the duty to tell the truth is not predictive of their subsequent lie- or truth-telling behaviour (Talwar, Lee, Bala & Lindsay, 2002), and neither is their performance on moral reasoning tasks predictive of their later behaviour (Bala, Lee, Lindsay & Talwar, 2010). The only pre-interview instruction that does significantly decrease the frequency of lying (Talwar et al., 2002) and influences children’s decision to tell the truth/lie (Evans & Lee, 2010) is promising to tell the truth. If these findings are considered alongside the problem of non-compliance, then a different issue presents itself; namely, that a new credibility tool is required, which is empirically tested and that child interviewers will use.

The strength of the above findings, primarily collected by the Child Witness Project based at Queen’s University in Canada, led to Canadian officials significantly reforming the country’s laws governing the process for determining the competence of child witnesses (Bala et al., 2010). In force since January 2006, the Canada Evidence Act now states under section 16.1 that children will be required to promise to tell the truth but will not be questioned regarding their understanding of the nature of the promise to tell the truth. In this sense, the promise has the same effect as swearing an oath if the evidence is later presented as evidence in court. This level of reformation may not, however, be possible in the UK, as ABE guidance clearly states that “there should be no attempt to get the witness to swear an oath, either before or after an interview” (Ministry of Justice, 2011, p. 72). The only recommendation that can currently be used is to remind children of the importance of telling the truth; the effect of which has yet to be empirically tested. There is, however, another line of research that has shown promise in uncovering adult’s deceit and could also be effective with children: cognitive lie detection.
1.1.1 Cognitive lie detection

During interviews, lying can be more mentally taxing than truth-telling (e.g. Christ, Essen, Watson, Brubaker & McDermott, 2009; Hartwig, Granhag, Strömwall & Kronkvist, 2006; Mann & Vrij, 2006). Increased cognitive effort for lie-tellers may be due to increased information-processing demands associated with telling a convincing lie. As lie-tellers can’t take their credibility for granted as much as truth-tellers can, they have to perform multiple tasks simultaneously to increase the likelihood that they will be believed. These include monitoring and regulating their behaviour in response to interviewer feedback as well as constructing and maintaining a consistent lie that does not leak truthful information that might give them away (Vrij, 2008). Cognitive lie detection techniques aim to tax lie-tellers’ cognitive systems even further to allow for greater behavioural (verbal and nonverbal) differences to occur between true and false reports. A recent meta-analysis (Vrij, Fisher & Blank, 2015) demonstrated the benefits of using cognitive lie detection (CLD) techniques to detect adults’ truths (67% accuracy with CLD, 57% accuracy without CLD) and lies (67% accuracy with CLD, 47% accuracy without CLD). Such improvement is in line with Hartwig and Bond’s (2011) review of human lie judgments. They found that detecting lies is more readily facilitated by increasing these behavioural differences than by training lie detectors about actual cues to deceit.

1.1.2 A gap in the literature

Cognitive lie detection techniques manipulate cognitive load through three different strategies: (1) imposing cognitive load, (2) encouraging interviewees to say more, and (3) asking unanticipated questions (Vrij, 2015). In their review of the child deception literature, Talwar and Crossman (2012) highlighted that child lie-tellers, whose ability to lie is closely associated with their developing cognitive abilities, could be particularly susceptible to cognitive lie detection methods. Children with under-developed cognitive
skills are more likely to reveal their deceit (e.g. Evans & Lee, 2011; Talwar & Crossman, 2011; Williams, Leduc, Crossman & Talwar, 2016) so putting additional strain on these cognitive systems might decrease their ability to maintain a lie and, therefore, increase the likelihood that their lie will be uncovered.

Liu et al. (2010) were the first to manipulate cognitive load with children using unanticipated questions. If lie-tellers correctly anticipate which questions will be asked at interview, then they can prepare responses that contain fewer cues to their deceit and lie-telling becomes easier (DePaulo et al., 2003). Asking the unanticipated exploits this forward planning by obliging the interviewee to come up with a plausible answer on the spot or risk looking suspicious (Vrij, 2015). Children might be even more susceptible to the effects of spontaneous responding (compared to adults) because they have less well-developed planning abilities (Talwar & Crossman, 2012). In other words, there may be a greater number of questions for which they have not already planned a response. Indeed, when children are asked unanticipated questions, this can have a negative impact on the verbal content of their lies (Vredeveldt & Wagenaar, 2013). Liu et al. (2010) asked children aged 10 to 12 years old to answer either anticipated (event-relevant; e.g. where did the event take place?) or unanticipated (event-irrelevant; e.g. what did you have in your pocket when it happened?) questions about either a life event they had (truth-tellers) or had not (lie-tellers) experienced. It should be noted that this manipulation of anticipation is not in line with adult research where unanticipated questions are always related to the core event (Vrij, 2015). They found that deceptive children were more likely to answer unanticipated questions than truthful children. In summary, asking the unexpected revealed a behavioural difference in responding between truth-tellers and lie-tellers. However, response length was not affected. The lack of an effect on verbosity could be explained by the fact that Liu et al. (2010) tested older children (10 to 12 years) with well-
developed planning abilities. Another explanation could be the use of a between-subjects design for question type. Liu et al. (2010) may have benefited from having type of question as a within-subjects variable: Differences elicited between adults’ truths and lies, using this method, have been found when comparing the level of detail in their responses for both anticipated and unanticipated questions (Lancaster, Vrij, Hope & Waller, 2013). In spite of these disappointing results, more research with children was required to investigate whether manipulating cognitive load could facilitate the discrimination between children’s true and false reports.

Roos af Hjelmsäter, Öhman, Granhag and Vrij (2014) conducted the next investigation of cognitive load with adolescents. In their study, young adolescents in groups of three either experienced (truth-tellers) or imagined that they had experienced (lie-tellers) an encounter with a stranger. When interviewed individually, each adolescent was asked to perform two recall tasks; an anticipated verbal description of what happened and an unanticipated spatial description by making marks on a sketch. Adult evaluators ($N = 200$) then judged the consistency of the descriptions from children in the same triad. For the anticipated verbal description, there were no perceived differences between children’s true and false reports. However, for the unanticipated spatial description, truth-telling triads were perceived to be significantly more consistent than lie-telling triads. Eliciting this spatial difference strengthens the support for using cognitive lie detection with children, notably with groups of adolescents, to determine their veracity. It demonstrates that, even if child lie-tellers can spontaneously respond to an unanticipated task, there are still other nonverbal markers that might highlight their deceit.

Nonetheless, at the start of this doctoral programme of research, there were no published studies, with children, on the use of imposing cognitive load or encouraging interviewees to say more: this was the gap in the literature that I endeavoured to fill.
1.1.3 Child deception

1.1.3.1 Type of lie

Lying is defined as “a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief, which the communicator considers to be untrue” (Vrij, 2008, p. 15). People can lie through omission of true information, known as false denials, or through the commission of false information, otherwise known as false allegations (Lyon et al., 2008). The prevalence of children's false denials in forensic investigations is difficult to measure because, unless there is sufficient physical evidence to prove that an incident has occurred (but that a child is denying), then false denials are hard to classify. Laboratory research has, however, highlighted children’s willingness to conceal information (i.e. keep secrets) about their own wrongdoing (e.g. Polak & Harris, 1999; Talwar & Lee, 2002) and the wrongdoing of an adult (e.g. Talwar, Lee, Bala & Lindsay, 2004; Tye, Amato, Honts, Devitt & Peters, 1999). Studies into the prevalence of children’s false allegations in forensic investigations suggest low rates based on clinical judgments (2-6% Everson & Boat, 1989; 2%, Jones & McGraw, 1987) and more rigorous review processes of case files (8% Anthony & Watkeys, 1991; 1.5%, Oates et al., 2000; 4%, Trocmé & Bala). However, it does appear that in court cases, and even with regard to allegations as serious as sexual abuse, some children do lie. Although prevalence rates may be relatively low, just one case of a false allegation can have far-reaching consequences for both child witnesses and the defendants (O'Donohue, Benuto & Fanetti, 2010). The current thesis investigates the effects of using cognitive lie detection techniques on false allegations to prevent these serious legal implications.

It is also important to take into consideration that a frequent strategy for reducing the cognitive load of lying is to incorporate truthful elements of unrelated events into the lie; these are called embedded lies (Vrij, Granhag & Porter, 2010). This means that most
lies that people tell are not complete fabrications. Children have reported including real-life components in their false reports as the most frequent verbal strategy that they employ, which could explain adults’ poor deception detection accuracy (Strömwall, Granhag & Landström, 2007). Embedded lies are particularly difficult to detect because they can be rich in detail, which could confuse verbal lie detection tools (Vrij et al., 2010). Referred to in the adult literature as ‘temporal displacement’ (Michael, 2013), embedded lies are most commonly used when the lie-teller must create a false story about a familiar event. In this doctoral thesis, the events about which the children were interviewed ranged in familiarity. However, for all events, the children could have, to some degree, included truthful information (whether based on their own experience or the experience of another) from a different event into their false reports. This provided a more realistic challenge for the cognitive lie detection techniques to overcome.

Finally, lies can be classified according to their perceived stakes: high versus low. Most laboratory research has focused on low-stakes lies with children, although it has been argued that what appears low stakes to adults might not be low stakes for children (Talwar & Crossman, 2012). Nevertheless, only a few studies have considered more ecologically relevant scenarios such as being ‘touched’ by an adult (Goodman, Luten, Edelstein & Ekman, 2006) and reporting the theft of an expensive item to a ‘police officer’ (Tye et al., 1999). In these scenarios, it could be considered that the stakes are perceived to be higher, resulting in an increase in stress. Brunet et al. (2013) showed, using a within-subjects design to control for individual differences, that children’s reports regarding highly stressful events were longer than those regarding less stressful events. As longer statements increase the opportunity for more cues to veracity to occur (Vrij, Mann, Kristen & Fisher, 2007), the benefit of using highly stressful events could be that the effects of veracity are further magnified. However, it is important to consider the ethics of inducing
high stress in children (i.e. a sensitive population). Brunet et al. (2013) used occasions where children had been bullied and reported feeling stressed in their study, thus stress was naturally occurring rather than induced. To avoid ethical complications in this doctoral thesis, non-stressful events were used with the view that any veracity effects would be further enhanced, if later investigated with stressful events.

1.1.3.2 Age of children

The term ‘children’ can refer to any person under the age of eighteen. Child deception research has primarily divided into testing young children aged 2 to 7-years-old (e.g. Ahern, Lyon & Quas, 2011; Evans & Lee, 2013, Talwar & Lee, 2002) and older children aged 7 to 14-years-old (e.g. Akehurst, Köhnken & Hofer, 2001; Brunet et al., 2013, McCarthy & Lee, 2009). The reason for this division around 7 years of age could be due to the finding that children find it difficult to successfully conceal this deceit before this age (Talwar & Lee, 2002), which may be associated with their ongoing development of certain cognitive and social skills (Talwar & Crossman, 2011). Talwar and Lee (2008) tested children aged 3 to 8 years-old and found that 77% of the older children were able to conceal their deceit compared to only 29% of the younger children. This ability to conceal information successfully was also related to their cognitive development of executive functions, such as working memory and inhibitory control, which improved with age. In their lifespan study, Debey, De Schryver, Logan, Suchotzki and Verschuere (2015) investigated patterns in lying behaviour and inhibitory control performance in persons aged 6 to 77-years-old. Their findings showed that both variables followed the same pattern, increasing over the course of childhood, peaking at adolescence and then decreasing again in adulthood. Thus, the ongoing development of certain cognitive skills across the entire lifespan is related to people’s propensity to deceive.
The development of certain social skills could also present a reason for a division in age groups. Lying requires a sense of self as well as an understanding of others’ mental states, also known as theory of mind. To create a false belief in the mind of another, a child must understand that a person’s mental perspective has an effect on his/her actions and emotions (Talwar & Crossman, 2011). Theory of mind further divides into first-order belief understanding and second-order belief understanding. First-order belief understanding refers to creating a belief that is simply false with regard to the true state of affairs (for example in a temptation resistance paradigm, if the child peeked and then said they didn’t, then this would be a false statement). Second-order belief understanding refers to children’s ability to sustain a lie and provide responses that are in line with what the interlocutor believes to be the truth. With regard to age, first-order belief understanding usually develops by the age of 3 years whereas second-order belief understanding is not developed until 6 to 7 years of age (Talwar & Crossman, 2011). Moreover, when it is developed, it has been linked to better lie maintenance (Talwar & Lee, 2008). Thus, studying children from 7-years-old onwards would provide more of a challenge for any study aiming to test new detection techniques.

Another social factor to be considered when interviewing children of different ages is children’s awareness of the impressions they make, referred to in the literature as self-monitoring or self-presentation (Eder & Jones, 1989) and their ability to manipulate these impressions. Although not studied in relation to child deception, self-monitoring could theoretically contribute to children using more verbal and nonverbal strategies to produce an honest impression, when in fact they are acting dishonestly. Furthermore, individual differences in children’s social evaluation concerns could explain why some child lietellers are more motivated to be believed than others because they would better practised in regulating their behaviour in response to social cues (Banerjee, 2002). In other words, they
would be better at adjusting their behaviour in response to the verbal and nonverbal feedback provided by the interviewer who they are attempting to convince.

In this doctoral thesis, the term ‘child/children’ refers to persons aged 7 to 11-years-old. This age group was chosen because the children would have sufficient verbal skills to provide longer deceptive narratives that would test the cognitive lie detection techniques (i.e. no ceiling effect due to generally poor lie-telling performance) but would not make it too challenging for this preliminary investigations (i.e. no floor effect due to general very good lie-telling performance).

It is, however, important to note that age remains a continuous variable and that individual differences in children’s development may result in different rates of development. For example, one seven-year-old might have a better false belief understanding than one eleven-year-old, and one nine-year-old might have a better working memory than one ten-year-old. We cannot, therefore, assume a linear relationship between children’s age and their lie-telling performance in the experiments in this thesis, and so it will be necessary to conduct preliminary analyses for each statistical test to uncover any age effects. Furthermore, where possible, we will endeavour to use within-subjects designs so that children’s lie-telling performance post-interview technique is compared to their lie-telling performance pre-interview technique. This will reduce grouping children into specific age groups and be based on their own level of development.

As well as considering the effects of cognitive lie detection on children’s lie-telling ability between 7 to 11-years-old, it is vital to consider whether these interview strategies could have adverse effects on their ability to provide truthful accounts of eyewitness events. Talwar and Crossman (2012) advocated the investigation of cognitive lie detection with children with the proviso that it did not affect truth-tellers. Any interview technique
should first and foremost have either a neutral, or preferably a positive effect on eliciting truthful information, and then an additional negative effect on deterring or uncovering lie-tellers. This is in line with the primary aim of information gathering during an investigation (Vrij et al., 2014). It was, therefore, important across all the studies in this doctoral research to pay close attention to the effect of the technique on truth-tellers, as well as lie-tellers.

1.2 Thesis outline

The central aim of this PhD thesis was to explore two cognitive lie detection interview techniques with child interviewees, namely imposing cognitive load and encouraging interviewees to say more, as potential means for facilitating truth-tellers and, thus, improving the detection of children’s false reports. This PhD thesis constitutes six studies described across five chapters. It is important to note that there is some repetition throughout the thesis. This is because some of the chapters have been written so that they are independent of all other chapters and so can be read and understood individually (i.e. as standalone journal articles). This means that there is a reference section at the end of each chapter.

Chapter 2: ‘Look this way’: Using gaze maintenance to facilitate the detection of children’s false reports

This chapter includes two experiments that investigate imposing cognitive load through gaze maintenance to facilitate child deception detection. As lie-telling requires more cognitive resources than truth-telling (Zuckerman, DePaulo & Rosenthal, 1981), lie-tellers will have fewer cognitive resources leftover, when compared to truth-tellers, to complete other tasks. Imposing cognitive load taxes lie-tellers’ cognitive resources further by adding in a secondary task to be performed whilst answering questions at interview
The cognitive load experienced by lie-tellers when managing both tasks may exceed that of truth-tellers putting even more strain on their attentional resources. Thus, it makes the interview setting itself more difficult. With adults, secondary tasks, such as telling a story backwards (Evans, Michael, Meissner & Brandon, 2013; Vrij, Leal, Mann & Fisher, 2012; Vrij et al., 2008), or telling a story whilst maintaining eye contact (Vrij, Fisher, Mann & Leal, 2010), have magnified differences between truth-tellers and lie-tellers as well as improving detection accuracy. Indeed, a meta-analysis by Vrij et al. (2015) demonstrated that imposing cognitive load is the most effective lie detection technique for improving lie detection accuracy (34% for standard techniques versus 63% for imposing cognitive load techniques) compared to encouraging interviewees to say more (69% for standard techniques versus 74% for ‘encouraging interviewees to say more’ techniques) and unexpected questions (53% for standard techniques versus 67% for unexpected questioning techniques). As imposing a secondary task puts additional strain on the interviewee’s working memory system, it is anticipated that child lie-tellers would be particularly vulnerable to its effects. Children whose under-developed working memory can already reveal their deceit (Talwar & Crossman, 2011) may be put under further pressure to stop themselves from leaking more cues to their deception. Since the beginning of this doctoral programme, only one very recent study has explored imposing cognitive load through reverse order recall with children (Saykaly, Crossman, Morris & Talwar, 2016). This chapter investigated an alternative secondary task, namely gaze maintenance.

Chapter 3: ‘Truly crime-relevant’: Using a model statement in an eyewitness context to differentiate between children’s true and false reports

In this chapter, child eyewitnesses were played an example of a detailed statement (i.e. a model statement), provided by another child, to investigate whether it would elicit
verbal differences in their true and false accounts. Encouraging interviewees to say more taxes lie-tellers’ cognitive resources further by challenging them to provide more detailed statements (Leal, Vrij, Warmelink, Vernham & Fisher, 2015). Providing a model statement encourages truth-tellers to become more talkative and causes lie-tellers to re-evaluate their tactics (Vrij, 2015). On the one hand, lie-tellers could stick to their original story, but this puts them at risk of looking suspicious. On the other hand, lie-tellers could embellish their story and rise to the interviewers’ standards, but this puts them at risk of revealing cues to their deceit. For child lie-tellers, longer statements may put further strain on their verbal working memory systems that are closely associated with their ability to maintain lies (Alloway, McCallum, Alloway & Hoicka, 2015), decreasing their lie-telling performance. This experiment builds on the only previous study conducted, which was conducted with adult insurance claimants (Leal et al., 2015), by investigating the effects of a model statement in an eyewitness context and with a baseline comparison measure.

Chapter 4: ‘Don’t mess with the detail’: Increasing interviewee productivity using a model statement hinders child credibility judgments

This chapter uses the interview clips from Chapter 3 to see whether the exaggerated verbal differences between truth-tellers and lie-tellers that were found using a model statement made it easier for human observers to distinguish between children’s true and false eyewitness accounts. With adults, cognitive lie detection techniques that elicit longer responses have resulted in significant improvements in truth detection accuracy: 57% for a standard technique and 81% for ‘encouraging interviewees to say more’ technique (Vrij et al., 2015). More specifically, Leal et al. (2015) found, using discriminant analyses, that observers’ plausibility ratings could correctly classify 75% of false accounts and 85% of truthful accounts when a model statement had been played. The data analysis in Chapter 4
goes further to use lens modelling to understand observers’ decision-making processes when faced with video stimuli.

**Chapter 5: ‘Own vs. other’: The effects of modelling on verbal cues to children’s false reports**

This chapter continues to investigate the use of a model statement, provided by another child, with children and, in particular, tests the social proof/social comparison theory behind its effects (Cialdini, 1993; Vrij, Leal, Mann, Vernham & Brankaert, 2015). Furthermore, this study investigates the transformation of children’s practice interviews into their own model statements (i.e. a model statement provided by themselves) as a ‘truth-teller friendly’ substitute to hearing another child’s model statement. Practice interviews are conducted as part of the pre-substantive phase of the National Institute of Child Health and Human Development (NICHD) Investigative Interview protocol (Lamb, Orbach, Hershkowitz, Esplin & Horowitz, 2007) and have been shown to elicit longer and more accurate narratives from child witnesses than when no practice interview is used (Price, Roberts, & Collins, 2013; Roberts, Brubacher, Powell & Price, 2011; Sternberg et al., 1997). Unlike model statements given by a peer, practice interviews have received much support for their beneficial effects on truth-tellers. The effects of using model statements compared to no model statement, and then another child’s model statement (AMS) compared to a child’s own model statement (OMS) were tested. It was anticipated that the potential negative effects of social comparison, that may occur when children hear AMS, could be avoided by introducing a self-comparison with the ‘OMS’ condition, leading to more exaggerated verbal differences between children’s true and false reports in this interview condition.
Chapter 6: Elicit, Evaluate, Empower and End product: A basic framework for designing interview strategies to uncover children’s false reports

In this chapter, a qualitative methodology was used to gather practitioner feedback on the viability of the techniques researched in Chapters 2 to 5. Presented as a knowledge exchange workshop, police officers who currently interview child witnesses in their job role were invited to take part in focus group discussions regarding the three interview strategies: (i) imposing cognitive load through gaze maintenance (Chapter 2), (ii) encouraging interviewees to say more using another child’s model statement (Chapters 3 & 4), and (iii) encouraging interviewees to say more using a child’s own model statement (Chapter 5). The focus group discussions were recorded and transcribed. Transcripts were thematically analysed to reveal the participants’ evaluation of each of the three techniques as well as providing an overview of ‘real world’ child interviewing processes in the UK today. Participants provided insightful feedback on the advantages and disadvantages of each interview technique as well as suggestions for use in future practice. The latter thematic analysis uncovered a basic framework of four themes (the 4Es) that could be taken into consideration when designing lie detection strategies for use with children; the interview techniques should (1) elicit as much information as possible, (2) allow the recipients of that information (i.e. police officers, judges, jurors) to evaluate its credibility, (3) empower the interviewees to give their ‘best evidence’, and (4) result in an end product that will stand up to criticism in court.

Chapter 7: General Discussion

The final chapter discusses the main findings of this PhD thesis, reflects on both the theoretical and practical implications of this doctoral research, and puts forward suggestions for future research.
1.3 References


Polak, A. & Harris, P.L. Deception by young children following noncompliance.


Chapter 2: ‘Look this way’: Using gaze maintenance to facilitate the detection of children’s false reports

The research presented in this chapter has been accepted for publication in Applied Cognitive Psychology.
2.1 Abstract

In two experiments, we investigated whether imposing a secondary task is an effective technique for detecting child deceit. Firstly, 85 children aged 8 to 11 years old provided either a true or false report of a recent school event. At interview, some children were asked to gaze towards either the interviewer’s face (IF) or a teddy bear’s face (TF), whereas some children were given no gaze instruction. In both the IF and TF conditions, lie-tellers provided significantly fewer details than truth-tellers. 192 adult evaluators then judged the credibility of ten children’s reports from one of the three ‘gaze’ conditions with and without guidance on level of detail. Evaluators discriminated truths from lies successfully when judging children instructed to look at IF, but not when children were asked to gaze towards TF. Evaluators who received guidance demonstrated better discrimination between true and false reports than evaluators who received no such information.
2.2 Introduction

Child deception research has focused on both the developmental origins of children’s lie-telling behaviours, and the forensic implications of deceptive child testimonies going undetected (see Talwar & Crossman, 2012 for a review). Past research has painted a bleak picture: Children not only have the potential to lie in forensic interviews (Tye, Amato, Honts, Devitt & Peters, 1999), but, when the video-recordings of their statements are presented to legal professionals (e.g. police officers, judges), they experience great difficulty in uncovering false testimonies (Bala, Ramakrishan, Lindsay & Lee, 2014; Leach, Talwar, Lee, Bala & Lindsay, 2004). Thus, if children do decide to provide deceptive reports, then they could easily slip through the net resulting in miscarriages of justice that are damaging to both the victims and defendants (O’Donohue, Benuto & Fanetti, 2010). Clearly, more effective deception detection strategies are needed.

Cognitive processing is an important factor in deception (Zuckerman, DePaulo & Rosenthal, 1981), particularly for children whose growing cognitive abilities are closely related to their ability to maintain false reports (Talwar & Crossman, 2011). Indeed, children’s development of global executive functioning (Gordon, Lyon & Lee, 2014) as well as their development of specific executive functions, such as inhibitory control, working memory, executive planning and forward search planning, significantly contribute to their ability to conceal incriminating information when questioned (Alloway, McCallum, Alloway & Hoicka, 2015; Evans & Lee, 2011; Talwar & Lee, 2008; Williams, Leduc, Crossman & Talwar, 2016). Furthermore, lie-telling proficiency follows the developmental patterns of cognitive processes, such as inhibitory control (Debey, De Schryver, Logen, Suchotzki & Verschuere, 2015). This suggests that child lie-tellers, who
are still developing certain cognitive skills that might facilitate their lie-telling, might be affected by any interview technique that impacts upon these skills.

Growing research into adult deception has highlighted cognitive lie detection (CLD) as a promising strategic tool. Based on the well-established premise that lying is more cognitively demanding than truth-telling (e.g. Christ, Van Essen, Watson, Brubaker & McDermott, 2009; Hartwig, Granhag, Strömwall & Kronkvist, 2006; Mann & Vrij, 2006), CLD manipulates cognitive load, which refers to information-processing demands associated with attentional and working memory (Block, Hancock & Zakay, 2010), and transforms it into a system variable (Vrij, 2015). As a result, CLD techniques exaggerate behavioural differences between truth-tellers and lie-tellers, ultimately leading to impressive improvements in correct judgements of truths: 57% for standard approach to 67% for CLD approach, and correct judgments of lies: 47% for standard approach, 67% for CLD (Vrij, Fisher & Blank, 2015). By taxing cognitive load further, CLD decreases lie-telling performance. Children should be particularly susceptible to the negative effects of increased cognitive demand because their developing cognitive abilities, which already reveal their deceit, would be put under further strain.

2.2.1 Imposing cognitive load

Imposing cognitive load transforms the cognitive demand experienced by interviewees into a system variable through the addition of a secondary task (Vrij, 2015). Knowles (1963) proposed that each person has a limited pool of attentional resources that are differentially allocated to tasks according to difficulty. A difficult task, such as lie-telling, would draw more resources from this pool than a less difficult task, such as truth-telling. Lie-tellers would, therefore, have fewer resources (than truth-tellers) remaining if the pool were finite.
This asymmetry in the availability of cognitive resources for truth-tellers and lie-tellers has two consequences for lie-tellers when a secondary task is imposed. First, lie-tellers experience an overall increase in cognitive demand, working at or near to full attentional capacity. This means that lie-tellers exhibit more behavioural cues indicative of cognitive load compared to truth-tellers. Second, interference between the tasks may arise. When working at cognitive capacity, performance will depend on a person’s ability to divide his or her attention in accordance with task demands. Attention can be flexibly allocated from moment to moment (Kahneman, 1973): As the secondary task becomes more difficult, additional resources can be allocated. If the tasks share a particular pool of resources, then diverting resources from the primary task to the secondary task should result in a trade-off (i.e. decreasing performance for the primary task and increasing performance for the secondary task).

Imposing cognitive load in order to detect deception could be particularly effective with a younger population whose ability to manage their attentional resources has not yet fully matured. Before the age of 11 years, children find it difficult to differentially allocate their attention in dual-task processing (Irwin-Chase & Burns, 2000). Furthermore, research has shown that the increase in cognitive load experienced, when moving from single tasks to dual-tasks, is greater for children than it is for adults (Karatekin, 2004). Although 10-year-olds can allocate their attention similarly to adults, their control over attention management in response to task difficulty is not yet fully developed. In the context of the current study, this suggests that child lie-tellers may overcompensate for the rising demands of a secondary task, diverting too many resources away from the primary task of lie-telling, thus decreasing their performance on this task. It is also possible that children may prioritise the primary task, sacrificing their performance on the secondary task.
To date, two studies have examined the effects of cognitive lie detection techniques on children. Firstly, Liu et al. (2010) asked unanticipated questions of children aged 10 to 12 years old about a non-experienced life event. They found that, compared to truth-tellers, child lie-tellers were more likely to respond to unexpected questions. Secondly, Saykaly, Crossman, Morris and Talwar (2016) imposed cognitive load by asking children to falsely allege or deny play with a certain toy using the ‘reverse order’ interview instruction. Their results revealed that reverse order recall made it harder for child lie-tellers to maintain their reports compared to child truth-tellers, suggesting that telling a story backwards does increase cognitive demands. In summary, both these studies indicated that, when children have to perform a secondary task (i.e. answering a difficult question) at the same time as maintaining their false reports, their ability to maintain the lie is negatively affected. In the current experiment, the secondary task, introduced at interview, was an instruction to maintain gaze with either the interviewer’s face or a teddy bear’s face: A secondary task that has yet to be investigated with children.

2.2.2 Gaze maintenance

Using a systematic approach, Glenberg, Schröder and Roberston (1998) demonstrated that as the cognitive demands (i.e. cognitive difficulty) of a task increase, adults naturally avert their gaze. This cognitive strategy of gaze aversion is functional, as adults performed better on moderately difficult questions when they disengaged from (i.e. closed their eyes), rather than engaged with (i.e. looked at the interviewer’s nose), disruptive visual components in their environment. Looking towards a visual/social stimulus, therefore, interfered with their task performance when the cognitive demands of the task were moderate. This behavioural response to avoid cognitive overload has also been investigated with children. Doherty-Sneddon, Bruce, Bonner, Longbotham and Doyle (2002) compared gaze aversion behaviour in children aged 5 and 8 years old in
response to easy (low cognitive load) and difficult (high cognitive load) questions. Results revealed that the older children averted their gaze away from the questioner’s face more frequently in response to rising question difficulty (i.e. cognitive effort), but that this gaze pattern was only observed for younger children for certain types of questions. This suggests that gaze aversion is used as an overt response to cognitive effort more consistently with increasing age. In addition, there is evidence to support that the primary function of gaze aversion is to manage cognitive demands rather than as a response to social difficulty. Doherty-Sneddon and Phelps (2005) measured gaze aversion in 8-year-old children who were questioned either face-to-face or via live video link. Results revealed that question difficulty strongly influenced gaze aversion in both interview conditions. In the current study, it was anticipated that, as children’s ages ranged from 8 to 11 years old, they would attempt to use gaze aversion to reduce cognitive effort more so in the ‘lie-telling’ condition where cognitive load is higher than in the ‘truth-telling’ condition. Requiring interviewees to maintain gaze during questioning, as was the case for this study, would disable this coping mechanism for lie-tellers and maintain the increased cognitive demands of providing a false report. Furthermore, as maintaining gaze is not a natural behaviour, it would be necessary for interviewees in this experiment to intentionally remind themselves to comply with our gaze instruction, creating additional cognitive load.

In a previous study, maintaining eye contact was used to impose cognitive load on adult interviewees (Vrij, Mann, Leal, & Fisher, 2010). The researchers found that requiring eye contact elicited two cognitive cues (out of 14 cues) that discriminated lie-tellers from truth-tellers; namely, deceitful accounts contained fewer spatial details and were more chronological compared to truthful accounts. No significant differences were elicited between truth-tellers and lie-tellers when interviewees were given no ‘eye contact’
instruction. In terms of detection accuracy, the small difference in elicited cues only improved lie detection accuracy from 44% in the ‘control’ condition to 53% in the ‘eye contact’ condition. If an improvement in accuracy rates is dependent on the exaggeration of behavioural differences between truth-tellers and lie-tellers, then eliciting two cognitive cues did not suffice. Vrij et al. (2010) suggest that these findings may be the product of anxiety for lie-tellers rather than increased cognitive load. Alternatively, previous research has shown that, even when adults find maintaining gaze with a person’s face to be more difficult than either looking at the floor or closing their eyes, this does not result in them performing worse in the former condition compared to the latter two gaze conditions (Doherty-Sneddon, Bonner & Bruce, 2001).

On the contrary, the difficulty that children experience when instructed to direct their gaze does translate into poorer performance compared to a control condition involving no gaze instruction. In their first experiment, Doherty-Sneddon et al. (2001) compared the effect of gaze instruction (look at the speaker vs. look at the floor vs. close your eyes) on both adults’ and 10-year-old children’s task performance. Like adults, most children (83%) found looking at the floor or closing their eyes to be the easiest (least cognitively demanding) conditions. Results showed that, when children looked at the floor, this reported ease translated into them performing significantly better than when they looked at the speaker. This difference in task performance was also found across Doherty-Sneddon et al.’s subsequent experiments for different tasks and for a younger age group (6 years old). Children, therefore, experienced great difficulty in moderating the negative effects of gaze maintenance, with looking towards a face resulting not only in increased levels of cognitive demand for children, but also diminished task performance (Doherty-Sneddon et al., 2001). Thus, it was anticipated, for the current study, that children’s interview performance would be affected by gaze maintenance.
2.3 Experiment 1

The aim of Experiment 1 was to investigate whether an instruction to maintain gaze would exaggerate differences between children’s true and false reports. With a view to the future practical value of this research, it was important to consider how appropriate an instruction to maintain gaze would be with a child population. As maintaining gaze has already been linked to anxiety (Vrij et al., 2010), asking a child to look at an interviewer’s face may intimidate some interviewees. In this study, we instructed some of the children to look at a face stimulus considered to be less intimidating; a teddy bear’s face. It should be noted that toys can be useful in child witness interviews (Wilson & Powell, 2001), and a teddy bear was chosen because it has a face and is non-gender specific.

In this experiment, we predicted that lie-tellers would experience more dual-task interference than truth-tellers when instructed to maintain gaze. That is, lie-tellers’ ability to provide a detailed account would be more negatively associated with their level of gaze compliance, compared to truth-tellers (Hypothesis 1). Secondly, we anticipated that this dual-task interference would exaggerate subtle differences in level of detail between true and false reports. Thus, it was expected that child lie-tellers would provide reports that were significantly less detailed than those provided by child truth-tellers, and this difference in detail would be greater for children instructed to maintain gaze compared to children given no gaze instruction (Hypothesis 2).

2.3.1 Method

2.3.1.1 Participants

Eighty-five children (37 boys, 48 girls) aged 8 to 11 years old ($M = 10.46$ years, $SD = .81$ years) were recruited from four primary schools in the United Kingdom. Participant information sheets were sent home to children’s legal guardians who returned a signed
written consent form. The general procedure was outlined to the children to obtain their verbal assent to participation, but they were naïve to the specific purpose of the study and to the anticipated effect of maintaining gaze. All children, who were asked to lie, complied with the request to lie. Verification was sought from teachers that they had not taken part in the event that they were interviewed about. All children received a certificate and a stationery set in exchange for taking part.

2.3.1.2 Procedure

The experiment took place in two quiet areas of each school and involved the Principal Investigator (PI) who ran the study and a Research Assistant who conducted all interviews and was blind to the aims and hypotheses of the study. All children were tested individually.

The PI invited each child to take part in a short interview about a recent event at their school, thus events differed across schools. These events included a school sports day, a visit to the local cathedral, a school play, and a music concert. Children were randomly assigned to a Veracity condition within each year group in each school so that there were roughly equal numbers of truth-tellers and lie-tellers for each of the four events. Truth-tellers \((n = 39, M_{age} = 10.28 \text{ years}, SD_{age} = .83 \text{ years})\) were interviewed after they had experienced the event and were asked to provide a truthful recollection of what happened. Lie-tellers \((n = 46, M_{age} = 10.43 \text{ years}, SD_{age} = .81 \text{ years})\), on the other hand, were interviewed about an event that they had not experienced and were asked to convince the interviewer that they had already taken part in the event, when in fact they had not. This is similar to the veracity allocation carried out by other researchers interested in eliciting false allegations from children (e.g. Akehurst, Köhnken & Höfer, 2001; Brunet et al., 2013; Lyon, Malloy, Quas & Talwar, 2008).
Within their veracity groups, children were also randomly assigned to a Gaze Instruction condition: Look at the interviewer’s face (IF, n = 28, M<sub>age</sub> = 10.25 years, SD<sub>age</sub> = .80 years) or Look at the teddy bear’s face (TF, n = 29, M<sub>age</sub> = 10.52, SD<sub>age</sub> =.74 years) or No gaze instruction (Control, n = 28, M<sub>age</sub> = 10.32 years, SD<sub>age</sub> = .91 years). The teddy bear was seated on the interviewer’s lap throughout all interviews (i.e. for all conditions). Prior to the interview, children in the IF and TF conditions were instructed by the PI to maintain gaze with the relevant face stimulus as much as they possibly could throughout the interview (i.e. to look at it as much as they could remember to do so). All children then received a sheet listing general themes that they could tell the interviewer about (e.g. talk about who was there, what happened, when it happened). This does not constitute coaching as neither truth-tellers nor lie-tellers were told exactly what they should say and they did not rehearse their story with the PI. Providing children with these themes was anticipated to elicit longer statements, allowing for more cues to deceit to occur (Leal, Vrij, Warmelink, Vernham & Fisher, 2015; Vrij, 2015). All participants were given approximately three minutes to prepare themselves before the PI escorted them to the interview room. Before entering the interview room, children in the IF and TF condition were given a final reminder by the PI to maintain gaze with the relevant face stimulus. This was done out of earshot of the interviewer so that she remained blind to the aims and hypotheses of the study.

The interview protocol reflected the initial stages of a Cognitive Interview (Fisher & Geiselman, 1992): A rapport-building phase (that took place off-camera) was followed by two open-ended questions. First, an invitation to provide a free, uninterrupted narrative (e.g. tell me everything that happened when you took part in your school sports day), and then, secondly, a request, to all interviewees, to provide one additional piece of information about an aspect of the event that they had not already mentioned. No other
questions were asked. All children were video-recorded, and their interviews later transcribed. All interviewees were asked the following question, which served as a manipulation check: Where were you instructed to look during the interview? The response options were ‘interviewer’s face’, ‘teddy bear’s face’ or ‘no instruction given’.

2.3.1.3 Coding for detail

Two independent coders rated the children’s interview transcripts for number of details included. To make the coding more precise, all transcripts were coded for five different types of details; visual details (e.g. “white clay head” contains three visual details), auditory details (e.g. “the teacher told us to take deep breaths” contains one auditory detail), spatial details (e.g. “he stood behind the curtain” contains one spatial detail), temporal details (e.g. “at the end we left” contains one temporal detail), and action details (e.g. “we played football” contains one action detail). One coder coded all of the transcripts for the current study, whilst the second coder rated a random sample of 20 transcripts. Considering that general level of detail is a reliable indicator of veracity (DePaulo et al., 2003), total number of details was calculated for each interviewee, by adding together the scores for all five detail types. Intra-class correlation coefficients (ICCs) were calculated for the two coders. Inter-rater reliability was high, with all ICCs demonstrating high levels of agreement between coders (visual details, ICC = .96; auditory details, ICC = .98; spatial details, ICC = .94; temporal details, ICC = .96; action details, ICC = .92; and total number of details, ICC = .98).

2.3.1.4 Coding for gaze maintenance

To provide an objective measure of gaze behaviour, two different independent judges, using INTERACT 14.0 software (Mangold, 2015), coded all interviews (from start to end) for the amount of time (in seconds) that the child interviewees gazed towards the interviewer’s face (IF) and the teddy bear’s face (TF). The duration of these gaze patterns
for both face stimuli were then added together to give the total number of seconds spent gazing at the IF and the TF for each child. Percentage of time spent gazing towards both the IF and the TF were calculated by taking the total number of seconds spent gazing towards each face stimuli, dividing it by the total length of the interview in seconds and multiplying it by 100. Percentage of time spent gazing elsewhere was calculated by adding together the percentages for IF and TF and subtracting this total from 100. First, both raters coded 17 interviews (20% of the total) to check for inter-rater reliability. Inter-rater reliability was high for time spent looking at the interviewer’s face (ICC = .99) and at the teddy bear’s face (ICC = .91). Rater 1 then coded the next 40% of the video recordings (n = 34) and Rater 2 coded the remaining 40% of the video recordings (n = 34). Percentage of time spent gazing at each face stimulus was calculated for each child by dividing the time spent gazing at the stimulus (in seconds) by the total duration of the interview (in seconds) and multiplying the result by 100.

2.3.2 Results

2.3.2.1 Manipulation checks

All 85 children correctly indicated where they had been asked to look during the interview. To test level of compliance more objectively, two-way ANOVAs were performed with Veracity and Gaze Instruction as the between-subjects factors. These were conducted to investigate differences in percentage of time spent gazing at (a) the interviewer’s face, (b) the teddy bear’s face, and (c) elsewhere (i.e. towards neither face stimulus). Figure 2.1 displays the distribution of gaze behaviour across ‘veracity’ conditions and Figure 2.2 across ‘gaze instruction’ conditions.

In terms of gazing towards the interviewer’s face, there was a significant main effect of Veracity, $F(1, 79) = 5.78, p = .019$. Children providing a false report ($M =$
45.80%, $SD = 22.41$) spent a higher percentage of their interviews looking at the interviewer’s face than children providing a true report ($M = 35.24\%, SD = 23.72$), $d = .46$, 95% CI [.03, .89]. There was also a significant main effect of Gaze Instruction, $F(2, 79) = 10.50$, $p<.001$. Pairwise comparisons using Bonferroni adjustment showed that children instructed to look at the interviewer’s face ($M = 55.93\%, SD = 24.97$) spent a greater portion of the interview gazing at the interviewer’s face than children instructed to look at the teddy bear’s face ($M = 31.61\%, SD = 23.41$, $p<.001$, $d = .98$, 95% CI [.42, 1.52], or given no gaze instruction ($M = 28.71\%, SD = 17.49$, $p = .001$, $d = 1.04$, 95% CI [.47, 1.59]). There was no difference between these latter conditions, $p = 1.00$. There was no significant interaction effect, $F(2, 79) = 1.10$, $p = .34$.

*Figure 2.1* Average percentage of time spent gazing at each face stimulus as a function of veracity
In terms of gazing towards the teddy bear’s face, there was no significant main effect of Veracity, $F(1, 79) = .32$, $p = .57$. There was, however, a significant main effect of Gaze Instruction, $F(2, 79) = 9.50$, $p < .001$. Pairwise comparisons using Bonferroni adjustment showed that instructing children to gaze at the teddy bear’s face ($M = 16.77\%$, $SD = 18.77$) resulted in a higher percentage of time looking at the teddy bear’s face than instructing children to look at the interviewer’s face ($M = 5.22\%$, $SD = 5.21$, $p = .001$, $d = .83$, 95% CI [.29, 1.37]), or giving no gaze instruction ($M = 4.66\%$, $SD = 2.66$, $p < .001$, $d = .90$, 95% CI [.35, 1.44]). There was no difference between these latter conditions, $p = 1.00$. There was no significant Veracity X Gaze Instruction interaction effect, $F(2, 79) = .28$, $p = .76$.

![Figure 2.2: Average percentage of time spent gazing at each face stimulus as a function of gaze instruction](image)

*Figure 2.2: Average percentage of time spent gazing at each face stimulus as a function of gaze instruction*
Finally, in terms of gazing elsewhere, there was a significant main effect of Veracity, $F(1, 79) = 7.15, p = .009$. Truth-tellers ($M = 56.66\%, SD = 24.22$) spent a higher proportion of the interview looking elsewhere compared to lie-tellers ($M = 44.48\%, SD = 21.37$), $d = .54$ (95% CI [.10, .97]). There was also a significant main effect of Gaze Instruction, $F(1, 79) = 7.99, p = .001$. Pairwise comparisons using Bonferroni adjustment showed that children given no gaze instruction ($M = 61.73\%, SD = 17.74$) spent more time looking elsewhere compared to children instructed to look at the interviewer’s face ($M = 38.85\%, SD = 24.21$), $p<.001, d = 1.08$ (95% CI [.51, 1.64]). Percentage of time looking elsewhere did, however, not differ between children in the ‘control’ condition and those in the ‘teddy bear’s face’ condition ($M = 49.63\%, SD = 22.61$), $p = .10$. There was also no significant difference in percentage of time spent gazing elsewhere between children in the ‘interviewer’s face’ condition and child in the ‘teddy bear’s face condition, $p = .20$. There was no significant interaction effect, $F(2, 79) = .80, p = .45$.

In sum, children were able to comply with the instruction to look at the interviewer’s face or the teddy bear’s face. That said, although our instructions did increase time spent gazing toward a specific face stimulus, overall compliance was relatively poor as the average participant complied with their gaze instruction for less than 50% of their interview. Furthermore, children in the ‘teddy bear’s face’ condition only spent 16% of the time looking at their specified stimulus and just as much time looking at the interviewer’s face and elsewhere as children in the ‘control’ condition. This lack of compliance may be because gazing at a static toy when responding to a person is an unnatural behaviour. It could also be because the location of the teddy bear was problematic; staring at the interviewer’s lap may have seemed strange.
2.3.2.2 Hypotheses-testing

Preliminary analyses revealed no significant effects of child age, child gender, or specific activity reported (e.g. sports day, school trip) during the interview, on any of the dependent variables. The data for all participants were, therefore, combined for subsequent analyses.

2.3.2.2.1 Dual task interference

We investigated whether lie-tellers experienced more dual-task interference than truth-tellers, when given the secondary task of maintaining gaze with either the interviewer’s face or the teddy bear’s face whilst being questioned. The ‘performance operating characteristic’ (POC, Norman & Bobrow, 1975) of truth-tellers and lie-tellers was calculated separately for children in both ‘gaze instruction’ conditions. By calculating Pearson’s correlations between the total number of details included in the interviewee’s account (i.e. level of detail) and the time they spent gazing towards either the interviewer’s or the teddy bear’s face (i.e. level of gaze compliance), we were able to examine to what extent the two tasks interfered with one another. High levels of interference would be characterised by a strong negative correlation between performances on both tasks (i.e. increasing compliance with the gaze instruction resulting in decreasing level of detail in responses).

First, when the secondary task required interviewees to look at the interviewer’s face, findings revealed a weak, negative correlation for truth-tellers, $r = -.28$, $p = .40$, and a small to moderate, positive correlation for lie-tellers, $r = .39$, $p = .16$. Although these correlations are not significant, this may be due to the effect of a limited sample size. Following the suggestion of Ferguson (2009), we therefore looked at the effect size of these correlations as “effect sizes are resistant to sample size influence, and thus provide a truer measure of the magnitude of effect between variables” (p. 532). Interpreting these $r$
values as effect sizes (Field, 2013), the data showed that there was a small effect for truth-tellers and a medium effect for lie-tellers. This suggests that there was mild interference between truth-tellers’ ability to provide detailed answers and their compliance with the gaze instruction. However, it also shows that there was no interference for lie-tellers, whose level of detail in fact increased with their level of compliance with the gaze instruction. Second, when interviewees were instructed to look at the teddy bear’s face, there was no correlation between level of detail and compliance with the gaze instruction for truth-tellers, $r = -.04, p = .91$, nor for lie-tellers, $r = .08, p = .78$.

### 2.3.2.2.2 Level of detail

Preliminary analyses showed that true reports ($M = 750.79, SD = 670.31$) contained significantly more words than false reports ($M = 508.33, SD = 560.88$), $t(83) = 1.82, p = .037, d = .40$ (95% CI [-.37, .82]). As longer reports allow for more details to occur, length of statement would have an effect on our analysis of total detail. To take this effect into account, length of statement (in words) was entered as a covariate in our analyses. This is similar to previous work by Strömwall and Granhag (2005) when analysing reality monitoring scores.

First, a 2 (Veracity) x 3 (Gaze Instruction) ANCOVA was performed with total number of details as the dependent variable. There was a significant main effect of Veracity, $F(1, 78) = 8.44, p = .005$, a significant main effect of Gaze Instruction, $F(2, 78) = 3.16, p = .048$, and a significant Veracity X Gaze Instruction interaction effect, $F(2, 78) = 4.22, p = .018$. Descriptive statistics for each of the experimental cells are displayed in Table 2.1.

Of interest for the hypotheses is the Veracity X Gaze Instruction interaction effect. Separate ANCOVAs were conducted: first, for each of the Gaze Instruction conditions
with Veracity as the independent variable, and second, for each of the Veracity conditions with Gaze Instruction as the independent variable. When children were instructed to gaze at the interviewer’s face, truth-tellers provided significantly more details compared to lie-tellers, \( F(1, 25) = 8.53, p = .007, d = .92 \) (95% CI [.13, 1.70]). Similarly, when children were instructed to look at the teddy bear’s face, truth-tellers provided more details in their statements than lie-tellers, \( F(1, 26) = 5.88, p = .023, d = .83 \) (95% CI [.058, 1.59]). Veracity did not have a significant effect on the number of details provided by children who were given no gaze instruction, \( F(1, 25) = .24, p = .63 \). Irrespective of whether they were providing a true report or a false report, children in the control condition included the same amount of detail.

Table 2.1

<table>
<thead>
<tr>
<th></th>
<th>True Report</th>
<th>False Report</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
<td>( M )</td>
</tr>
<tr>
<td>Interviewer’s face</td>
<td>178.77</td>
<td>152.50</td>
<td>79.00</td>
</tr>
<tr>
<td>Teddy bear’s face</td>
<td>152.46</td>
<td>86.93</td>
<td>92.94</td>
</tr>
<tr>
<td>Control</td>
<td>117.15</td>
<td>99.39</td>
<td>114.80</td>
</tr>
<tr>
<td>Total</td>
<td>95.52</td>
<td>79.27</td>
<td>149.46</td>
</tr>
</tbody>
</table>

For children who provided a truthful account, there was a significant effect of Gaze Instruction condition, \( F(2, 35) = 4.04, p = .026 \). Post-hoc testing using Bonferroni adjustment revealed that truth-tellers who looked at the interviewer’s face provided more details than truth-tellers who were given no gaze instruction, \( p = .03, d = .48 \) (95% CI [-.31, 1.25]). There was no difference in quantity of detail between truth-tellers looking at the interviewer’s face and those looking at the teddy bear’s face, \( p = 1.00 \), and no
difference between truth-tellers looking at the teddy bear’s face and those in the control condition, $p = .14$. For children who provided a fabricated account, there was no significant effect of Gaze Instruction, $F(2, 35) = .55$, $p = .58$.

### 2.3.3 Discussion

The analysis of the association between providing a detailed account and complying with the gaze instruction revealed a small positive effect for lie-tellers in the ‘interviewer’s face’ condition. That is, the more the lie-tellers looked at the interviewer’s face the more details they gave. This was contrary to Hypothesis 1. Furthermore, the instruction to look at the teddy bear’s face did not elicit dual task interference for the lie-tellers nor for the truth-tellers. Our theoretical assumption posited that lie-tellers, who have a more cognitively demanding primary task compared to truth-tellers, would reach the limit of their resources when a secondary task was imposed (Knowles, 1963), and, therefore, experience a high level of dual-task interference (Kahneman, 1973). However, our analysis of lie-tellers’ dual-task interference does not support this theoretical assumption. Indeed, the positive relationship between level of detail and gaze compliance for lie-tellers instructed to look at the interviewer’s face completely contradicts our hypothesis. This could be due to the cognitive resources required for each task originating from separate (limited) resources. Multiple resource theory (Wickens, 2002) posits that tasks that are structurally dissimilar, such as answering interview questions (verbal) and maintaining gaze (visual/social), will interfere less. This may explain why imposing cognitive load through constructing/maintaining a lie (verbal) and telling the lie in reverse order (verbal) had greater success in previous studies (Saykaly et al., 2016; Vrij et al., 2008) because the two tasks use similar cognitive processes.

An alternative explanation could be that the effect of gaze maintenance on task performance can vary dependent on the relevance of the visual stimulus to the primary task
(Doherty-Sneddon et al., 2001). It could be the case, in the current study, that child lie-tellers instructed to look towards the interviewer’s face found the information communicated by her face more task-relevant than truth-tellers. Lie-tellers, who are more concerned with appearing honest than truth-tellers (Vrij, 2015), might have monitored the interviewer’s face for feedback on how their deception was being received and used this to modify their responses (e.g. to say more to appear honest). However, this tactic works to their disadvantage, as longer statements are more likely to contain cues to deceit (Vrij et al., 2015). This would particularly be the case for child interviewees who tend to reveal their deceit verbally (Talwar & Lee, 2002). Furthermore, these unanticipated findings might be explained by differences in children’s developing cognitive capabilities that are associated with lie-telling ability, such as executive functioning (Talwar & Crossman, 2011). Child lie-tellers in our study may have had good working memory skills that allowed them to look at the interviewer’s face whilst telling their false report. Future research should investigate whether the effects of imposing cognitive load are moderated by children’s growing cognitive development.

Interestingly, truth-tellers instructed to look at the interviewer’s face did experience some dual-task interference. This unexpected finding requires further investigation. As memory can be data-limited (i.e. limited by a person’s ability to recall a past experience), it could be that factors other than gaze compliance influenced our child truth-tellers’ ability to provide a detailed account. Finally, the absence of dual-task interference for children instructed to look at the teddy bear’s face could be explained by the teddy bear’s face not being as cognitively effortful to look at as the interviewer’s face. As the teddy bear’s face did not provide any relevant feedback, it was not necessary for the interviewees to monitor it for suspicion. Nevertheless, we suggest caution in interpreting these correlations due to their non-significant nature.
Irrespective of the findings for dual-task interference, significant differences in level of detail between child truth-tellers and child lie-tellers were only elicited when a secondary task was imposed. For children instructed to look at the interviewer’s face, these findings are in line with previous work with adults (Vrij et al., 2010), which has also found exaggerated behavioural differences between truths and lies when gaze was maintained. For children instructed to look at the teddy bear’s face, these findings extend current knowledge and demonstrate that gazing towards a non-human stimulus could act as a less threatening, but still effective, substitute in practice. Although exaggerated differences occurred when a dual-task was imposed, it remains unclear from a theoretical standpoint why this was the case. The dual-task processes involved in providing a narrative and maintaining gaze require further examination to understand the theory behind this effect. Indeed, further probing of the significant interaction suggests that using different gaze instructions does not have an effect on false reports but rather has an effect on true reports. Thus, these exaggerated differences could be due to gaze maintenance facilitating longer truthful accounts rather than inhibiting false accounts. Our findings suggest that the request to look at the interviewer’s face elicited true reports that were significantly more detailed than when no gaze instruction was provided. This may be due to the demeanour of our interviewer; supportive interviewers have been shown to elicit longer reports (Vrij, 2015). However, it is not within the scope of this research to draw any firm conclusions regarding these results. Furthermore, these findings should be interpreted with caution. Due to small experimental cell sizes, there is a risk of Type I error. This study, therefore, requires replication with a larger sample size to verify that the interaction effect remains significant.

In this study we were not able to examine the memory accuracy of the truth-tellers’ detailed reports. Based on the information provided by the schools, we were only able to
establish whether the children had taken part in the events or not, but, due to the scope of the events, we were unable to capture all of the information regarding the events to code for correct and incorrect details. Future research is required to explore the relevance and accuracy of the reports provided by truth-tellers in the ‘gaze instruction’ conditions to understand the specific benefits of eliciting more details in true reports.

In the current study the interview protocol was short and non-elaborative. Using open-ended questions did allow us to go beyond the majority of past research, which has primarily focused on forced-choice questions using temptation resistance paradigms, to examine how gaze maintenance would affect children’s longer narratives. However, this does not reflect interview protocols in real-life police investigations with child witnesses, where a variety of question types are used. We can, therefore, not generalize these findings to a whole police interview, but only to the beginning of the police interview where an uninterrupted free narrative is requested. Finally, our study represents a ‘best case scenario’ in which a child provides a long narrative. As we reduced our interview protocol to focus on two open-ended questions, it was important to facilitate long responses by providing all of the children with examples of the type of information they could provide and some time to prepare. Child witnesses typically provide shorter statements than both their adolescent and adult counterparts (Jack, Leov & Zajac, 2014); this may be due to them not knowing what level of detail is required at interview (Lamb, Orbach, Hershkowitz, Esplin & Horowitz, 2007). Future research should continue to test the generalizability of these findings by using a procedure where no examples are provided.

Despite the exaggerated difference in level of detail elicited between child truth-tellers and child lie-tellers in the dual-task gaze condition (compared to the single-task control condition), the major concern still remained whether evaluators would be able to
discriminate between lie-tellers and truth-tellers more effectively when child interviewees were instructed to maintain gaze compared to when no gaze instructions were given. We investigated this issue in Experiment 2.

2.4 Experiment 2

In Experiment 2, we tested the prediction that evaluators would discriminate better between truth-tellers and lie-tellers instructed to maintain gaze, than truth-tellers and lie-tellers who were given no gaze instruction (Hypothesis 3).

We also examined whether telling evaluators that truth-tellers provide more detail in their reports than lie-tellers would improve discrimination accuracy. Previous research into training to improve lie detection has shown that informing evaluators about empirically-supported verbal cues to deceit has the largest effect on their detection accuracy (Hauch, Sporer, Michael & Meissner, 2014). Overall, level of detail has been found to be a key indicator of veracity (DePaulo et al., 2003). It is also one of the general characteristics coded for in Criteria-Based Content Analysis (Steller & Köhnken, 1989) that has received the most support for distinguishing between child truth-tellers and child lie-tellers in the predicted direction (Vrij 2005). It was, therefore, anticipated that evaluators who received this guidance regarding detail would demonstrate better discrimination than evaluators who received no guidance (Hypothesis 4). It was further predicted that an improvement in discrimination, as a result of guidance, would be most pronounced when judging the credibility of children instructed to maintain gaze, due to a greater difference in detail being elicited in these conditions in Experiment 1 (Hypothesis 5).

Successful discrimination depends on whether evaluators can interpret behavioural cues correctly. It was, therefore, important to recognise that gaze aversion can be
perceived as a strong indicator of deception (Global Deception Research Team, 2006), even though this cue is non-diagnostic (DePaulo et al., 2003). We could not rule out the possibility that gaze behaviour perceived to be somewhat ‘strange’ might impact on evaluators’ judgments of credibility. Half of the evaluators were, therefore, played visual-audio clips of the children’s interviews, and the other half were played audio-only clips. We anticipated that evaluators who watched the visual-audio presentations displaying the gaze maintenance behaviour would demonstrate a truth bias because gaze maintenance might be interpreted as a sign of truthfulness (Vrij et al., 2010) (Hypothesis 6).

2.4.1 Method

2.4.1.1 Participants

A sample of 192 adult evaluators (89 males, 103 females) with an age range of 18 to 76 years ($M = 27.14$ years, $SD = 11.71$ years) was recruited. One hundred and ten participants (52% of the total sample) were undergraduate students who received 0.5 course credit for their participation. The further 82 participants were members of the general public recruited via convenience sampling. The non-student participants were not compensated for their participation.

2.4.1.2 Interview clips

A total of 30 interview clips were selected from the sample of 85 children in Experiment 1. There were ten clips per ‘Gaze Instruction’ condition; within each of those three sets of ten clips, there were five truth-tellers and five lie-tellers. In the first round of the interview clip selection process, all recordings that contained noise interference (e.g. school bell, road traffic) were excluded ($n = 20$). Second, clips in which the first free recall lasted longer than 300 seconds were removed ($n = 7$). This criterion was chosen to limit the total duration of the study (50 minutes maximum), reducing potential fatigue effects on
evaluators’ performance. The remaining 58 clips were divided by Gaze Instruction condition (IF, n = 22; TF, n = 16; CONTROL, n = 20), and five truth-tellers and five lie-tellers were randomly selected for each condition. The final thirty clips were edited down so that they only contained the child interviewee’s first free recall. This selection process resulted in an even distribution of gender (3 boys to 2 girls, or 2 boys to 3 girls) in each Veracity x Gaze Instruction cell, except for the false reports in the ‘control’ condition, which were all provided by boys. It was not anticipated that this would bias results as no response bias has been previously found for adults judging boys’ credibility (Talwar, Crossman, Gulmi, Renaud & Williams, 2009). Interview clips lasted from 53 seconds to 239 seconds (M = 135.67 seconds, SD = 56.16 seconds). A 2 (Veracity) x 3 (Gaze Instruction) ANOVA was performed to ensure that there were no significant differences in length of clip across conditions. There was no significant main effect of Veracity, F(1, 24) = .13, p = .72, no significant main effect of Gaze Instruction, F(2, 24) = .05, p = .96, and there was no significant Veracity X Gaze Instruction interaction effect, F(2, 24) = .62, p = .55. For each ‘gaze instruction’ condition, four random rotations of the ten clips were created to reduce order effects.

2.4.1.3 Guidance on detail

Evaluators who received guidance were provided with a sheet stating that truth-tellers provided more detail overall in their accounts compared to lie-tellers, as this has been reported in previous deception research (DePaulo et al., 2003) and was also found in Experiment 1. To help evaluators understand what the experiment meant by the term ‘detail’, five different types of detail were presented in a table. For each type of detail, a description and an example of that detail were provided (i.e. ‘visual detail refers to what the interviewee said that they saw. For example, a red hat contains two visual details’). Participants were advised to refer back to the guidance sheet as much as they found useful
when watching/listening to the interview clips and were able to ask the experimenter for clarification on these types of detail before and during the experiment.

2.4.1.4 Procedure

The study took place in a quiet environment with few distractions. In order to prevent evaluators from working on the assumption that they would be presented with equal numbers of truth-tellers and lie-tellers, two steps were taken. First, participants were informed that they would be asked to evaluate the veracity of twelve child interviews in turn (actually they only evaluated ten clips in total). Second, they were told that it was just as likely for a child to be telling the truth as it was for them to be telling a lie.

First, evaluators were randomly assigned to a Gaze Instruction condition. That is, they judged the credibility of ten interview clips (five truth-tellers and five lie-tellers) from only one of the Gaze Instruction conditions in Experiment 1 (IF vs. TF vs. Control). Evaluators who were provided with guidance on detail received this at the beginning of the experiment. Half of the evaluators watched all of the interview clips in visual-audio format, whilst the other half listened to all interview clips in audio-only format. Participants who watched visual-audio presentations of the interviewees in the ‘interviewer’s face’ and the ‘teddy bear’s face’ conditions were informed that the child interviewees had been asked by the experimenter to direct their gaze during the interviews. Evaluators then watched and/or listened to the clips, one at a time, via a computer. Headphones were provided. To record their credibility judgments, evaluators were given a hard copy answer booklet. Following each interview clip, evaluators were asked to decide if the child interviewee was lying or telling the truth.

Participants’ dichotomous judgments (truth or lie) for each clip were used to measure hits (proportion of deceitful clips correctly identified as deceitful) and false alarms.
(proportion of truthful clips incorrectly identified as deceitful) for subsequent signal
detection analysis.

2.4.2 Results

2.4.2.1 Accuracy

Overall accuracy ($M = 51.72\%, SD = 16.23$) was not significantly different from
chance, $t(191) = 1.47, p = .14$, but truth accuracy ($M = 60.62\%, SD = 20.56$) was
significantly above chance, $t(191) = 7.16, p < .001, d = .52 (95\% CI [.37, .67])$, and lie
accuracy ($M = 42.81\%, SD = 21.23$) was significantly below chance, $t(191) = -4.69,
p < .001, d = .34 (95\% CI [.19, .48])$. When evaluators judged the credibility of children
instructed to look at the interviewer’s face ($M = 58.91\%, SD = 16.44$), they performed
significantly better than chance, $t(63) = 4.33, p < .001, d = .54 (95\% CI [.28, .80])$. When
judging children instructed to look at the teddy bear’s face ($M = 47.97\%, SD = 15.45$) or
children given no gaze instruction ($M = 48.28\% SD = 14.54$), they were no better than
chance ($p s > .05$). Moreover, when evaluators were guided to look out for differences in
detail ($M = 53.96\%, SD = 17.07$), they were better than chance, $t(95) = 2.27, p = .025, d = .23 (95\% CI [.03, .43])$, but not when no guidance was provided ($M = 49.48\%, SD = 15.11), t(95) = -.34, p = .74.

2.4.2.2 Signal detection analysis

The application of signal detection theory to deception detection research has been
largely recommended because it provides an opportunity to measure two conceptually
different parameters of accuracy (Meissner & Kassin, 2002): discrimination accuracy -
ability to discriminate lie-tellers from truth-tellers (in this experiment, referred to as $d’$),
and response bias – tendencies to favour a particular response (truth or lie), in this
experiment, referred to as $\beta$. Means and standard deviations for discrimination accuracy and response bias across all conditions are displayed in Table 2.2.

2.4.2.2.1 Discrimination accuracy

A 3 (Gaze Instruction) x 2 (Guidance Provision) ANOVA was performed with participants’ sensitivity scores ($d'$) as the dependent variable to examine their ability to discriminate between truth- and lie-tellers.

First, there was a significant main effect of Gaze Instruction, $F(2, 180) = 10.84$, $p<.001$. Post-hoc analyses using Bonferroni adjustment revealed that evaluators discriminated better between children’s truthful and deceptive accounts when the interviewees were instructed to look at the interviewer’s face compared to when the interviewees were instructed to look at the teddy bear’s face, $p<.001$, $d = .66$ (95% CI [.30, 1.02]), and when the interviewees were given no particular gaze instruction, $p<.001$, $d = .67$ (95% CI [.32, 1.03]). Evaluators’ performance did not differ significantly between those instructed to look at the teddy bear’s face and for those given no instruction ($p = 1.00$).

Second, there was a significant main effect of Guidance Provision, $F(1, 180) = 4.20$, $p = .042$. Pairwise comparisons using Bonferroni adjustment showed that evaluators who received guidance discriminated better between veracity groups than evaluators who received no guidance, $d = .27$ (95% CI [-.014, .55]).

Finally, there was a significant Gaze Instruction X Guidance Provision interaction effect, $F(2, 180) = 4.88$, $p = .009$. We performed univariate analyses to test the effect of providing guidance within each Gaze Instruction condition. There was a significant main effect of Guidance Provision for evaluators judging the credibility of child interviewees instructed to look at the teddy bear’s face, $F(1, 62) = 12.10$, $p = .001$. For evaluators in the ‘teddy bear’s face’ condition, those who received guidance ($M = .22$, $SD = .76$) were able
to discriminate better than those who received no guidance ($M = -.38, SD = .63, d = .87$ (95% CI [.35, 1.38])). There was no significant main effect of Guidance Provision for evaluators assigned to the ‘interviewer’s face’ condition, $F(1, 62) = 1.27, p = .26$, or the ‘control’ condition, $F(1, 62) = 1.15, p = .29$. There were no other significant interaction effects ($p$-values > .05).

In a second level of analysis, $d'$ values were compared to 0 (no ability to differentiate between children’s truths and lies) using one-sample $t$ tests. With regard to Gaze Instruction, evaluators could reliably discriminate child truth-tellers from child lietellers in the ‘interviewer’s face’ condition, $t(63) = 4.32, p < .001, d = .54$ (95% CI [.28, .80]), but not in the ‘teddy bear’s face’ condition, $t(63) = -.87, p = .39$, nor the ‘no gaze instruction’ condition, $t(63) = -.87, p = .38$. For Guidance Provision, evaluators were able to discriminate reliably when provided with guidance, $t(95) = 2.30, p = .024, d = .23$ (95% CI [.03, .44]), but not when guidance was withheld, $t(95) = -.20, p = .84$.

Finally, we compared $d'$ scores to 0 for the significant interaction between Gaze Instruction and Guidance Provision. When evaluators judged the credibility of children instructed to look at the interviewer’s face, they were able to discriminate lie-tellers from truth-tellers whether guidance was provided ($M = .55, SD = .85), t(31) = 3.63, p = .001, d = .64$ (95% CI [.26, 1.02]), or not ($M = .32, SD = .75), t(31) = 2.43, p = .021, d = .43$ (95% CI [.063, .79]). For children instructed to look at the teddy bear’s face, evaluators were not able to discriminate between children’s truths and lies when provided with guidance ($M = .22, SD = .76), t(31) = 1.65, p = .11$, nor when there was no guidance provision ($M = -.38, SD = .63), t(31) = -3.46, p = .002, d = .61 (95% CI [.23, .98]). That is, evaluators labelled the groups incorrectly (i.e. they tended to label lie-tellers as truthful and truth-tellers as deceitful). Finally, when children were given no gaze instructions, evaluators were not able to discriminate truthful from fabricated reports, with guidance provision, ($M = -.17,$
$SD = .78$), $t(31) = -1.25$, $p = .22$, or without guidance provision, $(M = .018, SD = .64)$, $t(31) = .16$, $p = .88$.

### 2.4.2.2.2 Response bias

Participants’ response bias ($\beta$ scores) was investigated to see whether they tended to identify children as lie-tellers or truth-tellers in any particular condition. A three-way ANOVA, with Gaze Instruction, Guidance Provision and Modality of Presentation of the clips as between-subjects factors, revealed significant main effects of Gaze Instruction, $F(2, 180) = 5.05$, $p = .007$, and Modality of Presentation, $F(1, 180) = 6.55$, $p = .011$. First, responses were more biased when judging the credibility of children instructed to look at the interviewer’s face ($M = 1.21, SD = .49$) compared to children instructed to look at the teddy bear’s face ($M = 1.02, SD = .37$), $p = .020$, $d = .46$ (95% CI [.10, .81]), and children given no particular gaze instruction ($M = 1.01, SD = .38$, 95% CI [.92, 1.11]), $p = .019$, $d = .45$ (95% CI [.10, .80]). Response bias did not significantly differ between evaluators judging child credibility in the latter two gaze conditions ($p = 1.00$). Second, evaluators demonstrated more bias in the ‘audio-only’ condition ($M = 1.16, SD = .48$) than in the ‘video-audio’ condition ($M = 1.01, SD = .35$), $d = .36$ (95% CI [.07, .64]). There was no significant main effect of Guidance Provision and there were no significant interaction effects ($p$-values >.10).

Using one-sample $t$ tests, each $\beta$ was compared to 1 (no bias). In signal detection theory, $\beta$ values below 1 signify a tendency to respond yes (or lie in the current study), whereas values above 1 signify a tendency to respond no (or truth in the current study; Stanislaw & Todorov, 1999). Therefore, the subsequent analyses examined the existence and the nature of the bias. With regard to Gaze Instruction, evaluators who judged the credibility of children instructed to look at the interviewer’s face were significantly biased to respond ‘truth’, $t(63) = 3.46$, $p = .001$, $d = .43$ (95% CI [.18, .69]), whereas no
significant response bias was found for evaluators who judged children instructed to look at the teddy bear’s face, \( t(63) = .35, p = .73 \), nor for evaluators who judged children in the ‘no gaze instruction’ condition, \( t(63) = .30, p = .77 \). In terms of Modality of Presentation, evaluators in the ‘audio only’ condition displayed a significant truth bias, \( t(95) = 3.18, p = .002, d = .33 \) (95% CI [.12, .53]), whereas evaluators in the ‘video-audio’ condition showed no bias, \( t(95) = .17, p = .87 \).

### 2.4.3 Discussion

Table 2.2

**Discrimination Accuracy (d’) and Response Bias (β) as a Function of Gaze Instruction, Guidance Provision and Modality of Presentation**

<table>
<thead>
<tr>
<th></th>
<th>( d' )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td><strong>Gaze Instruction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Look at interviewer's face</td>
<td>.43***</td>
<td>.80</td>
</tr>
<tr>
<td>Look at teddy bear's face</td>
<td>-.08</td>
<td>.75</td>
</tr>
<tr>
<td>No instruction (control)</td>
<td>-.08</td>
<td>.71</td>
</tr>
<tr>
<td><strong>Guidance Provision</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>.20*</td>
<td>.84</td>
</tr>
<tr>
<td>No</td>
<td>-.02</td>
<td>.73</td>
</tr>
<tr>
<td><strong>Modality of Presentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video-audio</td>
<td>-.04</td>
<td>.77</td>
</tr>
<tr>
<td>Audio only</td>
<td>.22**</td>
<td>.80</td>
</tr>
</tbody>
</table>

*Note. Statistical tests compared \( d' \) to 0 and \( \beta \) to 1.*

\( * p<.05 \quad ** p<.01 \quad *** p<.001 \)

Instructing child interviewees to maintain gaze with the interviewer’s face enabled evaluators to discriminate between true and false reports to a better degree than when no
instruction was given, in spite of a significant truth bias. However, discrimination accuracy was not affected when child interviewees were instructed to gaze towards the teddy’s bear face. Thus, Hypothesis 3 was partially supported. The ability to accurately detect deception for evaluators rating children instructed to gaze at the interviewer’s face may be due to differences in details provided by child truth-tellers and child lie-tellers. The cognitive lie detection approach posits that the ability to discriminate between truths and lies should increase with the activation and exaggeration of cognitive behavioural differences (Vrij, 2015). Considering that significant behavioural differences were elicited for both children instructed to look at the interviewer’s face and children instructed to look at the teddy bear’s face, it is possible that the exaggeration of these cues might need to reach a certain threshold, beyond which they become more apparent to an evaluator. It is possible that this threshold was only reached when child interviewees were instructed to look at the interviewer’s face, in turn, facilitating evaluators’ credibility judgments, but the threshold was not met when the children were asked to look at the teddy bear’s face.

Informing evaluators that truth-tellers provide more detailed reports compared to lie-tellers did improve their ability to detect deception, thus supporting Hypothesis 4. However, it is difficult to conclude to what extent evaluators applied this guidance to the interview clips. Although training in verbal content cues is recommended because it leads to the highest training effects, it is also important to note that false information regarding cues to deceit can work as effectively as true information (Hauch et al., 2014). To encourage evaluators to engage more with the guidance and base their final credibility judgments on this specific information, it would be better to use methods such as the Psychologically Based Credibility Assessment Tool (Evans, Michael, Meissner & Brandon, 2013) that include the rating of diagnostic cues in the final credibility assessment.
Contrary to Hypothesis 5, the provision of guidance was not more beneficial when judging children who were instructed to maintain gaze compared to those in the ‘control’ condition. Indeed, the only benefit of providing guidance was that it protected evaluators in the ‘teddy bear’s face’ condition from incorrectly labelling child veracity. As children in this condition were neither maintaining eye contact, nor free to look where they wished, their ‘strange’ gaze behaviour of looking at the interviewer’s lap might have been interpreted incorrectly as suspicious. Directing evaluators’ attention towards what the child was saying, through the use of our guidance, and encouraging them to base their credibility judgments on the child’s verbal behaviour, may have detracted from the misinterpretation of their ‘strange’ gazing towards the teddy bear.

Finally, although we predicted in Hypothesis 6 that evaluators who watched the visual-audio presentations displaying the gaze maintenance behaviour would demonstrate a truth bias, this was not the case. This lack of truth bias might be due to evaluators interpreting gaze maintenance behaviour differently from that suggested by the general deception literature. On the one hand, gaze aversion is believed to be a cue to deceit (Global Deception Research Team, 2006), but, on the other hand, nonverbal behaviour that deviates from the expected norm, such as staring, can also be perceived to be ‘fishy’ (Bond et al., 1992). It is not known to what extent gaze behaviour influenced evaluators’ judgments, or how much suspicion evaluators attached to this nonverbal cue; however, the lack of bias might suggest that opposing interpretations may have cancelled each other out. Alternatively, informing evaluators that children had been instructed to divert their gaze may have made them more aware of their own bias.

For the current study evaluators were exposed to ten interview clips. This may have led to evaluators comparing cues and information across interviews. In real police investigations and court proceedings, it is likely that these comparisons will occur between
children’s statements, adult’s statements and physical evidence. Future research should try to replicate this scenario to understand how a police officer or juror might judge the credibility of a child both in isolation and in comparison to other sources.

2.5 General Discussion

We conducted the first empirical investigation exploring the use of gaze maintenance to detect deception in child witnesses during investigative interviews. Similar to Vrij et al. (2010), we predicted that the interview strategy would magnify differences in level of detail between children’s true and false reports. We also expected that the exaggeration of this cue would facilitate evaluators’ ability to discriminate children’s lies from truths.

The present findings show that gaze maintenance can be effective for determining the credibility of child witnesses. In Experiment 1, lie-tellers provided significantly fewer details in their reports compared to truth-tellers but only when they were instructed to look towards either the interviewer’s face or a teddy bear’s face. No significant difference was elicited when a secondary task was absent. In Experiment 2, we found that the exaggeration of this diagnostic cue facilitated evaluators’ discrimination accuracy, but this was only when children were instructed to look at the interviewer’s face.

Theoretically, the effect of imposing a secondary task on interviewee performance remains unclear. The findings of Experiment 2 make it difficult to discern whether the secondary task had any negative impact on truth-tellers’ memory or whether lie-tellers experienced any additional cognitive load. The latter issue may be due to the nature of the secondary task in this study and the difficulty in pinning down the exact cognitive mechanisms involved. As previously mentioned, the development of certain cognitive skills is closely linked to children’s proficiency to tell and maintain lies (Talwar &
Crossman, 2011). It may therefore be wise, in future, to provide cognitive measures of the specific executive functions that the imposed secondary task aims to affect to be able to establish whether (a) there is a link between these cognitive skills and the performance on the tasks, and (b) whether children’s ability to perform these cognitive skills predicts the effectiveness of imposing cognitive load. When testing dual-task methodologies, it would also be beneficial to obtain baseline measures of an individual’s performance on single tasks (Task A only and Task B only) to which their performance on a dual-task (Tasks A and B simultaneously) could be compared.

Our findings provide further support for the practical value of manipulating cognitive load as a potential means for discriminating between children’s true and false reports. In particular, the results demonstrate that the effects of imposing cognitive load are not limited to asking children to tell their stories backwards. This is beneficial because Saykaly et al. (2016) found that reverse order recall can adversely affect the accuracy of both truthful and deceptive statements, suggesting that it might not be helpful in real police investigations. In our study, requiring child interviewees to perform the secondary task of maintaining gaze had a positive effect on truth-tellers, eliciting more information from them than when no gaze instruction was given. This finding is in line with the primary goal of any investigative interview, which is to elicit as much information as possible from the interviewee. This finding could be due the interviewer’s supportive demeanour, which has been found with adults to elicit more details from truth-tellers than lie-tellers (Mann et al., 2013). Further investigation is required to determine whether it is the combined effect of a gaze maintenance instruction to witnesses and supportive interviewer behaviour that helps truth-tellers but not lie-tellers, rather than the technique on its own.

A practical limitation of using gaze maintenance with child interviewees may be its appropriateness in certain contexts. Maintaining gaze with an authoritative figure, such as
a police officer, might be an intimidating task for children. Although none of the children instructed to look at the interviewer’s face reported any discomfort, the average child did not maintain gaze for more than half of their interview. A recent school event is far less traumatic to talk about than incidents of physical and/or sexual abuse, which can be the main focus of police investigations involving child witnesses. Future research must examine the scope of the beneficial effects elicited in this study and balance them with potential discomfort in certain contexts. As such, the preliminary findings relating to an instruction to concentrate on the less intimidating teddy bear (or similar) should be extended.

Maintaining gaze, particularly with an interviewer’s face, is an effective strategy for judging the credibility of children. Future research should continue to explore the application of dual-task processing to child interviews by examining strategies that target children’s under-developed executive functioning, with a view to creating more appropriate secondary tasks for this potentially sensitive context.

2.6 References


10.1037/a0015808


10.3758/BF03211385


10.1111/cdev.12301


10.1177/00993650214534974


Chapter 3: ‘Truly crime-relevant’: Using a model statement in an eyewitness context to differentiate between children’s true and false reports
3.1 Abstract

Encouraging interviewee productivity using a model statement (i.e. by providing an example of a detailed report) elicits more truthful information as well as more verbal cues to veracity in adults. The aim of this study was to investigate whether using a model statement could also create ‘productive’ child interviewees whilst at the same time helping to exaggerate differences between their true and false reports in an eyewitness context.

Sixty-three children aged 8 to 10 years old were interviewed about the staged theft of a mobile phone that they had (truth-tellers), or had not (lie-tellers) witnessed. At interview, children provided two free recall attempts: a baseline recall, pre-model statement, and a second recall, post-model statement. Both truth-tellers’ and lie-tellers’ statements were longer and more detailed post-model statement than pre-model statement. True reports contained a higher proportion of crime-relevant information overall, and were consistently more forensically useful (i.e. included a higher proportion of repeated crime-relevant information) than false reports. In terms of the quality of truth-tellers’ eyewitness memory, there was no difference in the amount of correct information provided across recall attempts, but there was an increase in incorrect information, resulting in a decrease in memory accuracy after the model statement had been heard. Playing child interviewees a model statement increases their productivity, and preserves differences in the forensic usefulness of the information truth-tellers and lie-tellers provide as well as the consistency of this information. However, using a model statement in an eyewitness context highlights how children’s pursuit of being informative might cause them to sacrifice the accuracy of their report.
3.2 Introduction

From as young as four years old, children demonstrate impressive abilities to provide substantial amounts of forensically relevant information during interviews (Lamb et al., 2003). Nevertheless, when compared to both adolescents (14-16 years) and adults (25-60 years), children (9-11 years), although equally accurate, are found to provide statements that are significantly less detailed than their older counterparts (Jack, Leov & Zajac, 2014). The brevity of children’s reports is problematic for investigators, who then have fewer leads to follow up. This also puts the child witnesses at a disadvantage in court, as statements that are richer in detail are the most likely to be believed (Bell & Loftus, 1989; Johnson, 2006). Inaccurate judgments of children’s credibility are a major issue in both police and court settings, with legal professionals experiencing much difficulty when it comes to determining the veracity of children’s statements (Bala, Ramakrishnan, Lindsay & Lee, 2005; Leach, Talwar, Lee, Bala & Lindsay, 2004). New cognitive strategies, such as using a model statement, could facilitate this decision-making process by exaggerating the subtle differences that occur between true and false reports (Vrij, Fisher & Blank, 2015). This would be particularly beneficial in the early stages of an investigation to save both time and resources. Moreover, at the beginning of investigations, child truth-tellers’ reports are the most accurate (Bruck, Ceci & Principe, 2006), and differences in children’s truth- and lie-telling behaviours are most likely to occur (Saykaly, Talwar, Lindsay, Bala & Lee, 2013).

3.2.1 Using a model statement

A model statement refers to a recorded statement, presented to an interviewee, of a spontaneous, detailed recall of another person (in this study another child) describing his/her experience of an event that is unrelated to the witnessed event (Leal, Vrij, Warmelink, Vernham & Fisher, 2015) and is a cognitive lie detection technique aimed at
encouraging interviewees to provide more information (Vrij, 2015). The underlying assumption is that listening to a model statement will change both truth-tellers’ and lietellers’ views on what is expected from them at interview. In both an adult insurance context (Leal et al., 2015) and a security context (Ewens et al., 2016), model statements provided by a peer (i.e. another adult) have elicited more information from interviewees. The benefits of these more informative statements are two-fold: First, they elicit more truthful information, and second, false reports can become easier to detect. For example, Leal et al. (2015) found that the additional detail given by the liars made them appear less plausible. These findings highlight the potential for the model statement to resolve the aforementioned issues with child testimonies; namely, their brevity and their perceived credibility by legal professionals. It also suggests that using a model statement can be an effective truth detection technique as well as an effect lie detection technique.

3.2.2 Interviewee productivity

Interviewee productivity, here, refers to an interviewee’s response length as well as the amount of detail they provide in their report. When the model statement was used with adult interviewees, Leal et al. (2015) found that interviewee productivity increased for both truth-tellers and lie-tellers. However, the effect of playing the model statement may trigger different processes for truth-tellers and lie-tellers.

For truth-tellers, playing a model statement provides a good example of what is expected from them at interview. According to social comparison theory (Festinger, 1954) and social proof theory (Cialdini, 1993), in the absence of objective information, truth-tellers would be compelled to compare themselves to similar others. By presenting the person providing the model statement as a peer, truth-tellers can examine the peer behaviour (i.e. social proofing), determine whether their behaviour is appropriate in the same situation (i.e. social comparison), and, if necessary, correct their own behaviour.
accordingly. In other words, if interviewees do not understand what is required of them in an interview, they will be likely to use other sources, such as a model statement, as a point of reference. Using a model statement as a benchmark, interviewees determine whether what they planned to tell the interviewer is more detailed (upward social comparison) or less detailed (downward social comparison) than the model statement (Ewens et al., 2016). In the case of a downward social comparison, interviewees can then adjust their response by increasing the length and detail in their statement to match the model statement. For the current study, it was anticipated that children, who do not often know what is expected from them at interview (Lamb, Orbach, Hershkowitz, Esplin & Horowitz, 2007), would engage in this social comparison adapting their true reports by increasing their interviewee productivity like the adults in Leal et al.’s (2015) research.

For lie-tellers, however, using a model statement instils an internal strategic conflict (Leal et al., 2015; Vrij, 2015). On the one hand, lie-tellers wish to conceal their deceit by sticking to their originally planned story to prevent leakage of any incriminating information (e.g. contradictions, inconsistencies). On the other hand, lie-tellers engage in self-presentational behaviours concerned with actively creating an honest impression (DePaulo et al., 2003). Thus, they can choose to ignore the raised expectations of the model statement at the risk of looking incredible, or they can choose to meet the expectations of a longer statement at the risk of revealing their deceit. An increase in interviewee productivity found by Leal et al. (2015) suggests that adult participants prioritised an honest impression over a well-rehearsed story. Indeed, the lie-tellers in their study contradicted the typical lie-telling behaviour of saying as little as possible (Vrij, 2015), which is counterproductive as longer responses provide more opportunity for verbal cues to veracity to occur; for example, CBCA and RM scores (Vrij, Mann, Kristen & Fisher, 2007). Liu et al. (2010) reported similar behaviour by children who were asked
unanticipated questions. Lie-tellers were significantly more likely to go beyond their prepared stories and answer the unexpected questions than truth-tellers. Furthermore, child lie-tellers have reported using the addition of Detail as a frequent verbal strategy (Strömwall, Granhag, & Landström, 2007). This decision to say more might be particularly detrimental to child lie-tellers, who tend to reveal deceit through verbal behaviours (Akehurst, Manton & Quandte, 2011; Talwar & Lee, 2002). If child lie-tellers, in the current study, choose to go beyond their prepared baseline response (i.e. what they consider to be enough deceitful information to be believed but not so much that they can’t remember all the details in their working memory), then this could elicit more verbal indicators of deceit in their false reports.

3.2.3 Forensic usefulness

When encouraging interviewees to provide more information (i.e. to increase their interviewee productivity) during an investigative interview, it is essential that this increase in productivity equates to an increase in information focusing on the target incident (i.e. the details that are forensically useful). For application in the real world, it is important that children’s reports contain a high proportion of crime-relevant information thus facilitating an investigation, or acting as ‘best evidence’ in subsequent court proceedings. As Leal et al. (2015) conducted their study in the context of insurance claimants, it is not known whether using a model statement, compared to not using one, will lead to more crime-relevant details and an eyewitness account that will, therefore, be more useful in investigative settings. A recent meta-analysis of the cognitive lie detection approach with adults (Vrij et al., 2015) revealed that interview techniques that encourage interviewees to provide more information are the most effective for truth detection. This suggests that these techniques might exaggerate differences between truth-tellers and lie-tellers, by facilitating the cognitive process of truth-telling but not lie-telling. Hence, child truth-
tellers, who are already able to provide forensically useful information during interviews (Lamb et al., 2003), might, on hearing a model statement, be encouraged to give more of this crime-relevant information. On the other hand, regardless of whether or not they have heard a model statement, lie-tellers might choose to talk about the ‘gist’ of what happened due to the unfamiliarity of the target incident (Michael, 2013). False reports would, therefore, not be as crime-relevant as true reports.

### 3.2.4 Between-statement consistency

There are two opposing hypotheses in terms of the relationship between the consistency and veracity of child witnesses. First, the Pinocchio hypothesis posits that child lie-tellers will continue to elaborate their story with every opportunity; so like their metaphoric noses, their stories continue to grow in length. Bruck, Ceci and Hembrooke (2002) found support for this hypothesis when they asked children to provide either a true or false report of a staged event on five occasions across a six-week period. Across recall attempts, child lie-tellers added in more new information (commissions) than child truth-tellers, whereas child truth-tellers repeated more of the same details than child lie-tellers. Thus, the Pinocchio hypothesis predicts that children’s truthful statements will be more consistent (i.e. include more repetitions) than fabricated statements.

Second, the repeat versus reconstruct hypothesis (Granhag, Strömwall & Jonsson, 2003) postulates that child lie-tellers, more so than truth-tellers, are concerned with being consistent to avoid casting any doubt on their accounts. According to this hypothesis, false reports will, therefore, contain more repetitions than true reports. Furthermore, it posits that true reports will be affected by certain memory processes: Memory fade may result in truth-tellers omitting more information than lie-tellers and reminiscence may result in more new details being added across truthful statements compared to lies. Strömwall and Granhag (2005) found support for each of these predicted differences between children’s
true and false reports. Thus the repeat versus reconstruct hypothesis predicts that children’s truthful statements will be less consistent (i.e. include more *commissions* and *omissions*) than fabricated statements.

In the context of the current study, an increase in interviewee productivity is predicted for both lie-tellers and truth-tellers due to the use of a model statement. With the encouragement to include more details, lie-tellers are, therefore, afforded the opportunity that they seek, according to the *Pinocchio hypothesis*, to elaborate on their initial story. Yet, the model statement may also promote more detailed reports from truth-tellers, who, according to the *repeat versus reconstruct hypothesis*, may enhance their account through reminiscence. For both veracity groups, the addition of new information after hearing a model statement would result in low internal consistency.

We were also interested to investigate how internal consistency interacted with forensic usefulness. If we anticipate that child truth-tellers, upon hearing a model statement, will increase both the crime-relevance of their account as well as the amount of new information they provide, then both these increases could combine to result in a higher proportion of crime-relevant commissions compared to lie-tellers (who will increase the amount of new information provided after a model statement but won’t increase the crime-relevance of this information).

### 3.2.5 Memory accuracy

It is important to ascertain whether longer statements elicited from truth-tellers result in a trade-off between quantity of information remembered and memory accuracy. Memory research with adults (Koriat & Goldsmith, 1996) as well as with children (Koriat, Goldsmith, Schneider & Nakash-Dura, 2001) has shown that sometimes an increase in quantity of reported details can result in a decrease in memory accuracy. The use of the
Cognitive Interview with both children and adults has been found to lead to an increase in incorrect details when accounts become more detailed (Köhnen, Milne, Memon & Bull, 1999). However, Köhnken et al. (1999) found children reported significantly fewer incorrect details than adults, and that the accuracy rates (percentage of correct details out of all details given in an interview) remained equal to those of a standard interview. It has therefore been argued that, “it is more desirable to collect 100 facts per interview at a 90% accuracy rate than to collect 10 facts at the same 90% accuracy rate” (McCauley & Fisher, 1995, p. 514). Using a model statements with truthful eyewitnesses only has found that it does not affect the accuracy of the information provided (Brackmann, Otgaar, Roos af Hjelmsäter & Sauerland, 2016). Further support for the absence of adverse effects of using a model statement on truth-tellers would continue to promote its practical significance.

3.2.6 Aims

The aims of the current experiment were to look at the effect of using a model statement on (i) interviewee productivity, (ii) forensic usefulness of true and false reports, (iii) internal consistency of true and false reports, and (iv) memory accuracy of true reports. First, it was hypothesised that, irrespective of veracity, interviewee productivity would increase following the model statement, resulting in longer and more detailed reports (Hypothesis 1). Second, it was anticipated that there would be a higher proportion of crime-relevant information (i.e. forensic usefulness) in truthful accounts than in fabricated accounts, and that there would be an interaction whereby this veracity difference would be exaggerated by using a model statement (Hypothesis 2). Third, it was expected that both truth-tellers and lie-tellers, having heard a model statement, would add in new information (commissions) during their second recall attempts. In other words, both child truth-tellers and child lie-tellers would demonstrate the same level of internal consistency (Hypothesis 3). However, we predicted that truth-tellers would provide a higher
proportion of crime-relevant commissions compared to lie-tellers (Hypothesis 4). Finally, we investigated whether the anticipated increase in detail given by truth-tellers post-model statement would have an effect on the overall accuracy of their accounts.

3.3 Method

3.3.1 Participants

Sixty-three children (26 males, 37 females) aged 8- to 10-years-old ($M = 9.00$ years, $SD = .70$ years) were recruited from one primary school. The majority of the child participants were White/Caucasian, with the remaining minorities (less than 5% of the total sample) being White Asian and Black African. Information sheets regarding the study were sent home to the children’s legal guardians who returned a signed written consent form agreeing to their children’s participation. Verbal assent was also obtained from the children. All children received a certificate and a stationery set in exchange for taking part.

3.3.2 Model statement

The model statement was an audio elicited from an interview with a ten-year-old girl from a previous study (Chapter 2). This audio extract was considered a model statement because it was a 230-word spontaneous free recall that, through coding for detail, had been shown to be the most detailed account from that dataset. Furthermore, following the guidance set out by Leal et al. (2015), the model statement was about an event (a school sports day) that was unrelated to the to-be-remembered event in question (a mobile phone theft). From the perspective of a source-monitoring framework (Johnson, Hashtroudi & Lindsay, 1993; Lindsay, Allen, Chan & Dahl, 2004), this conceptual dissimilarity between the ‘model statement event’ and the ‘to-be-remembered event’
should reduce the risk of information from the model statement interfering with the recall of the witnessed event.

3.3.3 Procedure

The experiment took place in two quiet areas within the school and involved three different people: the Principal Investigator (PI), an interviewer who was blind to the aims and hypotheses of the study, and a confederate who committed the mock theft. All child participants were naïve to the hypotheses of the study.

First, children were randomly assigned to a Veracity condition. Participants were yoked in terms of age and gender. Truth-tellers were invited to play a short board game with the PI. During the game, the PI went to the toilet and the confederate entered the room with the child, stole a mobile phone from a handbag nearby, and then left the room. The PI then returned and moments later found the phone was missing. Upon discovering its disappearance, the PI told the child that they would now be interviewed about what they had witnessed and were given approximately five minutes to think about what they would like to tell the interviewer. This procedure was standardized to enable the coding of correct and incorrect information. Children were not forewarned that a theft would occur during the game. The participants were then escorted to the interview room.

Lie-tellers, however, did not play a game with the PI, nor did they witness the theft of a mobile phone. They were informed that the PI had stolen her friend’s phone and were asked to ‘cover up’ the transgression by providing a false account of witnessing someone else taking the mobile phone. This set up reflects the ‘Child Absent’ condition in Talwar, Lee, Bala and Lindsay (2004)’s paradigm where the child was not present when a transgression took place and was later asked to cover up for the ‘perpetrator’. To aid their preparation, lie-tellers were provided with a crib sheet that outlined the standardized
structure of the truthful event and were encouraged to think about any additional details they could provide to embellish their statement (e.g. how would you describe the person who stole the phone?). After five minutes of preparation time for all participants, they were escorted to the interview room.

In both conditions, participants were told that it was very important that the interviewer thought that they were telling the truth. All participants were told that if the interviewer believed their story, then they would receive a certificate and a reward, but if they were not believed to be telling the truth, then they would have to write out their statement in full.

All children underwent the same interview protocol with the same interviewer. This consisted of the two components of the Cognitive Interview (Fisher & Geiselman, 1992) that are most frequently used by police officers; building rapport and an uninterrupted free narrative (Dando, Wilcock & Milne, 2008). Participants were instructed to refrain from guessing. Following an initial rapport-building phase (that took place off camera), all participants were informed that the interviewer was there to investigate the theft of the mobile phone and that the purpose of the interview was to obtain as much information as possible regarding the theft. Participants were then asked to provide two free recall attempts. The first recall attempt acted as a baseline measure (pre-model statement/control) where the interviewer asked the children to tell her everything that happened during their time with the PI in as much detail as possible. Following this, the interviewer introduced the model statement:

“Now, I know it’s sometimes difficult to know what I mean when I ask for lots of detail. So I’m going to play you a short recording of a girl talking about a school sports day. Listen carefully to what she says and to all the details she includes.”
Once the model statement had been played, the interviewer requested a second free recall attempt (post-model statement), which was later compared to the pre-model statement recall attempt (baseline):

“The girl you just heard included lots of detail about everything. Please could you tell me again, in as much detail as you possibly can, everything that happened when you were playing with [PI]? Don’t worry about repeating yourself; it’s fine to say what you said before and try to add in as much detail as you can. Tell me everything! Try to include as much detail as the girl you heard and tell me again everything that happened with you and [PI].”

In the final phase of the experiment, participants returned to the PI who asked them to state whether they had been instructed to tell the truth or tell a lie. All participants successfully answered this manipulation check. Regardless of performance, all of the children were given stationery sets as a reward for taking part. All interviews were transcribed for coding of verbal cues.

3.3.4 Coding schemes

For interviewee productivity, forensic usefulness and internal consistency, two independent raters coded all pre- and post-model statement recall attempts. One rater coded all transcripts and a second rater coded a random sample of 20 transcripts (31.7% of the total number of statements). For memory accuracy, the same first coder scored all transcripts and the second coder rated a random sample of 15 truth-tellers’ transcripts (46.9% of the total).

3.3.4.1 Interviewee productivity

First, both recall attempts were coded for total number of overall details. To make the coding more precise, all transcripts were coded for five different types of details; visual
(e.g. “dark brown hair” contains three visual details), auditory (e.g. “she told me she was borrowing the phone” contains one auditory detail), spatial (e.g. “the bag was next to the wall” contains one spatial detail), temporal (e.g. “first she rolled the dice then she moved her counter” contains two temporal details), and action (e.g. “we played a game” contains one action detail). Total number of details was the sum of the five types of detail per recall attempt. Intra-class correlation coefficients (ICCs) were calculated between the two coders based on the coded cues in the whole transcript for each child (both pre- and post-model statement). Inter-rater reliability was high, with all ICCs demonstrating high levels of agreement between coders (visual details, ICC = .98; auditory details, ICC = .99; spatial details, ICC = .93; temporal details, ICC = .97; action details, ICC = .98; and total number of details, ICC = .99).

3.3.4.2 Forensic usefulness

Each detail that was coded above was then coded as either crime-relevant (related to the criminal event, such as details about the perpetrator and her actions), or crime-irrelevant (unrelated to the criminal event, such as playing the board game). Proportions of crime-relevant details were calculated by dividing the total number of crime-relevant details by the total number of details recalled and then multiplying this by 100 for total number of details per recall attempt. High percentages reflected more forensically relevant reports that would be more useful to a police investigation or court case. The Intra-class correlation coefficient (ICC) was calculated between the two coders. Inter-rater reliability was high (% crime relevant details, ICC = .97).

3.3.4.3 Between-statement consistency

Requesting two free recall attempts from our child interviewees allowed us to record a baseline measure (pre-model statement) for truth-tellers and lie-tellers to which a second measure that had been experimentally manipulated (post-model statement) could be
compared. Details that were present in both the pre- and post-model statement free recalls were coded as repetitions. Commissions (information only provided post-model statement) were then calculated by subtracting the number of repetitions from the total number of details provided in the post-model statement recall attempt. Omissions (information only provided pre-model statement) were also calculated by subtracting the number of repetitions from the total number of details provided in the pre-model statement recall attempt. Intra-class correlation coefficients (ICCs) were calculated between the two coders. Inter-rater reliability was high, with all ICCs demonstrating high levels of agreement between coders (repetitions, ICC = .97; commissions, ICC = .96; and omissions, ICC = .97). Repetitions, commissions and omissions were further divided into crime-relevant and crime-irrelevant categories (according to the coding in the above section). Intra-class correlation coefficients (ICCs) were calculated between the two coders. Inter-rater reliability was high (% crime-relevant repetitions, ICC = .97; % crime-relevant commissions, ICC = .95; and % crime-relevant omissions, ICC = .90).

3.3.4.4 Memory accuracy

A list of information that could be provided about the perpetrator, the theft and the surroundings, and that could be coded as correct or incorrect with certainty, was constructed based on the standardized, staged theft. Any crime-relevant information that participants provided that was not already on the list was added, resulting in a total number of eight-sevent details that could be correctly reported. The crime-relevant information provided in the transcripts of the thirty-two truth-tellers was scored with reference to this list as correct or incorrect. Intra-class correlation coefficients (ICCs) were calculated between the two coders. Inter-rater reliability was high, with all ICCs demonstrating high levels of agreement between coders (correct details, ICC = .96; and incorrect details, ICC = .94). Memory accuracy, as referred to by Koriat and Goldsmith (1996), was calculated by
dividing the total number of correct responses provided by the total number of details (correct details + incorrect details) provided for each participant, multiplied by 100.

3.4 Results

3.4.1 Interviewee productivity

A mixed MANOVA was performed to test changes in length (in words) and level of detail across recall attempts as a function of Veracity (Hypothesis 1). There was a significant main effect of Recall Attempt on both response length, $F(1, 61) = 163.49, p<.001$, and level of detail, $F(1, 61) = 45.43, p<.001$. There was no significant main effect of Veracity, $F(1, 61) = 2.94, p = .09$, and no significant interaction effects ($ps >.05$).

Univariate analyses showed that the responses provided post-model statement ($M = 230.83$ words, $SD = 136.91$) were significantly longer than those provided pre-model statement ($M = 152.35$ words, $SD = 90.00$), $d = .62$. Interviewees also provided more details post-model statement ($M = 58.51$ details, $SD = 25.11$) than pre-model statement ($M = 42.08$ details, $SD = 21.15$), $d = .69$. According to these findings and as predicted by Hypothesis 1, interviewee productivity increased, irrespective of veracity, following the model statement.

3.4.2 Forensic usefulness

A mixed ANOVA, with Veracity and Recall Attempt (pre or post-model statement) as the independent variables, was used to examine differences in the proportion of crime-relevant details reported by interviewees. There was a significant main effect of Veracity, $F(1, 61) = 8.45, p = .005, d = .69$. Truth-tellers’ statements ($M = 61.89, SD = 13.01$) contained a higher percentage of crime-relevant details compared to lie-tellers’ ($M = 49.02, SD = 23.24$). There was, however, no main effect of Recall Attempt, $F(1, 61) = 1.65, p = .20$, and no interaction effect, $F(1, 61) = .03, p = .86$. The proportion of crime-relevant information included in accounts before and after the model statement did not change for
truth-tellers, nor did it change for lie-tellers. These findings provide partial support for Hypothesis 2: truth-tellers did provide more forensically useful reports than lie-tellers, but this difference was preserved, rather than exaggerated, by the model statement.

3.4.3 Between-statement consistency

A one-way MANOVA was performed with Veracity as the independent variable and with repetitions, commissions and omissions (coded from the post-model statement recall attempts) as the dependent variables. There was no significant effect of Veracity on number of repetitions, $F(1, 61) = .31, p = .58$, commissions, $F(1, 61) = 1.01, p = .32$, nor omissions, $F(1, 61) = .84, p = .36$. Thus, Hypothesis 3 was supported, as there was no difference between true and false reports in terms of the number of commissions (pieces of additional information provided) post model statement.

The crime-relevance of the children’s repetitions, commissions and omissions was tested using a one-way MANOVA, with Veracity as the independent variable. There was a significant effect of Veracity on the proportion of crime-relevant repetitions, $F(1, 61) = 9.97, p = .002$. Pairwise comparisons using Bonferroni adjustment\(^1\) showed that truth-tellers ($M = 62.31, SD = 22.40$) repeated a higher proportion of crime-relevant information compared to lie-tellers ($M = 44.12, SD = 23.32$), $p = .002, d = .80$. There was no effect of Veracity on proportion of crime-relevant commissions, $F(1, 61) = 2.04, p = .14$, nor on proportion of crime-relevant omissions, $F(1, 61) = 1.33, p = .27$. These findings do not support Hypothesis 4, as the only difference between true and false reports was in terms of the proportion of crime-relevant repetitions, not commissions.

\(^1\) Bonferroni adjustment throughout this thesis refers to the alpha level, which is set at .05, divided by the number of tests performed. This was performed to control for overall Type I error.
3.4.4 Memory accuracy

Paired samples t-tests were used to investigate the effect of the model statement on truth tellers’ accounts in terms of correct details, incorrect details and memory accuracy. There was no significant difference in the number of correct details provided pre-model statement ($M = 10.28, SD = 4.08$) and post-model statement ($M = 11.84, SD = 5.73$), $t(31) = 1.48, p = .15$. There was, however, a significant difference in the number of incorrect details, $t(31) = 3.43, p = .002, d = .61$. Truth-tellers gave more incorrect information post-model statement ($M = 2.81, SD = 2.46$) than pre-model statement ($M = 1.72, SD = 1.42$). As a consequence, memory accuracy, significantly decreased across the two recall attempts, $t(29) = 2.07, p = .04, d = .30$. Truth-tellers’ recall was more accurate before the model statement ($M = 86.02\%, SD = 10.60$) than after hearing the model statement ($M = 82.49\%, SD = 11.95$).

3.5 Discussion

In line with Leal et al. (2015), both truth-tellers and lie-tellers became more productive interviewees following the model statement. That is, they provided longer and more detailed statements in their second recall compared to their baseline attempt. No differences in the amount of details provided by truth-tellers and lie-tellers is surprising because lie-tellers only heard the story and were not exposed to other perceptual cues (e.g. seeing, smelling and feeling) that truth-tellers experienced. This difference in physical experience is what has previously led to the construction of verbal lie detection tools, such as Criteria-Based Content Analysis (CBCA) and Reality Monitoring (Vrij, 2008), which identified verbal indicators that could only pertain to real life experiences. In particular, accounts that are richer in detail have been consistently related to truthfulness (Vrij, 2005). In our study, lie-tellers had the challenging task to imagine the event described to them and
add to the information provided to put the blame on someone else. Their ability to imagine could explain why they were able to provide equally detailed accounts as truth-tellers. Imagination, referred to in the literature as fantasy proneness can affect the level of detail included in false reports, with high fantasy proneness associated with more detail-rich accounts (Merckelbach, 2004). Furthermore, high fantasy prone individual can even ‘fool’ verbal lie detection tools with their detailed false reports, achieving CBCA scores that are equal to and sometimes greater than those of true reports (Schelleman-Offermans & Merckelbach, 2010). Lie-tellers in our study could, therefore, have been particularly imaginative, overcoming this deficit caused by the design. Alternatively, poor truth-telling performance could also explain why lie-tellers didn’t have to perform as well to provide false reports that were as detailed as the true reports.

Overall, truth-tellers provided more crime-relevant information than lie-tellers. For truth-tellers, the model statement did not increase, but did preserve, the high proportion of crime-relevant information (and thus, forensic usefulness) included in their reports. Similarly, for lie-tellers, the model statement did not impact upon the amount of crime-relevant information provided; this remained significantly below the levels provided by truth-tellers. These findings highlight crime-relevance as a new indicator of children’s veracity and suggest that children’s true reports include more forensically useful information than children’s false reports, regardless of interview technique.

Our findings support the Pinocchio hypothesis (Bruck et al., 2002) for lie-tellers, and the repeat versus reconstruct hypothesis (Granhag et al., 2003) for truth-tellers, as there were no differences in the number of commissions for true and false reports. This would suggest that consistency was similar for both veracity groups. However, when crime relevance was taken into account across all measures of consistency, we did find that children’s true reports contained a higher proportion of crime-relevant repetitions.
compared to children’s false reports. The fact that truth-tellers demonstrated better internal consistency across the two recall attempts than lie-tellers for *forensically useful* information is important because witness consistency is one of the most important features of perceived witness credibility (Fisher, Brewer & Mitchell, 2009): Inconsistencies in crime-relevant information have the potential to destroy cases. Thus, as our truthful children were more consistent than our lie-telling children regarding details relating to the theft, this could increase the likelihood that they would be perceived as truthful. A model statement could, therefore, be a truth-detecting tool that can be used to elicit information and details from children in forensic settings.

Finally, although truth-tellers reported more crime-relevant detail than lie-tellers, the quality of this information decreased as a result of using the model statement. An increase in incorrect details across recall attempts, for truth-tellers, resulted in an overall decrease in memory accuracy, although this remained relatively high at 82%. No significant change in correct details elicited could be due to the children generally recalling low levels of the to-be-remembered information; on average, 10 details (out of 87) before the model statement and 11 details after the model statement. This could be attributed to truth-tellers being interviewed about something they had no forewarning would turn into the subject of an investigation; a common occurrence for real eyewitnesses. This means that they may not have paid the theft their full attention. Divided attention at encoding produces interference, leading to poor memory for stimuli (Fernandes & Moscovitch, 2000). These sub-optimal encoding conditions for truth-tellers could also explain the decrease in accuracy across the two free recall attempts. If truth-tellers did not pay the theft their full attention then they may have found the recall task particularly difficult, creating a floor effect. Furthermore, when they were later encouraged to say more after the model statement, they may have sacrificed the
accuracy of their statement in an attempt to be more informative (Ackerman & Goldsmith, 2008). Thus, even though the interview protocol discourages child interviewees from guessing, they may still have offered information that they were not truly confident was accurate in order to help the ‘investigation’.

A limitation of the current experiment was the absence of a ‘standard interview’ condition, in which children provided two free recalls without the presentation of a model statement between the recall attempts. As found by McCauley and Fisher (1995), the repetition of a standard interview can result in new information being elicited from child interviewees. In the present study, the child interviewees’ first free recall attempts were used as a control (baseline) to which subsequent changes in verbal behaviours were compared. However, this design does not enable us to rule out the effect of simple question repetition on our findings, particularly for the crime-relevant repetitions. Previous research has shown that using a model statement does elicit more information compared to no model statement (Leal et al., 2015). This now requires further evidence from future model statement studies in which two free recall attempts are elicited, where a control condition with no model statement between free recalls is tested.

Although this study used a forensic scenario (i.e. a mobile phone theft), this does not mean that the results would be replicable with all crimes involving children. For example, cases of alleged child sexual abuse are often characterised by the repetitive nature of the trauma (Lamb et al., 1997). In the current study children were talking about one event that happened on one occasion. We cannot, therefore, expect to find the same results when talking about one of many similar events. For repeated abuse cases, it might be necessary to ask more specific questions about each of the incidents recalled (Roberts & Powell, 2001). As the current study only looked at the effects of using a model statement on a single free recall, further testing is required to ascertain whether the same results can
be found through cued recall questioning. Furthermore, the type of crime witnessed could also affect the perceived relevance of details relating to that crime. In this study, relevant and irrelevant details were categorised according to the eyewitness literature (e.g. Gabbert, Hope & Fisher, 2009; Hope, Gabbert, Fisher & Jamieson, 2014; Vredeveldt et al., 2015). However, criminal relevance may vary and be subject to personal interpretation. Furthermore, it may rely upon what the victim/witness is willing to disclose about the incident. Indeed, in some cases of alleged child sexual abuse, children may avoid disclosing central information about the incident because of the negative emotions associated with the disclosure, such as internal feelings of shame, guilt and self-blame (Leander, Christianson & Granhag, 2007; Magnusson, Ernberg & Landström, in press). This demonstrates that the most forensically useful information may not always be obtainable, even when a child is telling the truth, and, therefore, cannot be the sole cue to their credibility.

Future research should also investigate the theoretical assumptions regarding the use of a model statement and test more directly the role that social comparison plays for truth-tellers in particular. In educational settings, the nature of social comparative feedback can be a strong determinant of performance through personal self-efficacy mechanisms (Bandura, 1993). Considering the statement as a ‘model’, therefore inferring it is better than their first recall attempt, could potentially lead to a downward social comparison by child truth-tellers. If this is the case, truth-tellers’ ability to say more might be impaired by a decrease in perceived self-efficacy and poor performance (Bandura, 1993). This could explain why truth-tellers did not provide significantly more crime-relevant information following the model statement and why they did not provide more details overall compared to the lie-tellers in our study.
In conclusion, we conducted the first empirical investigation of using a model statement with children. Our findings highlight the downside of using this technique in an eyewitness scenario where the event could be poorly encoded, but also reveal that forensic usefulness is an important verbal marker that could help to differentiate between children’s true and false reports.
3.6 References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting--

Akehurst, L., Manton, S., & Quandte, S. (2011). Careful calculation or a leap of faith? A
field study of the translation of CBCA ratings to final credibility

Bala, N., Ramakrishnan, K., Lindsay, R., & Lee, K. (2005). Judicial assessment of the

Bandura, A. (1993). Perceived self-efficacy in cognitive development and
functioning. *Educational psychologist, 28*(2), 117-148. doi:
10.1207/s15326985ep2802_3

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a
doi: 10.1037/0022-3514.56.5.669

approach to improve recall in different ages: Providing eye witnesses with a model
statement*. Manuscript submitted for publication.

memories. *Developmental review, 22*(3), 520-554. doi: 10.1016/S0273-2297(02)00006-0

practice. In K. A. Renninger, I. E. Sigel, W. Damon, & R. M. Lerner (Eds.),
Reference text:


Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J, & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. Applied Cognitive Psychology. Advance online publication. doi: 10.1002/acp.3270


95


Chapter 4: ‘Don’t mess with the detail’:
Increasing interviewee productivity using a model statement hinders child credibility judgments
4.1 Abstract

Child truth-tellers provide more forensically useful reports than child lie-tellers, and repeat a higher proportion of crime-relevant details across recall attempts after hearing a detailed account (also known as a model statement) compared to lie-tellers. However, these differences in verbal behaviour do not ensure that the detection of children’s true and false reports will be facilitated. In the current study, 114 adult evaluators judged the credibility of three different children (from a pool of 57 children): one provided a pre-model statement free recall, one provided a post-model statement free recall, and one provided both pre- and post-model statement free recalls. Detection accuracy hovered around chance levels regardless of the type of recall viewed by evaluators. This poor performance might be due to evaluators’ final credibility judgments being informed by perceptions of level of detail, which was not a reliable indicator of veracity for this study. In conclusion, evaluators were not good at detecting children’s deception, even when the model statement had elicited some differences between the true and false reports. Evaluators were not able to spot these differences, or if they did, they did not interpret them correctly.
4.2 Introduction

The ability to detect and interpret the presentation of deceptive cues in children’s investigative interviews is essential to reducing the risk of miscarriages of justice. Yet, it is frequently reported that both legal professionals (e.g. police officers, customs officers, judges, social workers) and laypersons (i.e. potential jurors) rarely perform at above chance levels (50%) when judging the credibility of children’s reports (Bala, Ramakrishnan, Lindsay & Lee, 2005; Leach, Talwar, Lee, Bala & Lindsay, 2004). Furthermore, this performance level plummets significantly below chance levels when a parent has rigorously coached a child to give a report (Talwar, Lee, Bala and Lindsay, 2006). In their study, Talwar et al. (2006) showed that adults were only able to detect just over a quarter of false reports (25.8%) made by children aged 4 to 7 years old, and that this only rose to 56% following cross-examination of the witness. These accuracy rates are even more disappointing when one considers that children under the age of 8 years have yet to develop sufficient verbal skills to conceal their lies (Talwar & Lee, 2002) and maintain them under questioning (Talwar & Lee, 2008). Thus, even young children’s false reports, that should be easier to detect, can slip through the net. Although children’s false allegations make up a small percentage of police cases (Anthony & Watkeys, 1991; Trocmé & Bala, 2005), it is important to recognise that, if undetected, each one of them could have serious, far reaching consequences for all persons involved (O’Donohue, Benuto & Fanetti, 2010). Thus, any technique that could facilitate the detection of false allegations would be useful during investigative interviews or even later when the evidence is played to judge and jury in court.

In Chapter 3, we conducted the first investigation into the use of a model statement in a child eyewitness context. With poor detection accuracy being largely attributable to the weak and subtle nature of naturally occurring behavioural differences between truth-
tellers and lie-tellers (DePaulo et al., 2003), the model statement constitutes a cognitive lie
detection approach aimed at enhancing these differences so that they are easier to detect
(Vrij, 2015). In Chapter 3, we found that true reports contained a higher proportion of
crime-relevant information compared to false reports. We also found that truth-tellers
were more consistent in the crime-relevant information they provided (i.e. there was a
higher proportion of crime-relevant repetitions) across recall attempts (i.e. pre- and post-
model statement) compared to lie-tellers. This latter difference, elicited by using a model
statement, is particularly important because testimonies that are perceived to be consistent
are more likely to be believed as truthful (Global Deception Research Team, 2006). So, if
child truth-tellers provided statements that were more consistent than child lie-tellers, then
child truth-tellers should be correctly judged as being more truthful than child lie-tellers.
However, we cannot simply assume that eliciting one verbal difference between children’s
true and false reports will definitely result in improved lie detection accuracy.

When Leal et al. (2015) played a model statement to adult insurance claimants,
they also elicited one difference; namely, false reports were rated as less plausible than true
reports. Analyses revealed that listening to the example of a detailed account resulted in
better discrimination between true and false reports. For those who heard the model
statement, independent raters’ plausibility scores correctly identified 80% of the reports as
true and false overall. On the other hand, when no model statement was played,
independent raters’ criteria-based content analysis scores only correctly identified 62.5%
of reports overall. These findings show how a ‘model interview’ can help to determine
veracity. Broken down, the addition of the model statement made little difference to lie
detection accuracy (72.7% for standard interview, 75% for model statement interview), but
a big difference to truth accuracy (52.4% for standard interview, 85% for model statement
interview). On a larger scale, this is in line with Vrij, Fisher and Blank’s (2015) meta-
analysis of the cognitive lie detection approach, which reported that techniques that encourage adult interviewees to say more are the most effective truth detection techniques. It remains unknown whether this will be the case with child interviewees.

**4.2.1 Lens modelling**

As part of our analyses, we used a lens model style of analysis to understand evaluators’ decision-making processes when judging child credibility. Brunswik’s (1952) lens model has been used for over six decades to understand human judgments and what factors affect this process in psychological literature (Karelaia & Hogarth, 2008). This framework explores how cues (such as number of details included in a statement) elicited by the target of a judgment (e.g. a child interviewee) can both reflect a quality of that target (veracity of his/her statement) and the judgment itself (truth/lie judgment) using correlations. This analysis demonstrates which cues are facilitating accurate judgments of a target’s quality (i.e. veracity) more than others. Thus, in a deception context, the lens model investigates not only how the behavioural cues communicated by the truth- or lie-tellers are related to their actual credibility, but also how these same behavioural cues are perceived by evaluators and how they inform a correct or incorrect credibility judgment. The lens modelling approach has already been applied to a meta-analysis of judgments of deception with adult interviewees. Hartwig and Bond (2011) concluded that increasing behavioural differences between liars and truth-tellers more readily improves detection accuracy than informing evaluators about valid cues to deceit.

**4.2.2 Aims**

The main aim of this study was to test whether using a model statement would enable evaluators to tell the difference between children’s true and false reports. Based on the findings in Chapter 3, we predicted that detection accuracy would be greater when evaluators judged children’s pre- and post-model statement recall attempts together
compared to when they simply watched either the pre-model statement recall attempt or the post-model statement recall attempt (Hypothesis 1). Furthermore, we anticipated that using a model statement would facilitate truth detection more than when no model statement was used (Hypothesis 2). Our secondary objective was to investigate how evaluators’ perceptions of different actual and believed cues to deceit changed across different recall attempts and how these perceptions were associated with actual cues to veracity using lens modelling.

4.3 Method

4.3.1 Participants

One-hundred-and-fourteen (28 males, 86 female) adult evaluators (\(M = 21.21\) years, \(SD = 6.87\) years) took part in the experiment. The majority of the participants (90.4%) were psychology undergraduate students recruited via a participant pool who received 0.5 course credit for their participation. The remaining participants were psychology post-graduate students recruited via convenience sampling within the Psychology Department. Regarding ethnicity, 79.8% were White/Caucasian, 7.9% were Black/African, 6.1% were Asian and 6.1% identified themselves as belonging to another ethnic group (unspecified).

4.3.2 Video clips

Fifty-seven of the sixty-three interviews (29 truth-tellers, 28 lie-tellers) from Chapter 3 were used in the current study. Six interviews were not used due to poor quality of recording (i.e. loud background noise). Each interview was edited into the three ‘recall attempt’ conditions to be presented: pre-model statement only, post-model statement only, and pre- and post-model statement together. There was a five-second gap between recall
attempts in the latter condition. Each evaluator saw one clip from each of the three ‘recall attempt’ conditions and the three clips they saw depicted three different child interviewees. Random listing was used to counterbalance truth-/lie-teller presentations (with each participant seeing at least one of each veracity condition). The order of the presentation of truth-tellers and lie-tellers and the order of the presentation of the three ‘recall attempt’ conditions were counterbalanced.

4.3.3 Response sheet for perceptual ratings

After each video clip, evaluators completed a questionnaire before watching the next video clip. Adapted from Evans, Michael, Meissner and Brandon (2013), the questionnaire included perceptual ratings (all with 9-point Likert scales) that related to both believed cues to deceit (e.g. nervousness, gaze aversion) and actual cues to deceit (e.g. level of detail, plausibility, cognitive demand) from previous literature (DePaulo et al., 2003; Vrij, 2008). A rating scale for perceptions of the prevalence of repeated information was included when evaluators were judging the pre- and post-model statement recall attempts to acknowledge the findings in Chapter 3. Evaluators were also asked to make a dichotomous truth/lie judgment, using their perceptual ratings to inform their decision.

4.3.4 ‘Coded cues’ for lens modelling

Total number of details and consistency measures across recall attempts (i.e. repetitions, commissions and omissions) as well as their proportionate crime-relevance, which were coded for in Chapter 3, were transferred into the SPSS database created to perform the lens modelling. They were used to see how they reflected the actual veracity of the children (already analysed in Chapter 3) and how they informed evaluators’ perceptual ratings that led to their final credibility judgments.
4.3.5 Procedure

Once evaluators had given written consent to their participation, they were asked to provide demographic information. They then received written instructions that they would watch three video clips of three different children being interviewed about the theft of a mobile phone at their school. They were instructed to watch the entire recording of each video clip only once. They were also informed that one of the children would provide two recall attempts and were told which clip this would be (i.e. first, second or third). An open folder on the computer screen contained the three video clips (labelled 1, 2 and 3), with instructions to watch these in order using the headphones provided. Participants were not told about the use of the model statement in the previous study. When they were told about the child who would provide two recall attempts, it was simply explained that the second recall was another opportunity for that child to say again everything that happened. Finally, evaluators were told that their task was to work out which of the children were lying and which of the child were telling the truth.

4.4 Results

Preliminary analyses showed that participant demographics (i.e. their age, gender and ethnicity), order of the veracity of the clips, ratio of true and false clips and order of recall attempts did not have any effect on the inferential statistics below. These variables were, therefore, collapsed in subsequent analyses.

4.4.1 Overall accuracy rates

The overall accuracy rate was 51\% (SD = 29.49), with a one sample t test revealing this accuracy rate did not significantly differ from chance (50\%), t(113) = .32, p = .75. A repeated measures ANOVA showed that there was no main effect of Recall Attempt, F(1, 112) = .30, p = .59, and thus overall accuracy was generally poor regardless of which part
of the children’s interviews the evaluators judged. Chi-squared analyses were performed for each of the three ‘recall attempt’ conditions separately to see whether truth and lie accuracy differed from chance level. For ‘pre-model statement only’, truth accuracy (56.9%) and lie accuracy (50%) did not significantly differ from chance, $X^2 = .55, p = .46$. For ‘post-model statement only’, truth accuracy (59.3%) was significantly above chance whereas lie accuracy (38.2%) was significantly below chance level, $X^2 = 5.09, p = .024, V = .21$. For ‘pre- and post-model statement’, neither truth accuracy (56.9%), nor lie accuracy (42.9%) differed from chance level, $X^2 = 2.45, p = .13$. Judging a ‘post-model statement’ recall attempt was therefore the only condition where truth accuracy rose significantly above chance level but it was also the only condition in which lie accuracy decreased significantly below chance level.

To summarise, listening to both the pre- and post-model statement free recalls together did not facilitate detection accuracy, contrary to Hypothesis 1. Overall detection accuracy did not significantly differ from chance regardless of which type of free recall was judged. Truth accuracy was, however, significantly above chance when evaluators assessed the credibility of a post-model statement free recall, supporting Hypothesis 2.

4.4.2 Perceptual ratings

Repeated measures ANOVAs were performed for each of the ‘cue’ ratings to see whether evaluators’ perceptions changed as a function of Recall Attempt. Ratings for crime-relevant repetitions, commissions and omissions were excluded from these analyses because they were only perceived in the ‘pre- and post-model statement’ recall attempt condition. Results showed that there was a significant main effect of Recall Attempt for perceptions of ‘level of detail’ only, $F(2, 226) = 18.35, p < .001$. Pairwise comparisons using Bonferroni adjustment showed that the ‘pre-model statement only’ recall attempt ($M = 5.10, SD = 2.25$) was perceived to be significantly less detailed than ‘post-model...
statement only’ recall attempt \((M = 6.46, SD = 2.29), p < .001, d = .40\), and the ‘pre- and post-model statement’ recall attempt\(^2\) \((M = 6.81, SD = 1.86), p < .001, d = .55\). The latter two ‘recall attempt’ conditions were perceived to be equally detailed, \(p = .73\). There was no main effect of Recall Attempt for ratings of plausibility, \(F(2, 226) = 1.50, p = .23\), nervousness, \(F(2, 226) = .57, p = .57\), cognitive demand, \(F(2, 226) = 1.29, p = .28\), or gaze aversion, \(F(2, 226) = .09, p = .92\).

### 4.4.3 Lens modelling

As the above analyses show that Recall Attempt (viewing pre-model statement recall only, post-model statement recall only or pre- and post-model statement recall attempts) did have an effect on evaluators’ perceptions of detail, three lens models were created to represent each of these experimental conditions.

For each of the three lens models displayed in Figures 4.1-4.3, Pearson’s correlation coefficients were calculated between (1) the veracity of the child’s statement and the ‘coded cues’, (2) the ‘coded cues’ and the ‘perceptual ratings’, and (3) the ‘perceptual ratings’ and the final credibility judgment. Finally, the relationship between the veracity of the child’s statement and the final credibility judgment given by the evaluator was also measured to provide an indication of detection accuracy. However, these correlations again showed poor accuracy for all conditions (pre-model statement only \((r(114) = .05, p = .57)\), post-model statement only \((r(114) = .02, p = .85)\) and pre- and post-model statement \((r(114) = -.02, p = .98)\)). Only significant \((p < .05)\) relationships are displayed in Figures 4.1-4.3 to improve the clarity of the model.

\(^2\) When participants watched both pre- and post-model statements (i.e. both recall attempts), their judgment was based on their overall perception of both recall together.
4.4.3.1 Pre-model statement only

Figure 4.1 shows the absence of any association between ‘coded cues’ and perceptual ratings. This suggests that the perceptual ratings, which all informed the final credibility judgment (apart from plausibility), were not based on any of the ‘coded’ cues observed in Chapter 3. The direction of the relationships between the ‘perceptual ratings’ and the final credibility judgment largely reflected previous literature. Children’s statements that were perceived to be more detailed were more likely to be perceived as truthful (DePaulo et al., 2003) and children who were perceived to be more nervous and as having to think harder were more likely to be perceived as deceitful (Mann & Vrij, 2006; Vrij, 2008). The only anomaly, for the current study, was that high levels of gaze aversion were associated with truthful statements. Gaze aversion is usually found to be associated with lie-telling in both adults (Global Deception Research Team, 2006) and children (McCarthy & Lee, 2009).
Figure 4.1 Decision-making process when evaluating ‘pre-model statement only’ recall attempts. Lines represent Pearson’s coefficients with a $p$ value about .05. For statement veracity and final credibility judgment, truth = 1 and lie = 0. *$p<.05$ **$p<.01$ ***$p<.001$

4.4.3.2 Post-model statement only

Figure 4.2 shows that perceptions of level of detail and plausibility were correctly informed by the actual number of details provided by the child interviewees in their post-model statement recall attempts, with a high number of details being perceived as more detailed and more plausible. These two perceptual ratings were then related to evaluators’ final credibility judgments, with statements that were perceived to be more plausible and more detailed resulting in ‘truth’ judgments. There was, however, a lack of continuity because the coded cue of ‘total details’, upon which these perceptions were based, was not linked to children’s actual veracity.
Figure 4.2 Decision-making process when evaluating ‘post-model statement only’ recall attempts. Lines represent Pearson’s coefficients with a $p$ value about .05. For statement veracity and final credibility judgment, truth = 1 and lie = 0. *$p<.05$ **$p<.01$ ***$p<.001$

4.4.3.3 Pre- and post-model statement

Figure 4.3 shows the only instance of continuity out of the three lens models. For this recall condition, when evaluators were able to view the pre-and post-model statement recall attempts, truthful statements contained a higher percentage of crime-relevant details compared to deceptive statements. The truthful statements were associated with higher ratings of crime-relevant repetitions, which, in turn, led to ‘truth’ judgments. Nevertheless, the lack of association between statement veracity and final credibility judgment suggests that this continuity was insufficient in enabling evaluators to discriminate between true and false reports. Similar to Figure 4.2, total number of details provided by the child interviewee was associated with perceptions of ‘level of detail’ and plausibility. Children who provided more details were perceived to be more detailed and more plausible,
resulting in ‘truth’ judgments. However, the lack of association, again, between total number of details and statement veracity demonstrated that this coded cue was not a reliable indicator. Indeed, it is likely that the aforementioned continuity in Figure 4.3 did not result in accurate credibility judgments because there was a stronger correlation between plausibility and final credibility judgment \((r(114) = .54)\) than between crime-relevance of repetitions and final credibility judgment \((r(114) = .22)\). In other words, the invalid cue of ‘total details’ was more influential than the valid cue of ‘% of crime-relevant repetitions’. Thus, evaluators appeared to be swayed by the inaccurate cue, similar to Fisher, Brewer and Mitchell (2009).

*Figure 4.3 Decision-making process when evaluating ‘pre- and post-model statement’ recall attempts. Lines represent Pearson’s coefficients with a \(p\) value about .05. For statement veracity and final credibility judgment, truth = 1 and lie = 0. \(*p<.05\) **\(p <.01\)** ***\(p<.001\)**
4.5 Discussion

Our study investigated whether using a model statement would help determine the veracity of children’s reports. The prediction that detection accuracy would be greater when evaluators judged a child’s pre- and post-model statement recall attempts together compared to when they simply watched either the pre-model statement recall attempt or the post-model statement recall attempt was not supported (Hypothesis 1). This is unexpected as watching both recall attempts revealed how children repeated a high proportion of crime-relevant information in their true reports but not in their false reports. As consistency is believed to be a key indicator of veracity (Fisher et al., 2009), but is not necessarily a reliable indicator of veracity (Vrij, 2008), this may explain why true reports were not judged to be more truthful. An overall accuracy rate of 51% is consistent with previous studies that show that adult detection rates for children’s lies hover around chance levels (see Talwar & Crossman, 2012 for an overview). However, using a model statement did facilitate truth accuracy. Evaluator performance for detecting true reports was only significantly better than chance when the report had been elicited following a model statement, supporting Hypothesis 2. This suggests that encouraging interviewees to say more might facilitate effective truth detection for children, as it does for adults (Leal et al., 2015; Vrij et al., 2015).

An in-depth examination of evaluators’ decision-making processes when judging child credibility, using lens modelling, revealed that the poor performance in our study might be due to a lack of continuity between behavioural cues that were indicative of statement veracity and perceptual ratings that subsequently informed final credibility judgments.

In the literature, detail is reported as a reliable indicator of veracity because generally truth-tellers provide more details than lie-tellers (DePaulo et al., 2003).
However, in the current study, this was not the case as true and false reports contained the same amount of detail. Thus, when evaluators’ credibility judgments were influenced by perceptions related to level of detail provided by the children, this did not result in an accurate judgment. Just as the model statement encouraged truth-tellers to delve back into their memory, it also challenged lie-tellers to ‘up their game’. The resulting increase in detail for both truth-tellers and lie-tellers may have confused evaluators and nullified the ‘total details’ cue on which they based their judgments. Furthermore, statements with a higher total number of details were also perceived to be more plausible, irrespective of veracity. In their study, Leal et al. (2015) found that, although truth-tellers and lie-tellers provided equally detailed statements, lie-tellers’ detailed accounts were perceived to be more implausible. It remains unknown whether this inconsistency is due to the general ‘detailed’ impression of the statement or cues specific to its content. However, the lack of an association between percentage of crime-relevant details and ratings of plausibility suggests that it is not due to how much forensically relevant content was included in the true or false statements.

Across all the lens models, final credibility judgments were consistently informed by perceptual ratings of unreliable cues (Vrij, 2008), namely gaze aversion and nervousness, which are not based upon any of the cues that we coded in Chapter 3. The use of invalid cues could, however, be due to participants being asked to explicitly rate these cues, which they might not have used otherwise. We incorporated unreliable cues into our perceptual ratings to see if they influenced credibility judgments. Previous research has demonstrated that people’s beliefs regarding valid cues to deceit are frequently erroneous (Global Deception Research Team, 2006) so we endeavoured to explore whether this would still be the case when given both valid and invalid cues to base their judgments on, similar to the Psychologically Based Credibility Assessment Tool used
in Evans, Michael, Meissner and Brandon (2013). In hindsight, an alternative method for tracking the use of valid and invalid cues could have been to ask participants to report what cues their decisions were based on and ask them to rate these cues in terms of importance. This would have avoided experimenter bias and would have still allowed us to see whether evaluators used the same cues as we coded for in Chapter 3. Furthermore, evaluators’ individual perceptions of what constitutes a detailed and plausible response from a child of a specific age might also vary. Future research should investigate how evaluators’ expectations of children’s competency to report detailed and plausible reports affect their credibility judgments.

Another limitation to consider is that clearly the coded cues in this work don’t account for all of the environmental cues to which the evaluators were exposed. Indeed, evaluators watched video clips of the interviews, which also contained nonverbal behaviours. Previous research, however, has often found no reliable differences in nonverbal cues between child truth-tellers and child lie-tellers (Talwar & Lee, 2002), suggesting that these are poor cues for judging child credibility. If evaluators in the current experiment were heavily influenced by these unreliable cues, then this would explain why they were unable to discriminate between true and false reports.

In the current study we only coded verbal behaviours. We focused on verbal behaviours because these are what the model statement aims to influence and these are the behaviours that are mostly likely to reveal child deceit (Akehurst, Manton & Quandte, 2011; Talwar & Lee, 2002). However, to understand fully just where evaluators are going wrong when judging child credibility we need to take into account in our lens model more of the cues that they are exposed to. Future research should continue to use lens modelling to understand how decisions regarding credibility are made. This might highlight behavioural cues that lead to correct credibility judgments, which should not be tampered
with when designing new techniques. For example, it was demonstrated in our study that total number of details influenced evaluators’ perceptions and that giving lie-tellers the opportunity to provide as many details as truth-tellers (by using a model statement) may, in fact, hinder their detection. Evaluators were blind to the interview technique used with the child interviewees. If using a model statement does elicit true and false reports that are equally detailed on a general level but differ on a specific level (e.g. number of crime relevant detail repetitions), then it would be helpful to explain this to evaluators. When police officers have been trained in using cognitive lie detection techniques, they first took part in a workshop where the rationale and results of using each technique were discussed (Vrij, Leal, Mann, Vernham & Brankaert, 2015). In that study, the use of these techniques combined with their knowledge of the anticipated outcomes of the techniques resulted in police officers’ detection accuracy increasing from 59% to 74%. Future research should investigate whether providing this information to laypersons, who are likely to be called for jury service, can also be beneficial.

In conclusion, using a model statement with children can facilitate truth detection but at the same time significantly hinders lie detection. This poor performance can be explained by insufficient behavioural differences being elicited between children’s true and false reports, as well as behavioural similarities in the level of overall detail provided by child truth-tellers and child lie-tellers. With no forewarning that using a model statement results in longer and more detailed statements from both truth-tellers and lie-tellers, evaluators were not able to accurately determine the veracity of the child witnesses.
4.6 References


Chapter 5: ‘Own vs. other’: The effects of modelling on verbal cues to children’s false reports
5.1 Abstract

The effect of playing a detailed free recall (i.e. a model statement) on verbal markers (i.e. productivity and consistency) indicating the veracity of children’s reports was investigated. A theoretical model comprising social comparison, interviewee self-efficacy and interviewee productivity was tested. Fifty-four children (8 to 10 years old) took part in one activity (true) and were told about a second activity that they did not experience (false). At interview, children provided a first and second recall attempt for each activity in an attempt to make the interviewer believe they had taken part in both activities. In between recall attempts, children heard another child’s model statement (AMS), or their own model statement (OMS), or they heard no model statement (NMS). Interviewee productivity increased for children’s true and false reports in the AMS condition, but not for the other two interview conditions. The model revealed one relationship: perceived self-efficacy was related to a positive change in interviewee productivity but only for the AMS condition. Between-statement consistency was lowest in the AMS condition compared to the OMS and NMS conditions, regardless of veracity. Memory accuracy remained unaffected by using model statements. To conclude, hearing AMS encouraged the most detailed but inconsistent reports. Further exploration is required to understand the socio-cognitive mechanisms behind the effects of using model statements.
5.2 Introduction

The assessment of children’s credibility in forensic settings has been widely acknowledged by researchers as an essential component of the justice system. With a growth in already high numbers of child witnesses testifying in court, it has become increasingly important to determine the veracity of their accounts (Talwar & Crossman, 2012). On the one hand, some studies have shown that children can provide informative and accurate accounts (e.g. Lamb et al., 2003; McCauley & Fisher, 1995) but, on the other hand, some studies have highlighted children’s willingness and ability to tell lies that fool parents, legal professionals and laypersons (i.e. potential mock jurors) alike (Bala, Ramakrishnan, Lindsay & Lee, 2005; Evans, Bender & Lee, 2016; Leach, Talwar, Lee, Bala & Lindsay, 2004). Consequently, child deception research now needs to focus on establishing interview methods to exaggerate differences between children’s truths and lies.

A particular area of weakness for child lie-tellers is their poor semantic leakage control (Talwar & Lee, 2002). Dependent on their developing cognitive abilities, children can find it difficult to tell and maintain a lie making them particularly prone to revealing their deceit (Gordon, Lyon & Lee, 2014; Talwar & Crossman, 2011; Talwar & Lee, 2008). This under-developed ability to control their verbal expressive behaviours not only reveals their deceit in short follow-up questions (Talwar & Lee, 2002, 2008), but also in longer narratives (Evans et al., 2012; Brunet et al., 2013; Williams, Talwar, Lindsay, Bala & Lee, 2014). It is particularly in these longer responses that verbal markers to children’s veracity have been found (Williams et al., 2014). These findings are in line with adult deception research that shows that longer reports are more likely to contain verbal cues to deceit (Vrij, Mann, Kristen & Fisher, 2007). Child lie-tellers could, therefore, be particularly vulnerable to cognitive lie detection techniques that encourage interviewees to say more
(such as using a model statement) because eliciting even longer responses should further exaggerate verbal differences between children’s true and false reports.

5.2.1 Interviewee productivity

*Interviewee productivity* refers to the amount of information provided by an interviewee and can be measured in terms of response length (number of words) or quantity of details. Productive interviewees are particularly important to the criminal justice system because accounts that are richer in detail are more likely to be believed in court (Bell & Loftus, 1989; Johnson, 2006) and can also provide more leads to facilitate an investigation (Vrij, Hope & Fisher, 2014). Thus, information elicitation is at the core of investigative interviewing.

A model statement refers to an audio-recorded example of a detailed, spontaneous free recall provided by a peer about an unrelated event that is played to an interviewee. Its primary aim is to encourage interviewees, both truthful and deceitful, to become more productive when interviewed (Vrij, 2015). It is, therefore, an information-gathering procedure that should help to overcome ambiguities surrounding what level of detail is required from truth-tellers, whilst also challenging lie-tellers to embellish their stories. Based on the premise that lie-telling is more cognitively demanding than truth-telling (Zuckerman, DePaulo & Rosenthal, 1981), using a model statement heightens lie-tellers’ cognitive load further by raising the expectations of what a detailed account sounds like (Vrij, 2015). Expecting more at interview creates an internal conflict in terms of how lie-tellers decide to react; on the one hand, they could elaborate on their original story to mimic truth-tellers but this runs the risk of leaking incriminating information, or, on the other hand, they could stick to their original story but this runs the risk of looking suspicious (Leal, Vrij, Warmelink, Vernham & Fisher, 2015).
All model statement studies, to date, have shown that both truth-tellers and lie-tellers, adults and children alike, increase their interviewee productivity in equal measures after hearing a model statement compared to both a control condition where no model statement is provided (Leal et al., 2015) and a baseline measure taken before a model statement is played (Chapter 3; Ewens et al., 2016). As such it was predicted, for the current study, that hearing a model statement would elicit a greater change in interviewee productivity than hearing no model statement for both children’s true and false reports.

As increased interviewee productivity has been evidenced for both truth- and lie-tellers, model statements do not appear to discriminate between true and false reports directly, but rather indirectly. For instance, Leal et al. (2015) found that, as the longer reports provided by deceitful insurance claimants were judged as less plausible than those provided by truthful insurance claimants, 80% of reports were correctly classified. That said, a significant difference in increased interviewee productivity between true and false reports would be most beneficial. More specifically, a larger increase in the amount of information given by truth-tellers compared to that given by lie-tellers (as a consequence of hearing a model statement) would help both police investigations and court trials. To enable this, it is important to take a step back and to understand how the underlying social and cognitive processes at play when hearing a model statement affect truth-tellers.

Model statements frame what constitutes a detailed account for truth-tellers. This is particularly helpful for child truth-tellers who frequently do not know what is expected of them at interview (Lamb, Orbach, Hershkowitz, Esplin & Horowitz, 2007) and can make the task of providing a free recall, usually in the absence of any framework, less cognitively difficult (Saykaly, Crossman, Morris & Talwar, 2016). Based on the theory of ‘social proof’ (Cialdini, 1993), the ‘appropriate behaviour’ required in the interview situation is based on the behaviour of the person providing the model statement (Vrij, Leal,
Mann, Vernham & Brankaert, 2015). Through the means of social comparison (Festinger, 1954), the model statement makes the majority of participants realise that they have not provided the expected level of detail required (Ewens et al., 2016). Upon making this downward social comparison, interviewees rectify this by adding to the information in their truthful accounts (i.e. increasing their interviewee productivity).

Despite the positive relationship between making a downward comparison and a positive change in interviewee productivity, there is also a theoretical negative relationship between making a downward social comparison and perceived self-efficacy for providing an informative and accurate account, which would decrease the change in interviewee productivity. To adjust the length and detail of their responses, truth-tellers may endeavour to delve back into their memory to try to recall more information that they have not already disclosed. This may involve assessing the quality of that memory and, importantly, their confidence in their ability to provide an informative, but also accurate account (Ackerman & Goldsmith, 2008). They may, therefore, reflect on their perceived self-efficacy to be able to carry out such a recall task. Downward social comparisons have been associated with low self-efficacy beliefs in educational settings that subsequently lead to poor performance (Bandura, 1993). This reflection could, therefore, create a negative effect on their ability to increase their interviewee productivity that opposes the positive effect of the downward social comparison. In other words, truth-tellers may not be increasing the length and quantity of detail in their accounts as much as they can because they don’t believe that they are able to provide an informative and accurate account, as demonstrated by a model statement given by another child. It was therefore hypothesised that the direction of social comparison would be skewed towards more downward social comparisons for children in the current study who heard another child’s model statement. Furthermore it was predicted that downward social comparison would be positively
associated with interviewee productivity but negatively related to perceived self-efficacy for children who heard another child’s model statement. This latter negative relationship would, in turn, reduce the strength of the positive relationship between perceived self-efficacy and change in interviewee productivity.

An alternative solution to using another child’s model statement might be to play children a model statement that they themselves have provided (i.e. their own model statement). In the current study, we investigated the transformation of children’s practice interviews (also known as practice narratives) into their own model statements. Model statements and practice interviews already share the common goals of teaching truthful interviewees what is expected of them and of increasing retrieval performance (Brackmann, Otgaar, Roos af Hjelmsäter & Sauerland, 2016). Forming part of the pre-declarative stage in the National Institute of Child Health and Human Development (NICHD) Investigative Interview protocol (Lamb et al., 2007), practice interviews aim to prepare children to become effective informers by encouraging them to talk about unrelated past events in as much detail as possible. This training in answering open-ended questions conveys the interviewer’s expectations for a detailed description and develops a lengthy responding pattern (Sternberg et al., 1997), similar to a model statement. Whereas the benefits of another child’s model statement on child truth-tellers have received little support (Brackmann et al., 2016), the benefits of using practice interviews to elicit longer and more accurate reports from child interviewees are well documented (Price, Roberts, & Collins, 2013; Roberts, Brubacher, Powell & Price, 2011; Sternberg et al., 1997). It was, therefore, anticipated that using practice narratives to act as children’s own model statements would have a positive effect on child truth-tellers increasing their interviewee productivity.
Playing an excerpt of their practice interviews back to child interviewees in the form of their own model statement might remind them of the level of detail that is required. For the children who have forgotten how to give a detailed statement following the time delay, it might highlight that they haven’t provided enough information in their first free recall of the true activity (downward self-comparison) and, therefore, they should increase their interviewee productivity for the second recall attempt, (as was expected for the children who hear another child’s model statement). However, there may not be a resultant negative relationship between downward self-comparison and perceived self-efficacy because the practice interview would demonstrate each child’s potential to be informative and accurate and would highlight the fact that the interviewer accepts that level of productivity. Thus, perceived self-efficacy has the potential to remain high and be positively and strongly correlated with an increase in interviewee productivity.

A stronger, positive relationship between perceived self-efficacy and change in interviewee productivity for children’s true reports after hearing their own model statement rather than another child’s model statement should result in a greater change in interviewee productivity for the former group’s true reports. This would lead to (a) a veracity difference for the own model statement condition whereby the increase in interviewee productivity would be greater for true reports than false reports, and (b) an interview technique difference whereby the increase in interviewee productivity would be greater for true reports elicited following a child’s own model statement compared to another child’s model statement.

5.2.2 Between-statement consistency

Between-statement consistency refers to the level of consistency between statements made by the same person. In the context of this study, it refers to the consistency of the information provided in the first and second recall attempts. Low
consistency is important because it is largely interpreted by laypersons as an indication of deceitfulness (DePaulo et al., 2003; Global Deception Research Team, 2006). Previous ‘consistency’ hypotheses for child deception have inferred a relationship between interviewee productivity and consistency. For the Pinocchio hypothesis (Bruck, Ceci & Hembrooke, 2002), as lie-tellers grow their stories (like their noses), it is suggested that their false reports may contain more commissions (details not given before) and fewer repetitions when compared to true reports. Alternatively, the repeat versus reconstruct hypothesis (Granhag, Strömwall & Jonsson, 2003; Strömwall & Granhag, 2005) predicts that as truth-tellers delve back into their memory and remember more information, their true reports will contain more commissions, fewer omissions (details that were given before but are left out in a second recall attempt), and fewer repetitions when compared to false reports. As such, for both hypotheses, it is suggested that longer reports become more inconsistent. The findings of Chapter 3 support this relationship as an equal increase in productivity following the use of another child’s model statement also resulted in equal changes in the general consistency of true and false reports. However, these findings were limited by the lack of control group to which these changes in consistency could be compared. In the current study, it was anticipated that both true and false reports elicited using model statements (both ‘own’ and ‘other’s’) would become more inconsistent than true and false reports elicited without the use of a model statement due to the expected increases in interviewee productivity. Further, due to the prediction that a greater increase in interviewee productivity will occur for children’s true reports after hearing their own model compared to children’s true reports after hearing another’s model statement, it was anticipated that truthful information provided by children who hear their own model statement will become more inconsistent than it will for those who hear another’s model statement. Finally, in terms of consistency, a greater increase in interviewee productivity
for true reports, compared to false reports after hearing one’s own model statement, might, in turn, result in the children’s true reports being more inconsistent than their false reports.

5.2.3 Memory accuracy

Talwar and Crossman (2012) presented cognitive lie detection as a promising means for facilitating child deception detection with the proviso that any potential adverse effects of the techniques on truth-tellers were investigated. As noted above, both Leal et al. (2015) and Chapter 3 of this thesis found that playing another’s model statement resulted in truth-tellers providing accounts that were significantly longer and more detailed than when no model statement was used. When eliciting more information from truthful interviewees it is important to ensure that the memory accuracy of that information is maintained. It would be undesirable if there were a trade-off between quantity of detail and accuracy (Koriat & Goldsmith, 1996; Koriat et al., 2001). Other interview techniques that encourage longer responses, such as the Cognitive Interview, result in an increase in both correct and incorrect information compared to standard interview protocols (Köhnken, Milne, Memon & Bull, 1999; Memon, Meissner & Fraser, 2010). However, this does not necessarily lead to a decrease in recall accuracy (McCauley & Fisher, 1995). Indeed, it is possible to increase productivity whilst also maintaining accuracy.

There are mixed results regarding the effect of using another child’s model statement on children’s memory accuracy. On the one hand, Chapter 3 shows that child truth-tellers’ longer, post-model statement accounts were significantly less accurate than their shorter, pre-model statement accounts. This was due to a significant increase in incorrect information, but not a significant increase in correct information. On the other hand, Brackmann et al. (2016) found that using another’s model statement had no effect on the accuracy of children’s eyewitness reports. In Chapter 3, it is suggested that the decreased accuracy might be due to the children not receiving any forewarning that the
target event was going to occur resulting in poor encoding of that event. This meant that, at interview, truthful children might have sacrificed their accuracy, making up detail that did not actually occur, in order to be perceived as informative, in line with the revised dual-criteria model of meta-memory processes (Ackerman & Goldsmith, 2008). Although Brackmann et al. (2016) did not inform their child participants that they would be interviewed about the staged event, it was performed in front of an entire class, which suggests that it had the children’s full attention (i.e. an optimal encoding condition). As the staged event, therefore, received all of the children’s attentional resources, it is likely that this facilitated encoding and increased memory quality of the event (Lane, 2006). To test our model statements in a similar ‘best case’ scenario in which the witness has encoded the event well, the experimenter conducting the present study forewarned children that they would be questioned about the activities they took part in. As such, and in line with the findings of Brackmann et al. (2016), it was predicted that there would be no difference in memory accuracy across the interview conditions. That is, it was anticipated that although the model statements would increase interviewee productivity when compared to the ‘no model statement’ condition, this would not be to the detriment of memory accuracy.

5.3 Method

5.3.1 Participants

Fifty-four children (23 boys, 31 girls) aged 8 to 10 years old ($M = 9.09$ years, $SD = .73$ years) were recruited from a local primary school in the United Kingdom. Children’s legal guardians received an information sheet regarding the study and returned a signed consent form agreeing to their child’s participation. Verbal assent was also obtained from
all children. All children received a certificate and a small gift (e.g. stickers, balloons) in exchange for taking part.

5.3.2 Materials

5.3.2.1 Adapted practice interview

Based on the NICHD interview protocol (Lamb et al., 2007), a mock practice interview was conducted with all children about an atypical activity that took place in the past nine months (e.g. attendance at a wedding). Legal guardians were asked to provide, on their returned consent forms, details of an event that their children could remember well (similar to Brubacher, Roberts & Powell, 2011). If an event was not provided, the PI asked the children if they could think of an activity that they had done for the first time within the last nine months. Average delay between time of event and subsequent recall was 4.63 months ($SD = 2.49$ months). As testing occurred during school hours, measures were taken to limit the amount of time children spent with the researchers. One such measure was to shorten the practice interview protocol. In our adapted version, the principal investigator (PI) told the children that she wanted to know more about them and the things that they do. It then combined lines 1a to 1c in the 2007 version of the NICHD protocol by asking the children to “think hard about the recent activity/event and to tell the PI what happened on that day from the time they got up that morning until they went to bed that night”. Following this, the children were asked if they could tell the PI any more information about the event. All practice interviews were audio-recorded.

5.3.2.2 Activities

Two activities were designed for this experiment. The activities were designed to be of equal length and to contain an equal amount of actions without being too similar in
content. To keep the children engaged, both activities were appropriate for their age and were interactive.

5.3.2.2.1 Lego activity

Instructions for building a Lego tower firstly directed the children to open up a box of Lego next to them and to tip all the Lego pieces onto the floor. Then, using a sheet with different coloured circles on it, they were asked to sort all the pieces into their corresponding colours (e.g. red bricks into the red circle). Next the children were asked to build a Lego tower depicted in the instructions as fast as they could. The children were told to try to copy the image as best they could (i.e. it was not necessary for the tower to be an exact replica). Finally, they put together the pieces of a Lego man and put him on top of the tower.

5.3.2.2 Dice activity

Children who played the dice game were first instructed to find a red, foam dice and an orange, whiteboard marker in a bag filled with other miscellaneous items. They were then asked to roll the dice twelve times. After each roll, they were asked to write down the number on the top of the dice in one of the twelve boxes provided on a score sheet that was provided. On this sheet, the twelve boxes were organized into four rows of calculations with three boxes included in each calculation. Each row was connected to a different activity: star jumps, stand on one leg, claps and skip on the spot. When all the boxes contained a number, the children were asked to complete the four calculations to find out how many times (star jumps and claps) or for how many seconds (standing on one leg or skipping on the spot), they would then need to do that activity. After they had completed each of the activities in turn, they were instructed to roll the dice one more time. If they rolled an odd number, they were asked to stand up and take a bow. If they rolled an even number, they were told to give themselves a pat on the back.
5.3.2.3 Another child’s model statement (AMS)

This model statement was an audio-extract from an interview with a nine-year-old boy who attended a different school from the rest of the sample. This was a 319-word (level of detail: 85, duration of statement: 145 seconds) spontaneous free recall that was elicited using the same instructions provided when conducting the ‘practice’ interview with children in the current study. Following the guidance set out by Leal et al. (2015), the boy spoke about an event (a holiday) that was unrelated to the target activities. From the perspective of a source-monitoring framework (Johnson, Hashtroudi & Lindsay, 1993; Lindsay, Allen, Chan & Dahl, 2004), this conceptual dissimilarity between the ‘model statement event’ and the ‘to-be-remembered event’ is key to reducing the risk of interference between the child’s own information (i.e. recall of the experienced event) and the information provided by an external source (i.e. the model statement).

5.3.2.4 Own model statement (OMS)

The ‘practice’ interview of children randomly assigned to hear his/her ‘own model statement’ at interview was edited to be the same length (in seconds) as the model statement in the AMS condition. However, the duration of the practice interviews for three of these children were shorter than the model statement in the AMS condition, resulting in them also being significantly shorter than the other ‘own model statements’ (more than 1 standard deviation away from the mean duration). These three outliers consequently skewed the overall data, resulting in an overall mean duration for OMS that was significantly shorter in duration ($M = 123.67$ seconds, $SD = 29.45$ seconds), $t(17) = -3.07$, $p = .007$, $d = .72$, significantly shorter in response length ($M = 265.67$ words, $SD = 98.86$ words), $t(17) = 2.29$, $p = .035$, $d = .54$, and significantly less detailed ($M = 68.67$ details, $SD = 24.53$ details), $t(17) = 2.82$, $p = .012$, $d = .66$, than the AMS. Nevertheless, one of the purposes of the ‘own model statement’ condition was to demonstrate that children’s
individual ability to provide a detailed account might not be equal to that of another child providing a model statement (i.e. AMS). Indeed, the large standard deviations for duration, response length and number of details for the model statements in the OMS condition suggest that there is much variation in children’s ability to provide a detailed account.

5.3.2.5 Social/self-comparison

Following their interviews, children in the ‘model statement’ conditions (AMS or OMS) were asked to compare their first and second free recalls for the true activity to the model statement that they heard during their interviews. For each free recall, they were asked to report whether they had performed better (upward social comparison), worse (downward social comparison), or the same (neutral social comparison) as the person in the recording (i.e. themselves (OMS) or the boy they did not know (AMS)). Type of comparison was coded as follows: 3 = upward, 2 = neutral and 1 = downward. Thus, a high ‘score’ reflected an upward comparison and a low ‘score’ reflected a downward comparison. We also asked children to rate the extent to which they compared themselves to the model statement that they heard (1 = did not compare themselves at all, 7 = compared themselves a lot).

5.3.2.6 Perceived self-efficacy

To measure children’s beliefs about their ability to provide an informative and accurate report of the true activity, children were asked to rate their level of agreement on a 5-pt Likert scale (1 = strongly disagree, 5 = strongly agree) with the following statements: (1) ‘I was able to remember everything that happened’ (informativeness), and (2) ‘I was able to only say things that I’m sure are correct’ (accuracy). Higher ratings indicated higher self-efficacy beliefs.
5.3.3 Procedure

The children were invited to take part in a science experiment that the principal investigator (PI) was learning to run for some of her colleagues at the university. They were told that this would consist of taking part in two activities that formed the experiment (Testing session 1), and then being interviewed by one of her colleagues two to three days later about what had happened (Testing session 2). It was explained that the purpose of the interview was to check that the PI was running the experiment correctly. All children were tested individually.

5.3.3.1 Testing session 1

The PI started this session with the adapted practice interview (see details above), which was audio recorded. Children were then invited to take part in the first activity, which was either building a Lego tower or playing a dice game. The type of activity was counterbalanced. Written instructions were provided for both activities with each activity being divided into four tasks. These were read aloud by the PI to the participants. On average, the activities took 323.46 seconds ($SD = 41.71$ seconds) to complete. Neither type of activity, $F(1, 50) = .31, p = .58$, nor participant gender, $F(1, 50) = .061, p = .81$, affected the duration of the activity. At the end of the first activity, the children were informed that the PI had forgotten to bring in the equipment necessary to complete the second activity and were told that the PI would get into trouble if her colleagues found out that she didn’t do both activities with the children. The PI then suggested to the children that she could tell them all about the second activity and that they could ‘cover’ for the PI by making her colleague believe during their interviews that they had taken part in both activities as planned. Once the children had verbally assented to providing a false report, the PI ran through the instructions for the activity that they had not experienced (i.e. playing the dice game if they had built the Lego tower or building the Lego tower if they
had played the dice game). Before leaving the first testing session, the PI reiterated to the children that they would be interviewed and would be required to tell the truth about the first activity that they had actually taken part in, and that they needed to tell a lie about the second activity that they had not taken part in by pretending that they had actually conducted both activities.

5.3.3.2 Testing session 2

This took place 2 days after testing session 1. To start, the PI reminded the children that they were only able to take part in one of the activities during the first testing session and would now be asked to ‘cover’ for the PI during their interviews and convince her colleague that they had taken part in both activities. The children were told that it was therefore important to remember as many details as they could about the ‘true’ activity, as it was to come up with details about the ‘false’ activity based on the outline that the PI had given to the children in the previous testing session. All participants were provided with the instructions for the ‘false’ activity that they did not experience. They were given a maximum of five minutes to prepare what they wished to tell the interviewer about both activities. If they were ready before the preparation time had elapsed, they notified the PI and were escorted to the interview room. The children took, on average, 153.19 seconds to prepare for their interviews ($SD = 90.05$ seconds). Neither participant gender, $F(1, 50) = .007, p = .93$, nor participant age, $F(1, 50) = 1.24, p = .27$, had an effect on preparation time. However, type of activity did, $F(1, 50) = 10.37, p = .002$. Children who prepared their false report about the dice activity ($M = 187.36$ seconds, $SD = 101.30$ seconds) took longer than children who prepared their false reports about the Lego activity ($M = 116.38$ seconds, $SD = 58.25$ seconds), $d = .85$.

Following best practice, the interview protocol used for this experiment incorporated aspects of the Cognitive Interview (Fisher & Geiselman, 1992): an initial
rapport-building phase followed by an open request for an uninterrupted free recall of the two activities separately. To improve accuracy, an instruction to try not to guess was also given. Following this initial baseline recall, interview protocol changed according to the interview condition to which the children had been randomly allocated: another’s model statement (AMS) versus own model statement (OMS) versus no model statement (NMS).

Children in the AMS condition were played the audio recording of the boy talking about his recent holiday. Following the boy’s model statement, the interviewer commented that the boy in the recording had included lots of details about everything that had happened. The interviewer then invited the children to provide another uninterrupted free recall of the two activities separately, asking them to try to add in as much detail as they could. Children in the OMS condition were played the edited audio recording of their practice interview from Testing Session 1. Following their own model statement, the interviewer commented that the children had included lots of details in their previous interview. All children were told this, regardless of their performance in the practice interview. This was because the practice interview represented each child’s own ability to be detailed and was, therefore, their own ‘model’. The interviewer then invited the children to provide another uninterrupted free recall of the two activities separately, asking them to try to add in as much detail as they could. Children in the ‘no model statement (NMS)’ condition were not played any statement and were simply invited to provide another uninterrupted free recall of the two activities separately. The interviewer asked them to try to add in as much detail as they could.

This interview protocol was much shorter than standard protocols used in police investigations due to the researchers wishing to limit the duration of the testing session. As children were asked to provide four detailed, free narratives, it was also anticipated that a shorter interview protocol would reduce fatigue effects on their interview performance.
All interviews were video-recorded and later transcribed verbatim for coding purposes. Note that order of recall was counterbalanced and remained the same across recall attempts (i.e. some children talked about the ‘true’ activity first in both free recalls and some children talked about the ‘false’ activity first in both free recalls).

After the interviews, children returned to the PI to rate their social comparison (if they were assigned to one of the ‘model statement’ conditions) and their perceived self-efficacy. Children then received an oral debriefing highlighting lying as an undesirable and discouraged behaviour. A written debriefing was also sent home to their legal guardians.

5.3.4 Coding schemes

Two independent raters coded the children’s transcripts. The first coder rated all transcripts and the second coder rated one third of transcripts (18 transcripts in total). Intra-class coefficients across all ‘coded’ variables were high (ICCs > .80).

5.3.4.1 Interviewee productivity

Participants’ four free recall attempts (first recall attempt of true activity, first recall attempt of false activity, second recall attempt of true activity, and second recall attempt of false activity) were coded separately for response length and quantity of detail. Response length was represented by word count. Any speech utterances such as hmm or uh were excluded. To increase precision when coding for quantity of detail, each of the free recalls were coded for five different types of details: visual details (e.g. red bricks = two visual details), spatial details (e.g. on top of the tower = one spatial detail), temporal details (e.g. at the end we rolled the dice = one temporal detail), action details (e.g. I sorted the bricks and built the tower = two action details), and auditory details (e.g. she said that I was very sporty = one auditory detail). Total number of details was then the sum of the frequencies
for each of these details added together. Length of recall attempt (in words) and number of details being highly correlated across all conditions \((r > .62, p < .01)\): as number of details is more important to a police investigation than number of words, this measure is hereafter used to represent interviewee productivity. Finally, change in interviewee productivity was calculated for each participant for true and false reports separately. This was calculated by subtracting the total number of details in the participants’ first recall attempt of the true/false activity away from the total number of details provided in their corresponding second recall of attempt of the true/false activity. A positive number reflected an increase in interviewee productivity; a negative number reflected a decrease in interviewee productivity.

**5.3.4.2 Between-statement consistency**

Consistency was measured by, first, counting how many of the details provided in the second recall attempts of the true and false activities separately were repetitions of information provided in the first recall attempts of the same activity. Second, the number of repeated details, for true and false activities separately, was subtracted from the total number of details provided in the corresponding second recall attempts to calculate the number of commissions (new information). Finally, the number of repeated details, for true and false activities separately, was subtracted from the total number of details provided in the corresponding first recall attempts to calculate the number of omissions (omitted information).

**5.3.4.3 Memory accuracy**

A list of information that could be provided regarding each of the activities and that could be classified with certainty as being correct was created using the standardized instructions. Each of the participants’ first and second recall attempts of the true activity was then coded separately. First, each detail in the list was categorized as either
‘mentioned’ or ‘not mentioned’ in free recalls. Second, all ‘mentioned’ details were coded as either correct or incorrect. Finally, the number of correct details provided was divided by the total number of ‘mentioned’ details (correct details + incorrect details) and multiplied by 100 to calculate memory accuracy.

5.4 Results

Preliminary analyses revealed no significant effects involving participant age, participant gender, interviewer, true/false activity, or veracity order at recall. Thus, all the data were collapsed for all subsequent analyses.

5.4.1 Interviewee productivity

A 2 (Veracity: True vs. false) x 3 (Interview Protocol: AMS vs. OMS vs. NMS) mixed ANOVA was used to test differences in change in interviewee productivity across recalls. There was no significant main effect of Veracity, \( F(1, 51) = 1.32, p = .26 \). Children increased the amount of detail from their first recall attempt to their second recall attempt equally for the true \( (M = 3.44, SD = 11.84) \) and false \( (M = 1.52, SD = 9.11) \) activities. There was a significant main effect of Interview Protocol, \( F(1, 51) = 6.95, p = .002 \). Pairwise comparisons using Bonferroni adjustment revealed that children in the AMS condition \( (M = 15.72, SD = 15.40) \) increased the amount of overall detail they provided for both activities more than those in the OMS condition \( (M = 2.00, SD = 11.66), p = .03, d = 1.00, \) and the NMS condition \( (M = -2.83, SD = 18.64), p = .002, d = 1.09 \).

There was no difference in the overall change in interviewee productivity for the OMS and NMS groups, \( p = 1.00 \). Finally, there was no significant interaction between Veracity and Interview Protocol, \( F(2, 51) = 1.15, p = .32 \). These findings partially support the prediction that there would be an increase in interviewee productivity for children who heard a model statement compared to those who didn’t. The prediction was correct for
those who heard another’s model statement but not for the children who heard their own model statement. This increase in productivity for the children in the AMS condition occurred for both their true and false reports.

5.4.2 Social/self-comparison, perceived self-efficacy and change in interviewee productivity

These three variables were tested for the true reports provided by children in the AMS condition and the OMS condition separately. First, chi-squared analyses were used to determine whether the three types of comparison (upward, neutral or downward) were equally selected. Second, the relationships between the three pairs of variables were tested using Pearson’s correlation coefficients.

5.4.2.1 Social/self-comparison and interview protocol

Contrary to predictions, selection of comparison type (upward, neutral or downward) was equally distributed for the AMS condition, $X^2(2) = 2.33, p = .31$, but not for the OMS condition, $X^2(2) = 9.00, p = .011$. When children heard AMS, 22.2% made an upward social comparison, 50% made a downward social comparison and 27.8% made a neutral comparison. When children heard their OMS, two-thirds (66.7%) reported that their performance on the first recall attempt was the same as their performance in their practice interview, when their own model statement was recorded (neutral self-comparison), 16.7% made an upward self-comparison and 16.7% made a downward self-comparison.

We looked to verify whether children’s self-reported comparisons generally reflected the difference, or lack thereof, between their first recall attempt and the model statement that they heard. For children who heard the model statement given by another child, a one-sample t-test was performed to compare each participant’s average response
length and total details given to those of the model statement given by the other child (length: 319 words, details: 85). On average, children in the AMS condition provided a first recall attempt of the true activity that was significantly shorter in length ($M = 115.83$ words, $SD = 64.81$ words), $t(17) = -13.30, p<.001, d = 3.21,$ and significantly less detailed, ($M = 29.67$ details, $SD = 17.26$ details), $t(17) = -13.60, p<.001, d = 3.13,$ than AMS. For children who heard their OMS, paired samples t-tests were used. On average, children in the OMS condition provided a first recall attempt of the true activity that was significantly shorter in length ($M = 110.50$ words, $SD = 46.84$ words), $t(17) = -7.64, p<.001, d = 1.03,$ and significantly less detailed ($M = 28.61$ details, $SD = 9.23$ details), $t(17) = -6.97, p<.001, d = 1.64,$ than their OMS. In summary, children in both ‘model statement’ conditions should have perceived downward social/self-comparisons.

### 5.4.2.2 Social/self-comparison and change in interviewee productivity

For the children in the AMS condition, there was no relationship between their type of social comparison and a change in interviewee productivity, $r(16) = .07, p = .78.$ Similarly, for children in the OMS condition, there was also no relationship between reported self-comparison and a change in interviewee productivity, $r(16) = .07, p = .77.$ Therefore no support was found for the prediction that downward comparison would spur on the children to increase their interviewee productivity.

### 5.4.2.3 Social/self-comparison and perceived self-efficacy

There was no relationship between reported social comparison and perceived self-efficacy for informativeness for children in the AMS condition, $r(16) = -.22, p = .37,$ neither was there a relationship between reported self-comparison and perceived self-efficacy for informativeness for children in the OMS condition, $r(16) = -.05, p = .83.$ There was also no relationship between social comparison and perceived self-efficacy for accuracy for children in the AMS condition, $r(16) = -.06, p = .81,$ and again no
relationship between self-comparison and perceived self-efficacy for accuracy for children in the OMS condition, $r(16) = .04, p = .89$. Therefore no support was found for the prediction that downward comparison would lead to reduced perceived self-efficacy.

5.4.2.4 Perceived self-efficacy and change in interviewee productivity

For children in the AMS condition, self-reported beliefs in their ability to provide an accurate report of what happened during the true activity were positively correlated with a change in interviewee productivity in their true reports, $r(16) = .49, p = .04$. There were no other relationships between perceived self-efficacy (neither accuracy, nor informativeness) and change in interviewee productivity ($ps > .05$).

5.4.3 Between-statement consistency

A 2 (Veracity: True vs. False) x 3 (Interview Protocol: AMS vs. OMS vs. NMS) mixed ANOVAs was used to test differences in commissions, omissions and repetitions during the children’s second recall attempts.

5.4.3.1 Commissions

There was no significant main effect of Veracity, $F(1, 51) = 1.69, p = .20$, on commissions. Children added the same amount of new information into their second recall attempt about the true activity ($M = 16.02, SD = 7.56$) as they did about the false activity ($M = 14.35, SD = 9.76$). There was a significant main effect of Interview Protocol, $F(2, 51) = 4.76, p = .013$. Pairwise comparisons using Bonferroni adjustment revealed that children in the AMS condition ($M = 40.22, SD = 23.06$) provided more commissions overall than those in the OMS condition ($M = 24.50, SD = 12.48$), $p = .02, d = .85$, and the control NMS condition ($M = 26.39, SD = 12.20$), $p = .04, d = .75$. There was no difference between the OMS and NMS groups, $p = 1.00$. Finally, there was no significant Veracity X Interview Protocol interaction effect, $F(2, 51) = .95, p = .39$. 

146
5.4.3.2 Omissions

There was no significant effect of Veracity, $F(1, 51) = .06, p = .81$, on omissions. Children omitted as much information overall from their true reports ($M = 12.57, SD = 8.61$) as from their false reports ($M = 12.83, SD = 7.48$). There was no significant effect of Interview Protocol, $F(2, 51) = 1.08, p = .35$. Number of overall omissions did not differ between children who heard AMS ($M = 24.50, SD = 17.24$), children who heard their OMS ($M = 22.50, SD = 8.42$) and children who heard no model statement ($M = 29.22, SD = 15.07$). There was also no significant Veracity X Interview Protocol interaction effect, $F(2, 51) = .83, p = .44$.

5.4.3.3 Repetitions

There was a significant main effect of Veracity, $F(1, 51) = 19.10, p < .001$, on repetitions. Children repeated significantly more details in their false reports ($M = 21.44, SD = 7.91$) than in their true reports ($M = 17.69, SD = 7.56$), $d = .60$. There was no significant main effect of Interview Protocol, $F(2, 51) = .17, p = .84$. Number of overall repetitions did not differ between children in the AMS condition ($M = 40.61, SD = 16.19$), OMS condition ($M = 38.94, SD = 13.37$), and NMS condition ($M = 37.83, SD = 13.32$). There was also no significant interaction effect between Veracity and Interview Protocol, $F(2, 51) = .95, p = .40$.

In summary, both true and false reports provided by children in the AMS condition contained significantly more new information than the reports of the children in the OMS and ‘NMS’ conditions. We predicted that the children in both model statement conditions would be more inconsistent that the children who did not hear a model statement. This was only the case for the children who heard another’s model statement. Children who heard their own model statement were no different, in terms of consistency, from those who did not hear a model statement. Furthermore, these findings do not support the
predictions that truthful children who heard their own model statement would be more inconsistent than lie-telling children who heard their own model statement nor do they support the prediction that children who heard their own model statement would be more inconsistent than those who heard another’s model statement.

5.4.4 Memory accuracy

A 2 (Recall Attempt: First vs. Second) x 3 (Interview Protocol: AMS vs. OMS vs. NMS) mixed ANOVA was performed to examine differences in total number of correct details, total number of incorrect details and overall memory accuracy for the truthful accounts.

5.4.4.1 Correct details

There was no main effect of Recall Attempt, $F(1, 51) = .15, p = .70$. The amount of correct details provided did not differ from the first recall attempt ($M = 13.24, SD = 5.53$) to the second recall attempt ($M = 13.50, SD = 5.93$). There was no main effect of Interview Protocol, $F(2, 51) = .38, p = .69$. There was no difference in the overall number of correct details provided by children in the AMS condition ($M = 27.78, SD = 14.75$), OMS condition ($M = 25.00, SD = 7.55$), and NMS condition ($M = 27.44, SD = 7.27$). There was also no Recall Attempt X Interview Protocol interaction effect, $F(2, 51) = 2.27, p = .11$.

5.4.4.2 Incorrect details

There was no main effect of Recall Attempt, $F(1, 51) = .51, p = .48$. Number of incorrect details did not change from the first recall attempt ($M = 1.78, SD = 1.49$) to the second recall attempt ($M = 1.94, SD = 1.56$). There was a significant main effect of Interview Protocol, $F(2, 51) = 3.43, p = .04$. However, pairwise comparisons using Bonferroni adjustment revealed no significant differences in overall number of incorrect
details between each of the three interview protocols. Children in the AMS condition ($M = 4.94, SD = 2.90$) did not provide significantly more or less incorrect details than those in the OMS condition ($M = 3.06, SD = 2.07), p = .07, or the NMS condition ($M = 3.17, SD = 2.23), p = .09. There was also no difference in incorrect details between the OMS condition and the NMS condition, $p = 1.00$. The stricter statistical rigour imposed by using Bonferroni adjustment could account for the lack of any significant differences at this level of analysis. There was no Recall Attempt X Interview Protocol interaction effect, $F(2, 51) < .001, p = 1.00$.

**5.4.4.3 Overall accuracy**

There was no main effect of Recall Attempt, $F(1, 51) = .04, p = .84$. Children were equally accurate in their first recall attempt ($M = 88.24\%, SD = 9.28\%) as they were in their second recall attempt ($M = 87.95\%, SD = 8.52\%). There was no main effect of Interview Protocol, $F(2, 51) = 2.56, p = .087$. Children maintained high accuracy irrespective of the interview condition to which they were allocated (AMS, $M = 85.05\%, SD = 7.56$; OMS, $M = 89.33\%, SD = 6.88\%$; NMS, $M = 89.90, SD = 6.60\%$). Finally, there was no interaction effect, $F(2, 51) = .88, p = .42$.

In summary, supporting our prediction, using model statements did not make the children’s true reports any less accurate than those that were elicited when no model statement was used.

**5.5 Discussion**

The present study aimed to investigate further the effects of using model statements on children’s true and false reports. It looked to extend previous research with children (Brackmann et al., 2016; Chapter 3) by transforming practice interviews into children’s
own model statements as a substitute for listening to the amount of detail another child can provide.

Our findings revealed that it was only when children heard another child’s model statement (AMS) that their interviewee productivity increased. Hearing an excerpt of their practice interview (i.e. their own model statement, OMS) had no effect on the number of details they provided in their second recall attempt when talking about either the true or the false activity. The lack of any significant change in interviewee productivity for children who heard no model statement (NMS) demonstrates that an increase for the AMS condition is not an artefact of repeated questioning. These findings for AMS support previous studies (Chapter 3; Ewens et al., 2016; Leal et al., 2015) that also found an increase in productivity for both true and false reports when using AMS. Examining the theoretical underpinnings of AMS revealed that this positive change in interviewee productivity was solely related to having high perceived self-efficacy beliefs for being able to provide an accurate account. Although children did report using AMS as a social proof (Cialdini, 1993) more often than not, social comparison type was not associated with any change in interviewee productivity. This finding is contrary to that of Ewens et al. (2016). It might be that their type of comparison (downward, upward or neural) affected the current participants’ interviewee productivity in a way that was not measured in the present study.

Using an extract of a child’s practice interview as his/her own model statement did not have the desired effect of increasing interviewee productivity, which could be due to no relationships being found between self-comparison, perceived self-efficacy and interviewee productivity. In the field, practice interviews are typically conducted just prior to the substantive phase of questioning (Lamb et al., 2007). In the present study, there was a two-to-three-day delay between the practice interview and the interview about the target
events. This could indicate that training effects of practicing retrieval wear off in the long-term (i.e. after 48 hours) and simply playing a reminder of how detailed the children were a couple of days prior to the interview is insufficient in terms of re-training. Another explanation could be the brevity of the practice interviews conducted in the current experiment. The benefits of conducting a practice interview can depend on the quality of the practice interview carried out, with only good interviews eliciting significantly more detail than no practice interview (Price et al., 2013). Short practice interviews may not have helped build up a framework for how to answer open-ended questions, meaning that when they were played back the children were not reminded about techniques for providing a complete report, and so performed the same as children in the no model statement condition.

The *between-statement consistency* of children’s true and false reports was solely affected by the use of AMS in that the children in the AMS condition provided significantly more new information (commissions) in their true and false reports compared to the children in the OMS and NMS conditions. This is line with previous hypotheses (Bruck et al., 2002; Granhag et al., 2003) that longer statements, whether truthful or deceitful, will be more inconsistent. Eliciting more truthful commissions is beneficial to a police investigation because it provides new leads. Moreover, in the current study, no omissions were made, that is, none of the original information given by the children was left out to make room for the new information. Increasing the level of inconsistency in false reports could be beneficial if the inconsistencies in these reports are accurately perceived by lay person juries (Fisher, Brewer & Mitchell, 2009). An increase in inconsistent information, however, works to the disadvantage of truth-tellers in the courtroom whose more inconsistent reports might not be believed by jurors.
In response to Talwar and Crossman (2012), we examined the adverse effects of using model statements on accurate reporting. Unlike Chapter 3 where we gave no forewarning that the target event would take place, we told our child participants that they would later be questioned about the activities. In line with Brackmann et al. (2016) whose child participants were also able to give the target event their full attention, memory accuracy was maintained regardless of interview protocol. Children were, therefore, able to become more informative following AMS without sacrificing the accuracy of the information they gave (Ackerman & Goldsmith, 2008). In the current study, informativeness can be measured by quantity of detail as well as the quantity of correct/incorrect details. No increase in the latter two variables runs counter to our result that those in the AMS condition became more productive after hearing the model statement. This contradiction might be a consequence of our coding scheme: as we were only able to code for standardised aspects of the activities (i.e. aspects that we could code as correct/incorrect with certainty), there were many details that were not coded. To carry out a more complete assessment of the effects of model statements on truth-tellers’ accuracy, the to-be-remembered event could be filmed to capture everything that happened. However, this would require hidden cameras so that children’s behaviour is not affected. This is a commodity that was not available to us.

For the current study we used a short and non-elaborative interview protocol, which made it practically dissimilar to real police investigations where more thorough interview protocols are used (e.g. NICHD investigative interview protocol, Lamb et al., 2007; Revised Cognitive Interview, McCauley & Fisher, 1995). The decision to only ask for free recalls was, first, to reduce fatigue effects on interviewee performance, and, second, to focus our attention on how using a model statement provides a framework for answering open-ended questions. ‘Tell me everything that happened’ is an open request that children
can find hard to answer without guidance (Lamb et al., 2007; Saykaly et al., 2016). The present study demonstrated that using AMS helped children to provide more details; particularly more new details in their true reports that were no less accurate than when responding to an initial ‘free recall’ request. The child participants in the current study were telling true or false reports about enjoyable play activities, which were arguably not generalizable to the topics of real police investigations. Level of stress can have a significant effect on the linguistic patterns in children’s true and false reports, resulting in different verbal predictors of veracity (Brunet et al., 2013). Replicating the present study with a stressful activity could, therefore, result in different verbal behaviours. Finally, we used simple self-reporting methods to evaluate comparisons to the model statements and perceptions of self-efficacy. However, these ratings are not always a true reflection of the actual state of affairs. In the present study, 50% of the children in the AMS condition reported that they performed the same or better than the child who gave the model statement for the AMS condition, yet analyses showed that, on average, the group performed significantly worse than him. Thus, the lack of relationships between these variables could be due to inaccurate and unreliable measures.

Future research should continue to examine and clarify the social and cognitive mechanisms that influence interviewee productivity when using model statements, such as perceived self-efficacy. In the present study, all children reported high levels of self-efficacy, thus the results are only generalisable to children who strongly believe that they can provide informative and accurate reports. A more rigorous measure of self-efficacy needs to be established that allows for greater variations in children’s perceived beliefs in their recall ability to be revealed (e.g. adapted from the Self-efficacy scale in the Motivated Strategies for Learning Questionnaire, Pintrich & De Groot; Pintrich, Smith, Garcia & McKeachie, 1993). This would help to tease apart how using a model statement creates a
relationship between perceived self-efficacy and increased interviewee productivity. It would also be more methodologically sound to capture participants’ self-efficacy beliefs after hearing a model statement (but before a second recall attempt) to measure the direct effect of hearing a model statement on their perceived ability to provide an account of the target event. In the present study, self-efficacy ratings were taken post-interview to ensure that children were blind to the aims and hypotheses of the study. In addition, we could have gone beyond self-efficacy to investigate performance motivation and questioned children about their goal orientation. In the educational literature, task-focused goals have been linked to high self-efficacy (Wolters, Yu & Pintrich, 1996), which could explain why self-efficacy was not associated with interviewee performance for all children.

In conclusion, there is something about the presentation of a model statement given by another child that makes children say more. This effect is not present when children are played their own practice narrative, nor is it when they are presented with no model statement at all. In the present study, interviewee productivity increased for both true and false reports. The theoretical mechanisms underpinning these findings require further exploration.
5.6 References


Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J, & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology*. Advance online publication. doi: 10.1002/acp.3270


Chapter 6: Elicit, evaluate, empower and end product: A basic framework for designing interview strategies to uncover children’s false reports
6.1 Abstract

Focus groups with police officers, who regularly interview child witnesses, were held to evaluate current research into cognitive lie detection strategies with children in terms of their relevance to child interviewing practice. Sixteen participants across three focus groups provided practitioner feedback on the scope for incorporating lie detection techniques into child interview practice. Conducting thematic analysis on the focus group transcripts revealed advantages, disadvantages and suggestions for use for some, but not all, of the techniques. Further thematic analysis of the transcripts highlighted four endeavours that these police officers were trying to pursue in their child interviews: The 4E’s. First, eliciting as much information as possible remains the primary focus of interviews. Second, evaluating the credibility of that information must be considered within the bigger picture of the incident. Third, enabling the interviewee is important both before and during the interview. Finally, producing an end product that will be accepted in court and will protect a child witness is key. These broad aims provide a simple framework against which future lie detection techniques with children could be assessed when considering their transition from theory into practice.
6.2 Introduction

As previously mentioned in the Introduction to this thesis (Chapter 1), the aim of the PhD programme was to explore the viability of cognitive lie detection techniques that had demonstrated promising results with adults and had yet to be tested with children. The previous chapters have demonstrated that imposing cognitive load using gaze maintenance (Chapter 2) and encouraging interviewees to say more using model statements (Chapter 3) can exaggerate behavioural differences between children’s true and false reports, with varying effects on improving adults’ ability to determine children’s credibility. Thus, there is already evidence that cognitive lie detection techniques could present a new credibility tool for uncovering children’s deceit. In this chapter, we will explore the issue of ‘practitioner use’ highlighted by the Joint Inspection conducted by Her Majesty’s Inspector of Constabularies and Her Majesty’s Crown Prosecution Service Inspectorate (Criminal Justice Joint Inspection, 2014). To recap, this report noted low compliance with techniques that are currently recommended, which aim to discourage dishonest behaviour. Any new technique must, therefore, be practically favourable, as well as being empirically supported by the literature. Interview techniques that are unlikely to be used by child interviewers may not have the desired effect of facilitating child deception detection.

The transition from laboratory studies to practical use in the field requires a two-way conversation between academics and practitioners. Through this feedback loop, academics first explain how psychological mechanisms can be manipulated in theory, and practitioners then highlight aspects of the target environment that might inhibit this effect in practice. Academics should then alter their research paradigms to account for these ‘real world’ considerations and so the conversation continues. This feedback loop has facilitated the revision of many investigative interview techniques for use in the field (e.g. the Cognitive Interview, Dando, Wilcock & Milne, 2008; the National Institute of Child

The practical value of using cognitive lie detection techniques with adult interviewees has already been evaluated. Vrij, Leal, Mann, Vernham and Brankaert (2015) conducted training workshops with experienced police detectives in which they were briefly trained in how various cognitive lie detection techniques work and how best to utilise them effectively. The police officers in this study were then free to interview a mock suspect incorporating as many of the newly learnt techniques as they wished. Only 7% of the questions asked during the post-training interviews were related to the taught techniques, and when police detectives reported that they had implemented certain techniques, it was not in line with the way in which they had been trained to do so. Poor training effects might be due to a brief training workshop rather than poor compliance; the workshop only lasted 5.5 hours and covered brief descriptions of 10 techniques meaning that the police officers may not have fully understood the techniques. Furthermore, the participants were not given the opportunity to actually practice incorporating the techniques into their interviews nor were they given any feedback. Instead of disliking the techniques, the police detectives may have simply not understood how to implement them. A further limitation of the Vrij et al. (2015) study was that only quantitative performance measures were taken. Therefore it was not possible to ascertain the police officers’ perceptions of the practical value of the cognitive lie detection techniques. As the aim of this study was to achieve a more in-depth appraisal of the use of cognitive lie detection techniques with children, a qualitative methodology was used.

Another advantage of using a qualitative design was that it aided our understanding of the broader context in which child interviewing takes place. It was hoped that this might
help us to understand the finer details of the investigative process that would come from practical experience rather than a literature review. Previous laboratory studies have aimed to mirror real world interviews with children as much as ethically possible. For example, by using samples of maltreated children who are already witnesses/victims within the criminal justice system (Lyon & Dorado, 2008; Lyon, Malloy, Quas & Talwar, 2008), ‘touching’ games whereby an adult touches parts of a child’s body (in the context of the game) and the children are later interviewed, (Edelstein, Luten, Ekman & Goodman, 2006), witnessing a theft and reporting it to a ‘police officer’ (Tye, Amato, Honts, Devitt & Peters, 1999), and talking about true and false stories about stressful injuries (Brunet et al., 2013). Nevertheless, these studies can never truly reflect the diversity of the children interviewed and the diversity of the circumstances that led up to those children being interviewed, which come with real field investigations. As strict ethical considerations apply when testing sensitive populations (Fisher, 2005) we must turn to the practitioners, who see these child witnesses’ progress from first disclosure to final verdict in court, to understand the intricacies of this process and how these intricacies might be in conflict with the methodologies that we are investigating.

The present study, therefore, aimed to put our child deception theory into child interviewing practice by asking police officers, who regularly interview children, to evaluate the interview techniques investigated in Chapters 2 to 5. A secondary objective was to understand the broader ‘child interviewing’ context in which the interview techniques would be implemented. It was hoped that this would help to identify any overarching compatibility issues that affect the transition from lab to field and could help when designing and/or adapting interview techniques in the future.
6.3 Method

6.3.1 Participants

Each focus group was made up of 5-6 police officers ($N = 16$, 10 females, 6 males) that regularly interview child witnesses. Focus groups 1 and 3 were from the same police constabulary in the South of England, whilst focus group 2 was from a different police force in the North of England. Overall, participants had 3 to 40 years of experience working for the police ($M = 17.09$ years, $SD = 8.16$) and 5 months to 22 years of experience interviewing children in their police role ($M = 8.56$ years, $SD = 6.16$ years). There were no differences between focus groups, in overall police experience, $F(2, 13) = .10, p = .90$, nor for interviewing children, $F(2, 13) = .41, p = .67$. In terms of police ranking, 7 participants (43.8%) were detective constables, 6 (37.5%) were police constables and the remaining were a police staff investigator and two Tier 5 interview advisors. All participants reported (except one person who didn’t respond) having undergone an ABE interviewing course. The age of the children that participants routinely interviewed ranged from 2 to 17 years old for the majority of the sample (62.5%). The rest of the sample interviewed adolescents (13 to 17 year-old children) only.

6.3.2 Procedure

Police officers were asked to give informed consent to their participation. Following this, the Principal Investigator (PI) outlined the layout of the sessions and ground rules for the focus group discussions. For each session, the PI gave a 20-minute presentation regarding her cognitive lie detection findings with children for the following techniques: imposing cognitive load through gaze maintenance (Chapter 2), using a model statement provided by another child (Chapters 3 and 3), and using a self-generated model statement derived from a practice interview (Chapter 5). For the first and second techniques, this involved a full explanation of the theoretical underpinnings of the studies.
as well as the corresponding results. As data collection had not yet been undertaken for the third technique (Chapter 5), only the theory behind the expected outcomes of transforming a practice interview into a self-generated model statement was presented.

The presentations were then followed by semi-structured focus group discussions covering the following topics: current training/techniques to assess child credibility, evaluation of cognitive lie detection techniques with child witnesses, the role of age in assessing credibility and finally the participants were asked to highlight anything else that the PI’s programme of research had overlooked. The focus groups were audio-recorded using a Dictaphone and later transcribed verbatim.

6.4 Results

6.4.1 Analytic strategy

The focus group transcripts were coded inductively using thematic analysis. As the primary focus of the research was to gain practitioner feedback on the viability of the interview techniques, discussions relating to each of the three techniques were first analysed separately for emerging themes. Next, all the data across all three focus groups was compiled to code for broader themes that represented how the police officers understood the process of child witness interviewing in the UK. For both levels of analysis, the focus group transcripts were, firstly, summarised and reduced down to smaller “packets” of information, which were then analysed for themes in the data (Boyatzis, 1998). Emerging themes had to reflect similarities across multiple participants’ responses. A table displaying the number of participants who represented each theme and the overall contribution each participant made is displayed in Figure 6.1 on the next page.
<table>
<thead>
<tr>
<th>FG</th>
<th>P</th>
<th>AD</th>
<th>DIS</th>
<th>SUG</th>
<th>Extraction Information</th>
<th>Evaluating Information</th>
<th>Empowering the Interviewee</th>
<th>End Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Figure 6.1 Summary of themes highlighting participant contribution per theme. FG = Focus group. P = Participant letter. AD = Advantages. DIS = Disadvantages. SUG = Suggestions for use.
For the evaluation of techniques, three themes emerged: *advantages, disadvantages* and *suggestions for use*. However, these three themes were not represented across all three techniques (see Figure 6.2).

For the broader child-interviewing context, four themes emerged: *eliciting* information, *evaluating* information, *empowering* the interviewee, and the all-important *end product*. Subsequent analysis broke these four themes down into sub-ordinate categories – two subordinate categories emerged from the data for each of these four superordinate themes (see Figure 6.3). The thematic structure was discussed and agreed upon with the PI’s PhD supervisor. For the following quotes, ‘FG’ refers to the focus group number and ‘P’ refers to the participant letter that was assigned to each member of each focus group. […] depicts the deletion of unnecessary text in the quote.
6.4.2 Evaluation of cognitive lie detection techniques

Three main themes emerged from participants’ responses that captured their appraisal of the three cognitive lie detection techniques presented: Advantages, Disadvantages, and Suggestions for use. Below we will present each of the superordinate themes with the relevant comments for the interview techniques.

6.4.2.1 Advantages

The first theme encompassed the advantages of using such techniques with child witnesses. That said the participants discussed the benefits only relating to the use of children’s own model statements. Thus, they did not identify any benefits of asking
children to maintain gaze nor did they speak positively about playing child witnesses another child’s model statement.

6.4.2.1 Child’s own model statement

Transforming a practice interview into a model statement was considered to be; “the most viable option” (FG2, PC), with 3 participants (19% of the sample) providing advantages regarding its use. Participants commented that it would be a good reminder of what was expected of the children at interview; “then they would realise the detail that was needed” (FG1, PB), and it would also have “trained them in five/ten minutes to answer open questions and to give higher levels of detail” (FG2, PC). The self-generated nature of using a child’s own model statement was considered to be particularly beneficial because it would provide a useful baseline for that child, “demonstrating that child’s cognitive ability to remember” (FG2, PA) and “establish[ing] before the interview what their ability is” (FG2, PC). The participants couldn’t see any problems inherent in playing children’s own model statements back to them because “it’s self-generated… it’s always going to be age-appropriate, development appropriate” (FG2, PC).

6.4.2.2 Disadvantages

This second theme encompassed participants’ perceptions of the problems inherent in using cognitive lie detection techniques with child witnesses. Although participants discussed the disadvantages for all of the techniques presented, there were similarities and differences in problems highlighted across techniques.

6.4.2.2.1 Maintaining gaze

The disadvantages of asking child interviewees to maintain gaze could be summarised as: (1) inappropriate for sensitive offences, (2) in conflict with current practice, and (3) simply impractical. None of these disadvantages overlapped with those of
using another child’s model statement or a child’s own model statement. Firstly, maintaining gaze was considered inappropriate for sensitive offences. From their practical experience, participants reported that when children were interviewed about a traumatic event, it was quite difficult to establish eye contact, even to the extent that “the only way she would talk to me was by doing her colouring so she wasn’t even looking at me at all” (FG1, PD). This could be due to the discomfort children feel when talking about sexual incidents; “they can’t look you in the eye because to them they shouldn’t be saying it in the first place” (FG2, PB). Feelings of embarrassment can also mean that they’ll ask to write words down because “there’ll just be some words that they just can’t say” (FG2, PC). Sometimes the trauma of the incident can be such a barrier to eliciting information from children that there was concern that they wouldn’t elicit “as much information from them” (FG1, PD) when using a gaze maintenance instruction.

The second disadvantage was that requiring children to look at them ran counter to the instructions that they already gave:

Quite often in the introductions we will say, “If you find it more comfortable to close your eyes or look somewhere else, please do so ‘cause if that helps you concentrate that’s better for us ‘cause we want as much information as we can get.” (FG1, PB)

On the one hand, participants could understand why asking children to maintain eye contact would make it harder for them to tell a lie because it would be an added pressure. On the other hand, they explicitly said that this wasn’t their aim; “we’re not after
trying to put the child under pressure, we’re after trying to establish the truth” (FG1, PB). It was felt that the interviewer should never be seen to be “overpowering” (FG3, PA), and that maintaining gaze “would be potentially considered oppressive in a practical setting” (FG2, PC). Indeed, the added burden of establishing eye contact with the interviewer could have an adverse effect on truth-tellers; “if they have established eye contact and they don’t tell us [what happened], we can’t safeguard them” (FG1, PB). In other words, participants highlighted that if the techniques don’t help to elicit more information from children then they are counterproductive in an interview setting.

Finally, the third disadvantage was that using gaze maintenance with children was impractical. During interviews, children are often mobile; “some of them, they don’t even sit on their bottoms, let alone sitting there and looking at you” (FG1, PA), or completing other activities, like drawing pictures, that help to elicit more information; “it’s very difficult to maintain eye contact in a real situation because [...] they’ll describe the room and they’ll [the interviewers] say, “draw that for me” so then they’re drawing” (FG2, PB). An instruction to draw what happened would, therefore, be in conflict with an instruction to maintain gaze and could leave children wondering with which instruction they should comply. It might also be impractical because children have short attention spans. People don’t naturally maintain gaze so “you would have to repeatedly remind them” (FG2, PC) and even then children might not comply because “for some young children that would be quite hard to do it all the way through” (FG1, PB) and some children have got attention difficulties; “with the ADHD and Autism Asperger they’re not going to sit there and maintain eye contact with you” (FG3, PC). Finally, requiring gaze maintenance was not considered feasible because it would increase the cognitive load experienced by the interviewer, which is already high:
The cognitive requirement of a child interviewer is unrecognised I think by a lot of people because, if you think, there’s you in the room with a child, you’re thinking about what do I know about the job, what’s my interview plan, my questioning style, keep them on track, encouraging them, welfare, who’s outside the room, what facts do I know that the child doesn’t know so that I don’t contaminate. (FG2, PC)

Furthermore, it would be difficult for practitioners to maintain eye contact in return because they are trained to take brief notes during interviews (Marlow & Hilbourne, 2013). It was noted that children might find it unfair if the interviewer isn’t always looking at them; “every now and again you’re looking down, writing a little word and then looking back up at them, they’re thinking, “well you’re not maintaining eye contact with me”” (FG2, PC). So, poor compliance could also be due to the children being put off by the lack of reciprocated gaze.

6.4.2.2.2 Another child’s model statement

The perceived disadvantages of playing a detailed example of another child talking about an unrelated (i.e. another child’s model statement) could be summarised as: (1) tainting, (2) longer interviews, and (3) providing an inappropriate example. The last two disadvantages were also expressed in relation to the ‘child’s own model statement’ technique. Firstly, participants stressed the importance of eliciting a completely untainted account and the criticism they could face if there were any signs of coaching; “we have to be totally transparent and impartial and we get criticised that we’re goading or that we’re influencing children or that we’re preparing them for the interview” (FG1, PA). In particular, the participants would be reluctant to play an account provided by another child because it could contaminate their child interviewee’s account; “it could be argued by the
defence that the child had picked up some details from the account to embellish their own account” (FG1, PD). One participant even went so far as to say that, “it would only take one tiny similarity for the defence to say, “well if that’s similar, the whole thing is nonsensical” (FG2, PD), which demonstrates that they think that playing another child’s model statement is a risk that simply isn’t worth taking. Research, into using another’s model statement, states that it should be unrelated to the to-be-remembered event (Leal et al., 2015; Vrij, 2015). However, the effects of using another’s model statement on source monitoring have yet to be directly tested and, thus, cannot be assumed. Source monitoring (Johnson, Hashtroudi & Lindsay, 1993; Lindsay, Allen, Chan & Dahl, 2004) could be a potential issue. The extent to which the ‘model statement’ event needs to be conceptually dissimilar to the witnessed event is likely to run along a continuum of high to low risk.

Secondly, lengthening the interview would “eat into your best evidence” (FG1, PA). As some participants had experience of interviewing young children, they reflected on how the children’s short attention spans already reduced the amount of time that they had to interview them in the first place. Adding in a model statement section “could basically burst the bubble” (FG3, PB), confusing the child and resulting in less information being elicited before they get tired or lose concentration.

Finally, playing another’s child model statement was considered to be inadvisable because it would be an inappropriate example in terms of its age appropriateness and in terms of giving child lie-tellers a template. Firstly, there were concerns that another child’s ability to provide a detailed account may not be realistic for the child being questioned; “it would be inappropriate expectations for that child [truth-teller]” (FG2, PC). In other words, the child being interviewed may simply not be able to give as much detail as the child in the model statement. Furthermore, the length of delay between event and recall could be longer or shorter for the child in the model statement compared to the child
witness, with the effects of delay on recall ability varying from one child to the next. Even playing the model statement of a child of the same age as the child interviewee would not be appropriate due to developmental differences; “another one recently she couldn’t string a sentence together and was four so even if you’re balancing it off against another four-year-old, this little girl’s not got the capability of even saying three words in a row” (FG2, PB). On the other hand, one participant questioned:

> Are we just giving them [child lie-tellers] more ammunition to say, “well this is how it’s done” so they’re already thinking and exaggerating their account a little bit to fill it up so that they’re giving you more detail? (FG3, PC)

Overall, participants were suggesting the possibility that another child’s model statement, provided by another child, might raise the standards too high for child truth-tellers but give child lie-tellers the means to achieve these higher standards.

**6.4.2.2.3 Child’s own model statement**

As previously mentioned, both ‘longer interviews’ and ‘providing an inappropriate example’ were disadvantages that also emerged when participants discussed using a child’s practice interview as their own model statement. The concerns regarding longer interviews largely overlapped with previous issues for using another child’s model statement. Participants felt that adding this extra component into the interview protocol would take away precious time that could be used to gather information. In particular, it might interrupt the child’s willingness to talk:
You take them to an interview, they’re ready and they’re eager to just […] talk to you. So if you then delay that by doing a little practice with them say, they might lose momentum. (FG3, PC)

Reflecting on current practice, one participant followed up that comment by saying, “we can limit some just with our introductions if they’re too long. Sometimes you’ve just got to let them talk straightaway” (FG3, PD).

The participants’ concerns regarding ‘providing an inappropriate detail’ differed, however, from another child’s model statement when discussing its effect on truth-telling children. Participants felt that practising talking about a non-traumatic event might elicit a detailed response but when this was used as a basis for talking about a traumatic detail, then the expectations of providing detail might be unachievable. There were two ways in which participants thought that children might deal with unrealistic expectations; “they might feel that they’re letting you down and themselves down” (FG1, PD), or “they [might] feel that they might need to add in more detail to the incident that they don’t remember much about” (FG1, PC). Either way children’s reactions might be negative. In terms of their own model statement being an ‘inappropriate example’ for lie-telling children, the feedback, similar to that for using another child’s model statement, suggested that the child lie-tellers would use this to elaborate their lie:

The ones that are lying, are we just giving them more ammunition to say, “well this is how we want the interview done” so then they’re already thinking and
exaggerating their account a little bit to fill it up so that they’re giving you more
detail? (FG3, PC)

The same participant also suggested that child lie-tellers might use the practice
interview to check that they can lie convincingly by lying then too. The interviewer might
then, inadvertently, reinforce lie-telling behaviour; “the only worry would be
reinforcement, “oh it’s very good because you’ve lied your way through that practice
interview and I’m telling you how perfectly you’ve done”” (FG3, PC).

6.4.2.3 Suggestions for use
The third theme in terms of the lie detection techniques concentrated on suggestions for
practical use, namely ways in which the techniques could be adapted for use in real life
investigations. Participants did not include any suggestions for use regarding the use of
another child’s unrelated narrative to encourage interviewees to say more. Participants did
provide suggestions for adaptations of the use of gaze maintenance and playing a child’s
own model statement in the field that largely linked to the disadvantages highlighted for
those techniques.

6.4.2.3.1 Maintaining gaze
To combat the additional load of continually asking a child to maintain gaze on the
interviewer, it was suggested that, “you would have to engage the help of maybe the note
taker or the co-interviewer who sits in another room” (FG2, PC). In other words, have a
co-worker who is already involved in the interview to share the load. Unrelated to the
disadvantages mentioned, participants also suggested that asking a child to maintain gaze
with the interviewer “could be used as a follow-up technique” when “there’s something not
right” (FG2, PC). It would, therefore, not be used during the whole interview but “at specific parts” (FG3, PD).

6.4.2.3.2 Child’s own model statement

Suggestions for use for this technique divided into ‘who’ it should be used with, ‘how long’ it should last and ‘when’ it should take place in the investigative process. Firstly, participants thought that, “you could do it for certain offences possibly [...] if it’s not something so traumatic” (FG1, PB). This would resolve the issue of children not being able to give as much detail as they did about the non-traumatic event in their practice interview when talking about the actual offence. Other individual characteristics, such as age; “for the older people” (FG3, PA), or, such as disposition; “you might get a really nervous one you would just need to get them talking and get them to feel comfortable then a practice interview might be good for that” (FG3, PC), might make it more suitable for inclusion in child interviewing practice.

To reduce its negative effect on the overall duration of the interview, it was suggested that brevity would be the aim of the game; “the practice interview is a good idea if it’s short and succinct” (FG1, PB), “just like a very short operation” (FG3, PA). If it was kept to five or ten minutes long, then this should not affect children’s concentration. Furthermore, if it was not captured on the disc shown in court, then that would be additionally beneficial because “it’s all about time, time, time in the courts” (FG2, PA).

Finally, in relation to keeping interviews short, it was proposed that “the best bet” (FG1, PA) would be for the practice interview to take place and be played back to the child as part of an intermediary assessment (i.e. as a separate process from the interview itself). At court, it could be entered as “unused material” (FG2, PC), which the judge could call upon if s/he wished to know how detailed the child could be (i.e. their baseline). It was
further suggested that researchers should ask for intermediaries’ feedback on using a child’s own model statement; “perhaps use the intermediaries to say actually this would be a really good way to assess this child’s ability” (FG3, PD).

### 6.4.3 Broader child interviewing context

Four superordinate themes emerged from participants’ responses that revealed aspects of the child interview context to consider when designing new lie detection strategies. These can be referred to as the 4E’s of child interviewing: *eliciting* information, *evaluating* information, *empowering* the interviewee, and the all-important *end product*.

#### 6.4.3.1 Eliciting information

Many participants agreed that the interview was primarily an information-gathering procedure rather than an assessment of credibility; “it’s not our role is it to think about are they lying? You know we’re just there to get the information” (FG1, PA). Discussions regarding the elicitation of information from child witnesses divided into two subordinate themes: *techniques* to elicit more information and *internal barriers* to eliciting information.

##### 6.4.3.1.1 Techniques

This sub-category referred to any methods that the interviewers use to encourage their child interviewees to say as much as possible. This included encouraging the children to concentrate; “if they’re not looking at you they’re not concentrating […] you say their name […] and then you can ask them the specific question again” (FG1, PE), or to do whatever makes them feel most comfortable; “we will say, “if you find it more comfortable to close your eyes or look somewhere else, please do so ‘cause if that helps you concentrate, that’s better for us” ‘cause we want to get as much information as possible” (FG1, PB). Participants from both focus groups in the South of England
mentioned a specific technique that they use at the beginning of interviews to elicit longer responses later on:

I’ve got a cup here, for example, and I’ll give them an example of how much detail I need. I might say to them, “so tell me what you see about that cup” and they’ll say, “oh it’s a cup, you know it’s cream in colour” and I’ll say, “but if you look at it now, the detail that I need is that it’s cream in colour, it’s got red writing on it, it’s got ‘wibble wobble’ and a picture of a jelly. Inside it’s dirty”. So I’d explain to them that that’s the detail that I’d want. (FG 1, PE)

I often provide an example and I try and tailor it to the individual. So if it’s a child, often I’ve picked the child up so I will talk about the journey we’ve taken to get there so it’s something they’ll identify with and often they’ll chip in with bits that I’ve missed out but I try to get them to consider all the senses and explain that as well […] which gives them an idea of the level of detail I’m looking for. (FG3, PD)

It is notable how much this example of detail, provided by the interviewer him/herself overlaps with playing a model statement from another child yet the former does not receive any opposition. These subtle nuances between what works with children and what would be inappropriate with this witness population may come from practitioners developing a certain sensitivity towards the difficulties experienced by children and the constraints of the court system. Unlike academic researchers who are ‘non-practicing’, the ‘practicing’ child interviewers have learnt to identify and subsequently navigate the
limitations of the child interviewing process, making it easier for them to highlight problems that are common-sense to them but not to the researcher. Future research should investigate the interviewer-generated model statement as a technique to facilitate information elicitation.

Overall, participants agreed that, “anything that increases the chances of that child disclosing more details is good” (FG2, PC). With fewer questions being particularly beneficial; “the best one is where you don’t have to ask them an awful lot of questions […] something that would get people to give more information without asking […] too many questions” (FG3, PC). Indeed, police officers do tend to ask more questions than are necessary during an investigative interview (Vrij et al., 2015). Participants in the current study understood that they needed to juggle between creating a witness that was both informative and reliable. The number of questions asked may, therefore, be dependent on interviewers’ intuition. They must know when they’ve elicited enough information for court and they must not push the boundaries too far so that any additional information thereafter becomes inaccurate.

6.4.3.1.2 Internal barriers

This sub-category referred to any individual characteristics of the child witnesses, which might prevent the interviewer from getting as much information as possible. Forgetfulness was one internal barrier to recovering information about an incident. Children, particularly young children, might unintentionally begin to forget what happened, or they might intentionally exclude an aspect of the incident that they don’t want to talk about. When considering complicating factors for disclosures of child sexual abuse, legal professionals have been reported to consider limited memory capacity as one of the least complicating factors, with emotional factors, such as guilt and shame being the most important (Leander, Christianson, Svedin & Granhag, 2007). Thus, this forgetfulness
may not be the result of a memory error, but rather deliberately triggered to avoid negative emotions related to the incident in question. Like the police officers in Leander et al. (2007), the participants in our study recognised several emotional factors, such as shame; “it’s just they’re either ashamed of what’s happened, they realise it was wrong what’s happened to them and they don’t want to talk about it” (FG1, PE), self-criticism; “they miss bits out […] where they think they may have been part of, you know, partly to blame or something because they put the blame on themselves” (FG1, PB), and embarrassment; “they’ll just be some words that they can’t say […] but we need them to actually say what that body part is and that’s sometimes a real barrier” (FG3, PC). Also in terms of willingness and readiness to talk, children who are not the first people to disclose may not wish to talk about what has happened to them (Leander, Christianson & Granhag, 2007):

Often the allegations that come to us haven’t come from the child themselves so it’s not that the child’s been ready to talk about it […] so we go out with our tick list of lines of inquiry […] but actually for that person, if we go in there too soon […] then they’ll just drip feed you a tiny bit of information and they won’t actually be telling you everything that’s happened to them. (FG3, PC)

Indeed, more formal disclosures of child sexual abuse do occur at interview, when children have disclosed the alleged abuse before interview (Magnusson, Ernberg & Landström, in press). Participants mentioned that it can even be difficult in the first instance to get children to assent to an interview; “you’ve got people that crikey it’s enough getting them through the door” (FG3, PA), and then actually getting them to give a statement is also hard because “they don’t want to be there in the first place” (FG3, PA). A
reluctance to be interviewed was considered by participants to be partly attributable to feelings of loyalty towards the offender, who is often a close family member; “I had a teenage girl who was sexually abused by her biological dad […] he got convicted but she, all the way through, found it so difficult talking about it ‘cause she still loved her dad […] very loyal” (FG1, PA). Feelings of loyalty towards a perpetrator, who is often known to the child (London, Bruck, Ceci & Shuman, 2005), can cause children to withhold sensitive information regarding the incident (Christianson, Azad, Leander & Selenius, 2013; Magnusson et al., in press). Furthermore, children can also understand the consequences for disclosing abuse and may wish to avoid these negative outcomes by keeping quiet:

Younger children, they’ll cover up for their parents […] it has happened to them but they don’t want to be removed and they know the consequences if they tell us things so they try and hide as much as possible. (FG2, PB)

Whether children choose to keep quiet due to their own internal barriers, or are asked by the perpetrator to keep the incident a secret (Magnusson et al., in press), the concealment or omission of relevant information does constitute lie-telling (Vrij, 2008). However, children keep these secrets due their emotional cost (Last & Aharoni-Etzioni, 1995), rather than an outright desire to convey information that is completely different from the truth. It is, therefore, important to recognise these emotional barriers and to design better interview practices that make children feel safer and reduce this perceived emotional cost if they do disclose.
Participants in the current study suggested that children’s loyalty towards their peers can also produce a ‘them-and-us’ mentality towards the police creating negative attitudes about talking to police officers:

They’ll call us pigs and they won’t talk to us and it’s how they have been brought up and the perception they have of police […] they won’t tell us what’s happened ‘cause they’ve learnt from a young age, don’t talk to police officers […] it’s quite bizarre when we get little ones telling us to f**k off. (FG1, PA)

The attitudes of parents can therefore trickle down to be the attitudes of the children; “it’s not the children, it’s the parental guidance. The parents are demonstrating an indifference to anything formal” (FG2, PC), resulting in difficulties when it comes to eliciting information at interview. When asked whether a social worker could conduct the interview instead, the response was in line with ABE guidance (Ministry of Defence, 2011) that this would be a good idea if the social worker has better rapport with the interviewee than the police officer. However, it was explained that social worker require much training to lead an interview and to be trained to “home in on that offence” (FG 1, PB). This sometimes led to a reluctance from social workers to interview because “if they ask the wrong question they’ve got the possibility that they could be called to court to be asked why they asked that question” (FG, PE) so ultimately “they don’t want to put themselves at risk” (FG1, PD). Furthermore, it was noted that, “with teenagers, you’ll find they quite often don’t get on with their social workers” (FG1, PF). So it isn’t always an alternative solution.
This subordinate theme highlights the police officers’ knowledge of the internal barriers that could potentially disrupt the effectiveness of the techniques described in the first subordinate theme. This level of insight enables practitioners to develop intuitive solutions to ensure that their interviews are conducted to the best of their ability. This intuition is, however, not inherent to an academic researcher who has never interviewed children and wouldn’t look to incorporate these barriers into their research design. The child interviewers portray the child interviewing process somewhat akin to an obstacle course, which stands in stark contrast to the standardised, straightforward procedure used in most experimental designs.

6.4.3.2 Evaluating information

Whilst all participants in all focus groups agreed that it was not the role of the police officer to judge the credibility of an individual child, there were comments that they might get an inkling as to whether the child is telling the whole truth or not; “we get a feeling straightaway when we’re interviewing children. I had one the other day and I thought she’s not telling me the truth” (FG1, PA). The question is, therefore, what protocol should the interviewers follow when they get this ‘gut feeling’ that they should evaluate the veracity of the information they are receiving? From this superordinate theme, two sub-ordinate themes emerged: (1) how to assess child credibility, and (2) what if there are other factors to take into account when evaluating information provided by children.

6.4.3.2.1 How to assess child credibility

Although participants had not received any official training on assessing child credibility, there were general beliefs regarding how to determine veracity either through active techniques, or through passive evaluation of specific cues to deceit. Participants suggested that a good indicator of deceit was a lack of detail in children’s responses. This
could be a less detailed statement overall; “it is hard to get detail from someone that may be lying and they do sort of tie themselves up in knots and you will come out with not a lot of information” (FG1, PB), or few details regarding a specific aspect of the incident, “they can tell you details about everything around it but actually when you come to the act, that’s where you can often detect the lie because the detail within that is usually small” (FG3, PD). This evidence is in line with previous research on the general characteristics assessed by Criteria-Based Content Analysis (CBCA) that shows that higher levels of detail are consistently more indicative of a truthful statement and vice versa (Vrij, 2005). The participants in the current study, therefore, signalled a cue to deception that has been supported by empirical research. On the other hand, this belief sits in contrast to the above subordinate theme of internal barriers whereby silence or lack of detail could be due to the interviewer not being able to overcome the child’s internal issues to elicit more information. Moreover, if the child has been subjected to maltreatment, it could be that a lack of details is the result of poorly encoding the traumatic event (Gordon, Baker-Ward & Ornstein, 2001), and, thus not being able to recall as many details during their free recall (Eisen, Goodman, Qin, Davis & Crayton, 2007). Equally, it could be due to language and verbal difficulties related to the age of the child in question (Magnusson et al., in press).

Participants’ ability to navigate this minefield of reasons for less detailed reports highlights another nuance in interviewers’ perceptions. It would be of interest to explore further how police officers differentiate between lack of detail due to internal barriers and lack of detail due to deception.

Another indicator mentioned by participants was the age-appropriateness of the language used by children; “they might use a word that you think they wouldn’t normally use at their age, you know, that’s come from the mum or the dad or whoever […] it’s normally quite obvious” (FG1, PB). How obvious these linguistic markers are might be a
product of practitioners’ sophisticated sensitivity to the language of children from different age groups and backgrounds, which has come from years of interviewing a variety of children. However, we cannot yet know the accuracy of this assumption, since, to date, there are no studies that look at the effects of coaching on perceived age-appropriateness of language used by children in their subsequent reports. Police officers also need to be aware of the language used in the home and its effects on the children’s use of language.

In terms of active techniques to elicit these indicators and to evaluate them further, participants commented on “just ask[ing] the right questions” (FG1, PC). At first glance, this remark appears both common sense (i.e. obviously questioning style is critical) and ambiguous (i.e. what is considered to be the ‘right’ question could change from person to person). One participant shed light on one type of questioning style that could help:

People who prepare a story that they’re prepared to give you do not anticipate us going into the depth we do and the finer grain detail and you sometimes see elements of vagueness arising […] and it does then start to initiate your thought process […] I find some of the techniques about detail are where I get indicators, where I start to become uncomfortable about what somebody’s saying. (FG2, PA)

In other words, the use of probing questions following the free recall can be particularly revealing when it comes to evaluating the credibility of the information provided; “when you actually get down to the detail […] you can see it starts to unravel” (FG 3, PE). Participants reported that they were taught to make brief notes during interviews and that it was most time-efficient to “write down some key words for myself
so that I know right I’m going to come back and ask about that” (FG3, PC). Coming back to inconsistencies at the end of the interview was a point also stressed by the first focus group, who, on top of that, reinforced that “if there’s an issue it would be […] “I don’t quite understand, I’m a little bit confused about what you said earlier on so can we just go through that again” and try and clarify it, never challenge” (FG1, PB). Participants pointed out that the interviewer cannot directly accuse a child of lying and cannot be biased at interview so it’s about “trying to test [their] account at the very end without saying, “you haven’t told me the truth or you’ve told me different things.” We’ve got to be very careful” (FG1, PA). The interviewer has to strike that delicate balance between clarifying the information provided sufficiently to satisfy the prosecution but not too much that the defence can accuse them of treating the interviewee as if s/he was deceitful.

Participants recognised that techniques for evaluating the information provided could also take place after interviews have finished. In particular, looking at the bigger picture, it was considered that “credibility is corroboration basically” (FG3, PB). Thus, participants were inclined to take a step back and look at all the information elicited during interviews with multiple people when making credibility judgments; “that’s about getting accounts from different people. From the suspect, from the witness, stuff like that and then it builds a bigger picture rather than trying to work out if that individual’s being truthful or not” (FG1, PD). Knowing additional information regarding the case should also be left until after the interview has taken place so as not to taint the interview; “if you know too much about the investigation you will be inclined to ask questions about things they [the children] haven’t even told you about” (FG1, PF). Notably, asking the ‘right questions’ would, therefore, seem to mean questions that are impartial and only related to information that the child has already mentioned but are perhaps inconsistent with their other reports.
Finally, participants explained the difficulty of evaluating children’s statements that appeared to contain both truths and lies; “they tell half and half. Some of it might be true, some of it might not?” (FG2, PC). Known in the literature as embedded lies, this type of lie appeared to be a particular issue when it came to reports of incidents of sexual assault. Participants in Focus Group 3 mentioned that teenage girls would report that an ex-boyfriend, who they had recently split up with, has raped them on one occasion; “they’ll give you information about a time when they had sex, which you know is all truthful, and it’s the ‘how it was forced or ‘how they didn’t consent” […] which is the difficult bit” (FG3, PC). One participant in Focus Group 2 recalled a case of another girl who reported a catalogue of sexual offences against her, some of which were confirmed and some of which were not confirmed. However, the addition of false information discredited her entire account and all for the sake of more attention:

She used elements of truthfulness and confabulated lies around that […] and what exposed her was when she was telling us about events we were then able to check those events and recover CCTV footage and when we viewed the CCTV what she was saying happened, which came across very plausible and understandable, was something completely different […] the frustrating thing is we believed that some of it was truthful but she did so much damage to the investigation by telling us stuff we were able to show didn’t happen […] it was deliberately false memory that she was giving us to try and create all the attention she was getting. (FG2, PA)

In summary, embedded lies might be easy to detect from a practitioner’s perspective but unless the false information is verifiable, then it might be hard to prove
what is true and false. Furthermore, both child truth-tellers and child lie-tellers might embed falsities in their accounts.

6.4.3.2.2 Accounting for other factors

Even though practitioners are told, “it’s an indication that if people can’t provide detail that it’s not truthful” (FG3, PD), there are various contextual factors to take into consideration when appraising the quality of the information provided by children. Thus, the ‘bigger picture’ surrounding children can also be an indicator to their veracity. Firstly, it was considered unlikely for very young children to be lying in the first place because “they don’t really know what lying is” (FG1, PF). It was felt that as children became more capable of lying as they aged, they would be more likely to lie:

They [young children] are often brutally honest at that age. It’s not until they are sort of four/five that their minds, you know, allow them to [lie] and I think a lot of that comes from maybe being at school with older children that are able to lie about things. (FG1, PF)

On their own, very young children were considered to be incapable lie-tellers who required more mature models to learn vicariously the ability to lie. This idea of innocence for younger children was also related to them reporting on topics that were age-inappropriate; “you know, if they’re describing seeing […] their dad’s penis or something and they’re saying, “well it was pointing up and doing this”, the only way they would know that information is if they’d seen it” (FG3, PC). This subtle distinction between age-inappropriate language as a sign of deceitfulness and age-inappropriate situations as a sign
of truthfulness is, again, a nuance that might only occur to a police officer in the field but not to a researcher. That said the CBCA criterion ‘Accurately reported details misunderstood’ that is used as a cue to truthfulness, maps very well on to this observation from the field.

Secondly, participants put forward alternative explanations for brief statements that lack detail. If the children were younger, they were expected to give shorter statements and require more prompting compared to older children, who were expected to give “more self-initiated detail […] because they’re more articulate, they’re more intelligent” (FG2, PA). In other words, it was suggested that it might take more questioning to elicit the same amount of detail from very young truth-tellers compared to when older truth-tellers were being interviewed. It was also important to consider the effect of the trauma of the incident on the ability of the children to recall everything that happened:

Given that we’re often dealing with serious sexual offences that [saying they can’t remember part of the incident] might be as a result of trauma. So actually they can’t piece together things in a structured, chronological order, that it twists things in their mind or they can’t talk about it because it’s purely too traumatic. (FG3, PD)

Gaps in children’s memories might, therefore, by the result of poor encoding caused by the trauma of the event (Gordon et al., 2001). However, this is more likely to occur when children have experienced neglect in the home (Eisen et al., 2007). Practitioners, therefore, demonstrated an understanding of the effects of trauma on memory retrieval (i.e. psychology in an applied context). However, how our participants
differentiate between omitted details due to memory errors caused by trauma and omitted details due to the internal barriers mentioned above is still unclear.

Finally, the motivation to lie could be a result of children’s home life. On the one hand, children might be motivated to escape from a negative home life by exaggerating their story:

Every time she’s been interviewed she’s added more and added more [...] she’s developed this thing where six different men have apparently abused her to make sure she never goes back in and, I’m not saying it’s a definite lie, but it’s not really that plausible and she’s done it because, I suppose, in her life that’s her way to get out. (FG2, PB)

On the other hand, children might be motivated by people at home to deceive the police to be able to stay at home:

It tends to be with different cultures where they have coached, especially if child services have put them with a family member and they’ll have been told not to say anything. (FG2, PB)

Working with children from different domestic and cultural backgrounds, who most likely differ a great deal from the eager school-attending children who take part in
scientific research, gives the practitioner an edge to understanding how some children can be driven to deceive in serious situations.

6.4.3.3 Empowering the interviewee

Many of the participants commented that the usefulness of any interview technique would depend on the child they were interviewing. Participants stressed that it was very important to tailor their interviews to each individual child interviewee to get the most information from him/her:

Ultimately we need to safeguard the children that we’re interviewing, and that’s the most important thing. So if they want to tell [us] something, it’s got to be the way they want to tell us. (FG1, PB)

Summarised by one of the participants; “I don’t think there’s such a thing as one size fits all” (FG3, PB). Enabling the interviewee required much preparation prior to the interview to ascertain the best ways to strengthen the interviewee, but flexibility was also required during the interview to adapt to any unforeseen individual differences that may arise.

6.4.3.3.1 Strengthen the interviewee

Strengthening the interviewee referred to any preparation prior to the interview that the interviewer could undertake to tailor their interview techniques to get the best out of that particular interviewee. The primary means for this was by conducting Intermediary assessments. There was overall consensus that Intermediary assessments are a great source of information for assessing each child’s suitability for interview and how they respond to different questions:
When the intermediaries do the assessment […] we’re often present […] so you get a feeling as to how they respond to certain types of questions and, you know, the number of sort of key words you put in a question. (FG3, PC)

Participants felt that an Intermediary assessment was also beneficial for highlighting a child’s suggestibility, “I had one the other day and there was a picture of summer and she [the intermediary] kept saying to the child, “in the winter so this is winter” and the child was just agreeing” (FG1, PA), and for recommending how best to interview a child with learning disabilities; “you [have] got to assess the individual […] and depending upon the learning disability, and perhaps in consultation with an Intermediary as well, you might only do five minutes ‘cause their attention span is so short” (FG3, PD). This suggests that the previous comments regarding asking the ‘right’ questions may also be child-specific and that the interviewer requires external input from the Intermediary to shape what they consider to be the ‘right’ questions to that particular child interviewee.

It was also mentioned, however, that a delay whilst waiting for the right Intermediary (i.e. one that is specialised in assessing a particular child interviewee) can have a negative effect on a young child’s recall ability:

For little children, they [Intermediaries] can’t help them remember. So we will still have to interview them a lot later down the process and if they’re struggling to remember something then the Intermediary, although they can give us how we should be forming our questions, it can’t help them remember. (FG1, PB)
Thus, intermediary assessments facilitate best practice for eliciting as much information from that individual child as possible but these benefits begin to wane if the interviewer has to wait too long for the assessment to take place. Under these circumstances, it might even be justifiable to conduct an interview without performing an Intermediary assessment. Here the participants in the current study demonstrated some knowledge about the sensitivity of memory to time delays.

6.4.3.3.2 Flexibility

Despite preparing for the interview, participants discussed the need to be flexible and be able to adapt to any individual differences that aren’t covered in an Intermediary assessment. For example, the child’s background may make it difficult to predict how the child will behave during the interview:

The kids we’re going to get, whether they’re been abused or not, whether they’re truth- or lie-tellers, are going to be chaotic people who are in deprived backgrounds, don’t know where their next meal’s coming from, no discipline.

(FG2, PC)

A chaotic person at interview can be predicted as acting unpredictably so it’s up to interviewers to use their expertise and experience to react accordingly. The best advice from one participant was; “just roll with it” (FG3, PD), demonstrating that, to some extent, the interviewer should just go with how the interviewee wants to present themselves and their story.
A child’s personal disposition (i.e. “no self-esteem and lack of confidence” (FG3, PC)) may also mean adapting the flow of the interview to that person:

We might be going in and going out of the subject, you know, if you’re seeing that you’re upsetting them. (FG3, PA)

The interviewer can recognise the effect of the subject of the interview on the interviewee and be sensitive to its effect on their willingness to recall the target incident. However, as much as the interviewer can try and construct a safe environment to increase the likelihood of disclosure, sometimes this is not always possible:

I had one years ago and I did three attempts ‘cause they were just struggling to talk about it and they wouldn’t talk about the offence. Like you say they would talk about everything around it but it was just too horrific for them to talk about. So three times we tried and, in the end, we never got it. (FG1, PF)

The interviewer must, therefore, accept that being flexible to an interviewee also includes terminating an interview before sufficient evidence has been gathered which may mean that a case cannot be progressed.

6.4.3.4 End product

The final theme that emerged from participants’ discussions of the broad context of interviewing children was the importance of ensuring a good ‘end product’ that could be
played in court to ensure that children would be safeguarded and potentially achieve justice:

The whole idea is we’re trying to help that child get this to courtroom and get that offender convicted. If we do something wrong [...] they’ll throw the interview out and the case is lost and we haven’t protected that child. (FG1, PB)

In other words, interviewing child interviewees is a ‘high stakes’ business and if interviewers put a foot wrong, then there could be serious legal implications. This superordinate theme was divided into two sub-ordinate categories: preparing the evidence so that it is admissible in court and trying to pre-empt the cross-examination when that evidence is presented in court.

6.4.3.4.1 Preparing the evidence

Across all focus groups, comments regarding evidence were mainly about the pitfalls of child interviewing that may render evidence inadmissible in court. Participants noted that police officers must do what’s right for the children at the same time as doing what’s right for the court. They, therefore, need to be in sync, and not at odds, with all of the components of the judicial system. Interviews need to be conducted according to the guidelines; “we have to show that we did everything to the book” (FG1, PA) but, taken to the extreme, this resulted in some participant resentment towards the system:

We avoid transporting children to interviews and having conversations with them in the car because [...] you have to write down the whole conversation in the white
book so that they know. Obviously we’re police officers, we can’t be trusted. We have to write down to show what topics we’ve spoken about so that we’ve not been coaching them in the car. (FG3, PC)

Police officers regulated their behaviour to try and avoid the mistrust of court officials at trial to ensure that the children’s evidence was not considered as tainted. However, participants did express how unnatural this regulated behaviour felt when confronted with children in obvious distress:

They’ll [the defence lawyers] say things like […] “it’s not a fair trial”, you’ve got to be so careful […] and it’s awful when you’ve got a five-year-old or six-year-old that’s sitting there crying, especially being a female as well, you want to comfort that child. (FG1, PA)

Indeed, this lack of comforting also extended to not giving verbal encouragements and basically neutralising their behaviour completely; “we’re not even allowed to say, “well done, thank you” nothing. They sit in that interview and we are like a robot” (FG1, PD). These comments reflect the constraints that participants perceived the court system to impose upon their practice. There was also a sense of ‘walking on eggshells’, especially when participants did not agree with changes in procedural guidelines. One participant reflected on the contrast between previous and current practice in empathic responsivity:
Before if a victim used to cry, we’d be able to put our arms around the victim and we’d given them a bit of comfort. Now, it’s just about getting them a box of tissues, “are you okay to continue? Would you like a break? Would you like to go and see whoever is downstairs?” (FG 1, PA)

Interviewer behaviour aside, difficulties also arose from the politics of getting evidence to court. One participant described a case where a boy told a story that was only half supported by eyewitness evidence; other parts of the child’s account were not supported by other evidence. The police officer described that, in the end, the courts only looked to prosecute the mother for what was supported, ignoring all the other detail given by the child, about other crimes, that was not supported. The participant could not understand what the Crown Prosecution Service (CPS) used to make their credibility judgements:

CPS [...] what are their qualifications to make these decisions? ‘Cause that’s what they do, day in day out, is make decisions about whether a kid is telling the truth based on pretty much nothing I can establish. (FG2, PC)

To summarise, when police officers prepare evidence, they reported that they must follow many ‘rules’ but these ‘rules’ are more pre-emptive of what will happen in court and what the CPS will accept rather than ‘rules’ that are actually written down as official guidance. The fundamental issue is that police officers are information gatherers for the courts.
6.4.3.4.2 Cross-examination

Participants spoke about their experiences in court when they are cross-examined about how they elicited evidence from their child witnesses. This experience was resoundingly negative across all focus groups. In particular, participants expressed a lot of negativity regarding defence teams:

You’ve got your horrible defence barristers who will just throw anything. They haven’t got to prove anything. They’ve got that luxury where they can just stand there and make a suggestion. (FG3, PD)

They explained how defence teams are more likely to question their interviews rather than the actual content of the interviews themselves:

They look at procedure, not necessarily the content and product of the interview itself because invariably there is good material in that […] the last thing the defence team want to do is touch the evidence because it’s usually pretty compelling. (FG2, PA)

Participants’ understanding of the court dynamics explains their caution when preparing their evidence for trial. If they are the most likely to come under fire when examining child witness’ testimonies, they need to be able to demonstrate that they made defensible decisions throughout the investigative process to reduce any heat that they
might receive from the defence team. Participants talked about trying to avoid criticisms, which would later occur in court, when they were interviewing children. Anything that could be seen to be leading child interviewees was avoided:

We have to be totally transparent and impartial and we could get criticised that we’re goading or that we’re influencing or that we’re preparing them for the interview. (FG1, PA)

In addition, participants explained how they cut the length of the recorded material down, as this was another constant criticism from the courts, “we get criticised enough for how long our interviews are” (FG3, PC). However, this isn’t always possible due to the nature of the incidents that the interviewee is talking about:

We always argue you know if somebody has come to talk to you about a sequence of events that have been happening to them over a period of years, you can’t get that person to give you all that information in a half an hour interview, ‘cause somebody that’s talking about one event would come and take twenty, twenty-five minutes to tell you that information. (FG3, PC)

In other words, it was perceived by participants that the courts were not sympathetic to the diversity of the children that the participants interviewed; they didn’t understand that one size (or duration, in this case) did not fit all. Overall, it was very clear
that the cross-examination process in courts means that police officers have to conform and adjust to the standards of the courts and not the other way round.

6.5 Discussion

This qualitative study aimed to gain practitioner feedback regarding new cognitive lie detection techniques that could be used at interview with child interviewees, and to understand the child interviewing process in which these techniques would be implemented.

According to the participants in the current study, using a child’s own model statement stood out as the most viable option for use in real police investigations. It was the only interview technique that received positive feedback alongside suggestions for adaptation for use in an applied setting. At the time of the focus groups, the results of Chapter 5 were not available for presentation so these comments were based on the predicted outcomes of using a child’s own model statement as a tool for better discriminating between truth-tellers and lie-tellers. Practitioners may have felt differently about using this interview technique in practice if they knew that ‘own model statements’ did not produce the same effect as using another child’s model statement. However, their comments demonstrated that this technique was the most compatible with real-world constraints, and so future research should look to resolve the methodological problems of Chapter 5 to see if own model statements can be tested in a different capacity (e.g. using narrative elaboration techniques, Camparo, Wagner & Saywitz, 2001; Saywitz & Snyder, 1996). Instructing interviewees to maintain gaze came in second place, receiving some suggestions for use but still no mentioned benefits with child witnesses. Finally, practitioners identified only disadvantages to using another child’s model statement, writing it off primarily due to potential issues with tainting children’s accounts. This last finding is slightly perplexing as practitioners did talk about providing an example of a
detailed description of an object themselves to encourage more detailed reports. The subtle
nuance between how that does not constitute tainting but another child’s model statement
might taint requires further investigation. Overall, representation for each of themes within
each technique (see Figure 6.1) ranged from 13% to 69% of participants. Participants were
the least in agreement regarding the suggestions for use for gaze maintenance and most in
agreement for the disadvantages of playing another child’s model statement.

Discussion in the focus groups of the broader child interviewing process
highlighted many constraints that police officers face from both the child interviewees
themselves and the court process. On the one hand, interviewers must adapt to the
individual differences of child interviewees and their circumstances. On the other hand,
interviewers must always be careful to avoid putting children’s evidence at risk in terms of
its admissibility for the courts. Thus, practitioners find themselves stuck in the middle
trying to work out the best course of action for each individual case. This would explain
their self-reported ability to consider the effects of different witness characteristics, such as
age, background and trauma, on children’s verbal and nonverbal behaviour when being
interviewed. This sensitivity is, of course, of great value to researchers who may not be
aware of certain effects. The four themes that emerged – the 4 E’s – already provide a
simple and straightforward framework against which to evaluate the viability of interview
techniques with children. Any interview technique should (1) elicit as much information
from children as possible, (2) allow practitioners to accurately evaluate the credibility of
that information, (3) empower child interviewees and be flexible to their specific needs,
and (4) create a good end product that will stand up in court with little criticism. Although
the first two themes can be tested empirically, the second two themes would require the
advice and recommendation of someone in the field who understands the nuanced
practicalities of working with child interviewees. We should, therefore, continue to
empower practitioners by seeking and valuing their feedback regarding our research at the early stages of research design.

The present study is not generalisable with only three focus groups across two UK police constabularies. This means that the themes and comments made may only reflect the views of practitioners within these police forces. That said, it is of interest to note that there was a great deal of agreement between practitioners based in the North of England (Focus group 2) and practitioners based in the South of England (Focus groups 1 and 3), which suggests that these views might be more widespread. Another issue is that the participants were only presented with research issues from this PhD thesis and were not made aware of other uses of cognitive lie detection techniques with children (i.e. unanticipated questions - Liu et al., 2010; Roos af Hjelmsäter et al., 2014). Thus, using a child’s own model statement was the favoured technique out of those shown but might not have been considered the most viable if other options were given. This study provides a pilot study that shows that the academic-practitioner conversation is worth investing in to inform future research.

Indeed, future research should investigate some of the ‘intuitive’ techniques that the police already use to determine the veracity of a child’s statement. For example, it would be of interest to know the weighting of a child’s statement in the bigger picture of the investigation when the case goes to trial. If the child’s statement is in conflict with other key statements, then it would be of note to understand who layperson jurors are most likely to believe. Further examination of the description of an object as an example of detail (i.e. an interviewer-generated model statement) by the interviewer, prior to interview, should also be conducted. Studies should look to explore whether this helps or hinders lie detection, and whether children of all ages can understand this abstract concept.
In conclusion, practitioners are an extremely valuable source of information who are arguably overlooked during the design of research programmes by academics. They can provide not only insightful feedback on current research into new interview techniques, but also an eye-opening portrayal of the child interviewing process as a whole. It is likely that joint researcher-practitioner ventures are those most likely to succeed in developing techniques that are practical in the field as well as being supported by empirical evidence.
6.6 References


Chapter 7: General Discussion
7.1 General Discussion Outline

The general discussion will first provide a summary of the main findings of this thesis, drawing attention to the verbal markers associated with the veracity of child witness statements that were elicited when cognitive lie detection techniques were used (see Section 7.2). Second, the theoretical implications of these new child deception research findings are discussed with a focus on the underlying cognitive processes targeted by cognitive lie detection (see Section 7.3). Third, the practical implications of the doctoral research are discussed, highlighting the transition from lab to field (see Section 7.4). Fourth, the methodological issues are considered (see Section 7.5) followed by suggestions for further research (see Section 7.6). Finally, this doctoral thesis is concluded (see Section 7.7).

7.2 Overview of Main Findings

Child deception research has highlighted adults’ difficulties in determining the veracity of children’s statements (Talwar & Crossman 2012; Gongola, Scurich & Quas, 2016) as well as the ineffectiveness of current passive techniques used to deter children from lying in investigative interviews (Bala, Lee, Lindsay & Talwar, 2010; Lyon, Malloy, Quas & Talwar, 2008; Talwar, Lee, Bala & Lindsay, 2002). The main aim of this thesis was to explore the use of cognitive lie detection techniques with child witnesses. More specifically, this doctoral research aimed to go beyond previous studies that had used only unanticipated questioning with younger populations (Liu et al., 2010; Roos af Hjelmsäter, Öhman, Granhag & Vrij, 2014) to examine the effectiveness of (a) imposing cognitive load through gaze maintenance, (b) encouraging interviewees to say more using another child’s model statement, and (c) encouraging interviewees to say more using children’s own model statements. The final purpose of this research programme was to investigate the
perceptions of police practitioners, working in the child-interviewing field, to survey their opinions with regard to applying the lab findings to the real world.

Two experiments examined whether imposing cognitive load through gaze maintenance could facilitate child deception detection (these were reported in Chapter 2). In the first experiment, child interviewees were instructed to either maintain gaze (towards an interviewer’s face or a teddy bear’s face), or not, whilst telling the truth or lying about taking part in a recent school event. Overall, truth-tellers provided significantly more details than lie-tellers, but only when they had been given a gaze instruction to look at either the interviewer’s face or the teddy bear’s face. This result is in line with adult cognitive lie detection literature, which has shown that imposing cognitive load can create significant verbal differences between true and false reports (Vrij et al., 2008; Vrij, Mann, Leal & Fisher, 2010). Furthermore, an increase in detail was only found between gaze conditions for truth-tellers, whereas the amount of details that lie-tellers provided was the same, regardless of where they looked. In a second experiment, adult evaluators were asked to determine the credibility of the children’s interview clips. Only evaluators who judged children that had been instructed to look at the interviewer’s face were able to discriminate between truth-tellers and lie-tellers. This was also the only condition in which detection accuracy was significantly above chance (58%). The lack of a beneficial effect when evaluating children that had been instructed to look at the teddy bear’s face could be a consequence of the true and false reports not being sufficiently different in detail. That is, what was a statistically significant difference in quantity of detail was not perceived by human observers.

The effect of hearing another child’s model statement (AMS; i.e. an example of a detailed recall of an unrelated event) on the elicitation of verbal differences in the reports of children who had, or had not, witnessed a mock mobile phone theft was investigated in
Chapter 3. The interview protocol involved eliciting a baseline free recall attempt, followed by the presentation of AMS and then a second free recall attempt. Building on the coding scheme of Chapter 2, Chapter 3 went beyond interviewee productivity (i.e. quantity of detail, response length) to also code forensic usefulness (i.e. crime-relevance of the details) and between statement-consistency. It also investigated the memory accuracy of the truthful reports. Overall, both truth-tellers and lie-tellers provided significantly more details after hearing AMS, similar to the adult insurance claimants in Leal, Vrij, Warmelink, Vernham and Fisher’s (2015) study. When the two free recall attempts were compared, both truth-tellers and lie-tellers added the same amount of new information (i.e. commission) to their reports, as expected based on the Pinocchio hypothesis (Bruck, Ceci & Hembrooke, 2002) and the repeat versus reconstruct hypothesis (Granhag, Strömwall & Jonsson, 2003). In addition, truth-tellers repeated a significantly higher proportion of crime-relevant information (i.e. information pertaining to the criminal event according to eyewitness literature) compared to lie-tellers. Nevertheless, truth-tellers’ reports also became significantly less accurate following AMS. Although this study used children’s baseline free recall attempts as a control, the lack of a control condition where no model statement was played in between recall attempts made it impossible to conclude that these findings were not simply the artefact of repeated questioning, which can result in longer responses (McCauley & Fisher, 1995).

Chapter 4, building on Chapter 3, presented a follow-up lie detection experiment where adult evaluators assessed the credibility of pre- and post-AMS interview clips. Using AMS did not facilitate the detection of children’s truths and lies. A lens modelling approach, which has been used in previous lie detection research (e.g. Hartwig & Bond, 2011), was applied to tease apart the relationship between children’s verbal behaviours and adults’ perceptual ratings when deciding if children were being truthful or not. This
revealed that adults strongly rely on quantity of detail as a key indicator of children’s veracity; a believed cue that is empirically and consistently supported (Vrij, 2005). Thus, the poor detection accuracy in Chapter 4 could be the consequence of both truth-tellers and lie-tellers increasing their interviewee productivity following AMS, meaning no differences in level of detail existed between the two groups. In other words, in Chapter 3, quantity of detail was not a reliable indicator of deceit but this was not explained to participants prior to them making their judgments. It is suggested that evaluators need to understand the anticipated effects of the model statement interview technique, in terms of its elicitation of differences between truthful and fabricated accounts, before detection ability will improve.

Chapter 5, again following on from Chapter 3, investigated the use of another child’s model statement (AMS) compared to a child’s own model statement (OMS). This study also included a ‘no model statement’ condition (NMS) to ensure that the results were not an artefact of repeated questioning. Furthermore, this study reduced the effect of individual differences by requiring children to tell both a true and false report at interview. First, we found that only children who heard AMS increased the amount of detail in their second recall attempt of both the true and false activities. This is in line with the findings of Chapter 3 and previous studies with adults (Ewens et al., 2016; Leal et al., 2015). This increase in productivity for children who heard AMS (and not for the OMS or NMS conditions) resulted in the children in the AMS condition providing the most new information (i.e. most commissions). In addition, the new information that children who heard AMS included in their true reports did not affect the overall accuracy of their second recall attempts. Indeed, they remained as accurate as the children in the other interview conditions, even though they said more; a finding that has also been found when eliciting longer child statements using the Revised Cognitive Interview (McCauley & Fisher, 1995).
None of the predictions regarding the use of OMS were supported. The increase in interviewee productivity could be the consequence of shorter practice interviews that were conducted with a long time delay. Overall, the study reported in Chapter 5 demonstrated that a change in interviewee productivity for true reports was due to listening to another child’s model statement and not due to repeatedly asking for two free recalls. In addition, it also provided evidence that a decrease in overall accuracy highlighted in the earlier study, reported in Chapter 3, may have been the consequence of children poorly encoding the ‘to-be-remembered’ event. In the Chapter 5 study, where children were informed that they would be questioned about the true activity, their accuracy was not affected. Thus, they would have been aware that they should encode the event well, putting them in a ‘satisficing knowledge’ state (Ackerman & Goldsmith, 2008), whereby they would not need to sacrifice the accuracy of the information they provided when they became more informative.

Finally, Chapter 6 reported on a qualitative study that aimed to increase our understanding of the transition process from laboratory to field of the cognitive lie detection techniques investigated in Chapters 2 to 5. In focus groups, police officers who regularly interview children were presented with the interview techniques that had been investigated in this thesis and were asked to provide practitioner feedback on their relevance and feasibility for implementation in current interviewing practice. Using a child’s own model statement came out as the most viable option, although this was selected on the basis of the theoretical findings rather than the actual findings of Chapter 5 (which were not available at the time that the field investigation was conducted). However, participants stated that even this interview technique would require some adaptation, in terms of when it should be used, how long it should last and who it should be
used with, before it would be eligible for use in real police investigations involving child witnesses.

As well as evaluating the interview techniques, the police officers provided an insightful description of the child interviewing process and the journey from disclosure to court. From this discussion, four major themes emerged (the 4E’s) that could be used as a basic framework for future child deception research in this area. For an interview technique to be viable it needs to **elicit** as much information as possible, allow the practitioner to **evaluate** the credibility of that information, be flexible so that it can adapt to **empower** each interviewee, and, finally, produce a sound **end product** that will stand up in court. The field investigation highlighted the difficult position in which child interviewers find themselves: Constrained by the limits of each child interviewee but also constrained by the standards of the court. It also promoted the importance of empowering practitioners to bring to academics’ attention subtle nuances in the child interviewing process that would only occur to a police officer who has practical experience.

To summarise, cognitive lie detection can have the desired effect of exaggerating behavioural differences and facilitating the detection for both truthful and deceitful child witnesses. However, this is not true of all techniques. Imposing cognitive load through gaze maintenance achieved both goals successfully. Yet, encouraging interviewees to say more using another child’s model statement in an eyewitness context elicited only verbal differences on a very specific level (i.e. percentage of crime-relevant repetitions) and was, therefore, not able to improve detection accuracy that seems to rely on more easily observed differences. On the other hand, using another’s child model statement in a non-forensic context did elicit more general differences in terms of quantity of detail and consistency.
7.3 Theoretical Implications

7.3.1 Cognitive lie detection theory

Cognitive lie detection is based on the premise that lie-telling is more cognitively complex and, thus, more cognitively demanding than truth-telling (Zuckerman, DePaulo & Rosenthal, 1981). Each of the cognitive lie detection techniques put forward by Vrij (2015) aims to explore this difference in experiencing cognitive load by taxing lie-tellers’ cognitive resources even further through the manipulation of certain cognitive mechanisms. In Chapter 2, additional cognitive load was imposed on child lie-tellers through the request to perform a dual task. In Chapters 3 and 5, child lie-tellers’ cognitive resources were challenged through the use of social proof (Cialdini, 1993). That is, the expectations of a detailed account were presented for comparison (i.e. by playing a model statement) thereby encouraging child lie-tellers to say more and put further strain on their already under-developed verbal working memory (Alloway, McCallum, Alloway & Hoicka, 2015).

7.3.1.1 Imposing cognitive load through dual task performance

Imposing cognitive load aims to make the interview setting more difficult by imposing an additional request that should affect lie-tellers, who have fewer cognitive resources available, more so than truth-tellers. Vrij et al. (2008) proposed two mechanisms of attention that might account for the debilitating effect of imposing cognitive load, on deceitful interviewees, through a secondary task. First, if cognitive resources originate from the same limited pool, then lie-tellers will have fewer resources ‘leftover’ to attend to the secondary task compared to truth-tellers (Knowles, 1963). If the second task is, therefore, also cognitively demanding, lie-tellers will be pushed to the limits of their attentional capacity, and thus experience more cognitive overload compared to truth-tellers. To cope with both tasks, a second mechanism of attention management must be
triggered whereby lie-tellers can flexibly allocate their attention between the two tasks in accordance with the demands of each task (Kahneman, 1973). This, however, could cause dual-task interference between the tasks, as diverting attentional resources away from the primary task to the secondary task might increase task performance for the latter and decrease task performance for the former. Prior to this doctoral research, these theoretical assumptions had yet to be directly tested in any study that investigated imposing cognitive load. In Chapter 2, we analysed this anticipated negative relationship between the primary task of lie-telling (measured by number of details provided) and the secondary task of maintaining gaze (measured by percentage of time spent gazing towards the face stimulus). Our findings found no dual-task interference for either secondary task. In fact, we found a small, non-significant positive association between lie-telling and maintaining gaze with the interviewer’s face, suggesting a beneficial effect of performing both of these tasks simultaneously. Although our findings did not support the theory, this could be explained by the secondary task (visual) not drawing attention from the same pool of resources as the primary task (verbal). Multiple resource theory (Wickens, 2002) suggests that cognitive resources that come from structurally dissimilar pools (i.e. verbal versus visual) are likely to interfere less. Testing this theory with reverse order recall as the additional task might produce different results as recounting the story backwards is also a verbal task (Saykaly, Crossman, Morris & Talwar, 2016). Furthermore, due to the design of Experiment 1, reported in Chapter 2, we were unable to account for children’s ability to perform each of the tasks on their own allowing us to then control for this when they performed both tasks together. This method of understanding dual-task performance has been used in previous attention research with children and has helped to dispel some previous claims about children’s ability to manage their attention (Irwrin-Chase & Burns, 2000). Implementing a similar design in future imposing cognitive load research would help to dig deeper into
the cognitive processes involved when performing a secondary task at interview alongside lie-telling.

7.3.1.2 Encouraging interviewees to say more through social proof

Using a model statement as social proof of what a truth-tellers’ detailed account sounds like sets the standard for what level of detail is required at interview and encourages both truthful and deceitful interviewees to compare their expectations to this standard, ultimately resulting in them saying more (Vrij, 2015; Vrij, Leal, Mann, Vernham & Brankaert, 2015). Although these theoretical underpinnings were not directly tested in the study presented in Chapter 3, increases in interviewee productivity for both true and false reports were in line with this theory. Chapter 5 directly investigated the relationship between social comparison and increased interviewee productivity for true reports by looking at the underlying social-cognitive processes at work when truth-tellers hear another’s or their own model statement. The findings were contrary to a theory of social proof (Cialdini, 1993) because there was no relationship between the type of social comparison made by the children when comparing themselves to the model statement they heard (i.e. upward, downward or neutral) and their change in interviewee productivity. However, this may have been due to unreliable self-reports. Children in both the AMS and OMS conditions did report that they compared themselves to the model statement more often than not. Yet, the types of comparisons they made were not an accurate reflection of their performance. Indeed, on average, both children in the AMS condition and in the OMS condition, performed significantly worse than the model statements they heard. Nonetheless, neither group was significantly skewed towards making accurate downward comparisons. Social proof theory can, therefore, not be ruled out as the underlying mechanism for the effect of using AMS. Indeed, our findings would suggest that future
research should use a different method for measuring social comparison that might be more accurate than self-reporting.

7.4 Practical Implications

As mentioned in the Introduction of this thesis (Chapter 1), techniques are required for use during police interviews that (1) have been empirically shown to facilitate child deception detection, and (2) will be used by practitioners. In addition, any lie detection technique should (3) not have any adverse effects on child truth-tellers’ ability to give a full and accurate account (Talwar & Crossman, 2012). It was important to ensure that the interview techniques investigated in this thesis could not be criticised in court for having a negative effect on the account of truthful child interviewees. As noted in Chapter 6, the techniques should facilitate the elicitation of information from child truth-tellers. Moreover, as noted in Chapter 3, eliciting more information from child interviewees should not decrease the accuracy of that information.

First, imposing cognitive load by instructing interviewees to look at the interviewer’s face was shown to elicit significant verbal differences in general level of detail provided in true and false reports, which then facilitated the detection of children’s deceit. Although there were no advantages highlighted in the focus group study in Chapter 6 for using gaze maintenance, practitioners did provide some suggestions in relation to this technique; namely, its use in certain cases not involving serious sexual trauma and with older child witnesses. Although practitioners thought that it may not elicit as much information from child witnesses as letting them look wherever they choose, this is in contrast to the findings of Chapter 2 that found truth-tellers instructed to look at the interviewer’s face provided the most details. Nonetheless, the finding of Chapter 2 was limited by the use of a non-staged event that did not allow us to code the accuracy of the
children’s statements. Staged events were, therefore, used in the subsequent chapters to investigate the possibility of negative effects of interventions on truth-tellers’ accuracy. Thus, the practical use of gaze maintenance with child interviewees is suggested with caution in some cases. The technique should be subject to verification that the longer statements elicited, when eye contact is required, do not reduce the accuracy of children’s truthful reports.

Second, using another child’s model statement (AMS) had mixed practical significance. On the one hand, it did elicit significant verbal differences between children’s true and false reports on both a specific level (i.e. percentage of crime-relevant repetitions in Chapter 3). On the other hand, these specific differences did not translate into better detection accuracy in Chapter 4. From the practitioners’ perspective, using AMS was seen to be disadvantageous because introducing an external source of information would taint child witnesses’ subsequent reporting; a seemingly valid concern. That said, the police officers did report using a similar technique whereby they, as interviewers, provide an example of giving a lot of detail (for example, by giving a lot of detail about the description of a mug), which could also taint children’s accounts. This subtle nuance suggests that AMS could be adapted using this already existing technique to form a hybrid that increases productivity and would be accepted by police officers for use in their investigations. Regarding the effect of AMS on truth-tellers, Chapters 3 and 5 found that children who heard AMS provided subsequent statements that were significantly more productive than their control or baseline counterparts. However, the effects of this increase in interviewee productivity on overall accuracy were mixed. Whereas Chapter 3 showed a significant decrease in the overall accuracy of true reports, Chapter 5 did not show any significant change in accuracy across recall attempts. To understand how accuracy would be affected by AMS in police investigations, it could be
argued that the experimental procedure of Chapter 3 is most relevant as, like most eyewitness events, the children were interviewed about something they had no forewarning would turn into the subject of an investigation. Under these conditions, where they may have viewed/heard the to-be-remembered event under sub-optimal encoding conditions, interference is likely to occur and very poor memory stimuli are likely to result (Fernandes & Moscovitch, 2000). In this ‘unsatisficing knowledge state’, children might be inclined to sacrifice the accuracy of the information they provide in an attempt to be more informative (Ackerman & Goldsmith, 2008). Thus, based on the results of the ‘model statement’ studies in this thesis, AMS should be recommended for use with interviewees who are aware of the importance of remembering the to-be-remembered event and, thus, encode the event well.

Third, transforming children’s practice narratives into ‘own model statements’ did not have the desired effect of eliciting any significant differences between true and false reports: A lack of exaggerated differences, therefore, make it an unlikely facilitator of child deception detection (Hartwig & Bond, 2011). From a theoretical standpoint, practitioners considered it the most viable option out of those presented. However, it would still require much adaptation before it could be used as part of police investigations with child witnesses. Nevertheless, the noted advantages of using OMS will help to refine this interview technique or, at the very least, design other interview techniques to boast the same benefits. In terms of adverse effects, it did not elicit more information from child participants and it, therefore, did not affect the accuracy of their truthful accounts. As it stands the practical significance of OMS remains low, yet, with adaptation and a stronger emphasis on the practice interview aspect of this technique, it is possible that future research could demonstrate its potential to facilitate the detection of children’s deceit by helping truth-tellers. Indeed, by taking practice interviews back to their roots and using
narrative elaboration techniques to focus more clearly on the specific types of detail that are required during an investigative interview (i.e. people, objects, setting, actions, Saywitz & Snyder, 1996), the effect of the own model statement on true reports could be strengthened.

To conclude, instructing child interviewees to maintain gaze ticks the ‘empirically-supported’ box, using AMS ticks the ‘anti-adverse-effects’ box (but is dependent on the memory encoding for that event), and using OMS ticks the ‘practitioner-friendly’ box!

7.5 Methodological Considerations

7.5.1 Limited age range of children

All the children that participated in this doctoral research were aged between 8 and 11 years. This age group was chosen predominantly because it has been shown that, under the age of 8 years, children perform poorly when telling and maintaining lies (Talwar & Lee, 2002), and so we wanted to test the cognitive lie detection techniques with a slightly more challenging child population. Furthermore, for practical reasons, we decided to stay within the same school grouping (i.e. up until the end of primary/junior school in the UK school system). Nevertheless, reducing our samples to this specific age bracket makes it hard to speculate on the wider effects of age and cognitive development on the effectiveness of the investigated techniques. Looking across the lifespan (6 to 77 years old), Debey, Schryver, Logan, Suchotzki and Verschuere (2015) found that participants’ ability to inhibit their behaviour followed the same pattern as their lying frequency behaviour. This suggests that the development of cognitive mechanisms associated with ‘good’ lie-telling, such as inhibitory control, promotes more engagement in being deceitful (Talwar & Lee, 2008). As our participants find themselves in mid-childhood, they may have developed some of the cognitive skills necessary for successful lie-telling but they
may not have mastered the art. As children’s cognitive ability is closely related to their ability to tell lies (Talwar & Crossman, 2011), it would be important for future research to test the interview techniques investigated in this thesis with younger children and older children to understand fully the impact of age and cognitive development on the effectiveness of the interview techniques.

7.5.2 Absence of individual differences

Another important limitation is the lack of individual differences investigated in the studies in this doctoral thesis. Although we did control for age and gender of the children, we did not look beyond demographic variables to explore cognitive and social ability as well as personality and deceptive strategies. Random allocation of children to experimental conditions should help to control for differences in social and cognitive development that could influence deceptive ability (Talwar & Crossman, 2011; Talwar & Lee, 2008). However, it would have been of interest to know whether theory of mind or executive functioning had a mediating influence on the effects of each interview technique. Furthermore, the effects of a particular interview technique could depend on how that technique is understood by child lie-tellers and what strategies they put in place to negate the negative impact said technique is supposed to have on their lie-telling ability. Children have reported using a variety of verbal and nonverbal strategies to be successful lie-tellers (Strömwall, Granhag & Landström, 2007). It would be of interest to know whether our child participants employed any of these strategies, and whether they significantly inhibited the effects of the interview techniques. Finally, it is important to know what motivates children to lie and lie well when asked to in an experiment. Anecdotally, one of the children reported that they would have preferred to have told the lie if they had chosen to do so on their own accord. Thus, deliberately asking children to lie and allowing them to choose to lie might result in different levels of motivation and perhaps performance.
7.5.3 Irrelevance of context to target field

The majority of our findings are also limited by the non-forensic events about which the children were interviewed. Previous research on child deception detection has already shown the impact of context on evaluators’ perceptions of veracity. For example, when they hear children talking about being touched (in the context of a game) by an experimenter, adults are more inclined to judge children as deceitful (Edelstein, Luten, Ekman & Goodman, 2006), whereas when they hear children talking about a coached, non-forensic event, they are more inclined to judge them as truthful (Talwar et al., 2006). This differentiation in the criminal relevance of context was also found between Chapters 3 (forensic) and 5 (non-forensic) of this thesis where it affected the memory accuracy for the event. To recap, reflecting a typical eyewitness event (i.e. a mobile phone theft in Chapter 3), where the eyewitness was not informed that the event would occur, led to an increase in inaccurate information when the eyewitnesses were encouraged to be more informative. Alternatively in Chapter 5, providing the eyewitnesses with prior warning that the event would take place and should be remembered for later recall resulted in accuracy being maintained. Accounting for more factors relating to ‘being an eyewitness’, such as not knowing that a to-be-remembered event will take place, and, thus, the quality of encoding of unexpected events would have increased the practical significance of the findings.

7.5.4 Sole focus on verbal behaviours

The effect of using the cognitive lie detection techniques in this thesis on exaggerating behavioural differences between child truth-tellers and child lie-tellers has solely focused on differences in their verbal behaviours. This is due to children’s verbal behaviours being more indicative of their deceit than their nonverbal behaviours (Akehurst, Manton & Quandte, 2011; Talwar & Lee, 2002). Nevertheless, when it comes to child deception detection, a combination of verbal and nonverbal cues (e.g. total CBCA + limb
movements + cognitive operations + speech disturbances for 5-6 year-olds) has been shown to be the most accurate at discriminating between their true and false reporting (Vrij, Akehust, Soukara & Bull, 2004). This was highlighted in the lens model reported in Chapter 3 that was unable to account for all of the cues that evaluators used when making their final credibility judgments. Although children as vulnerable witnesses are entitled to special measures in court, such as video-recorded evidence-in-chief, live video links (Ministry of Defence, 2011), the jurors and judge who will be determining the credibility of their account will still be able to see how the children behave when giving evidence and under cross-examination. In Chapter 2, manipulating the presentation of the children’s statements (audio only versus visual and audio) resulted in a truth bias when evaluators only listened to the children’s testimonies (i.e. when they were not able to see the children), whereas they showed no response bias when they saw and heard the video clips. This latter finding suggests that children’s nonverbal behaviour did have an impact on evaluators’ credibility judgments. We cannot, therefore, rule out that the cognitive lie detection techniques had an effect on our child interviewees’ nonverbal behaviour and that this effect did not subsequently influence adults’ credibility judgments.

7.6 Overview of Future Research

This doctoral thesis has scratched the surface of using cognitive lie detection techniques with children. In this chapter, as well as across the Discussion sections of Chapters 2, 3, 4 and 5, suggestions have been made regarding future research. To summarise, the main ideas proposed throughout this thesis include: (1) Measuring how the mechanisms manipulated by each technique are affected by developmental differences in cognitive skills associated with lie maintenance (e.g. verbal working memory, inhibitory control, and general executive functioning); (2) Testing and clarifying the underlying
social and cognitive theoretical assumptions of cognitive lie detection more directly (i.e.
dual-task interference, social proof); (3) Coding for nonverbal cues to understand all of the
behaviours that evaluators are perceiving when making their credibility judgements and to
identify pitfalls in their decision-making; (4) Informing evaluators about the mechanisms
behind the interview techniques to facilitate their understanding and detection accuracy;
(5) Continuing to examine the effects of cognitive lie detection on truth-tellers’ reports in
terms of productivity, relevance and accuracy of information; and (6) Involving
practitioners in the review process to ensure that all aspects of the child interviewing
process are taken into account when redesigning and adapting the interview techniques for
laboratory testing.

7.7 Conclusion

The aim of this thesis was to apply two cognitive lie detection techniques
(imposing cognitive load and encouraging interviewees to say more) to child witnesses to
facilitate the detection of their false reports. This thesis demonstrated that cognitive lie
detection is a promising method that can create significant verbal differences between child
truth-tellers and child lie-tellers that can, sometimes, improve detection accuracy.
However, this success was largely due to the first technique of maintaining gaze, with the
subsequent two techniques of using a model statement receiving less support. This
research suggests that the development of cognitive skills for lie maintenance, and the
effects of cognitive lie detection on those skills, need to be better understood if the
techniques are to work effectively. It also highlights that a practical knowledge of the
child interviewing process is necessary to transfer the findings of this doctoral research to
the real world.
7.8 References

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting--


Ewens, S., Vrij, A., Leal, S., Mann, S., Jo, E., Shaboltas, A., Ivanova, M., Granskaya, J, & Houston, K. (2016). Using the model statement to elicit information and cues to deceit from native speakers, non-native speakers and those talking through an interpreter. *Applied Cognitive Psychology*. Advance online publication. doi: 10.1002/acp.3270


Appendices
Appendix A: Favourable Ethical Opinion (Chapter 2)

STAFF AND PhD STUDENT PSYCHOLOGY DEPARTMENT RESEARCH ETHICS FORM 13/14

NAME: Or Lucy Atkehurst, Dr Amy Leach & Hannah Lawrence
NAME OF SUPERVISOR (if applicable) N/A
☐PhD research
✓ Staff research

Psychologists and students of psychology shall act in accordance with the ethical principles outlined in The British Psychological Society Code of Human Research Ethics (2010).

Please note: if you require approval from any of the following bodies you will need to go through the IRAS system. See https://www.myresearchproject.org.uk

- Ministry of Justice
- NHS / HSE NHSB office
- NHSE/HHSE MSC Research Ethics Committees
- National Information Governance Board (NHSB)
- National Offender Management Service (NOMS)
- Social Care Research Ethics Committee

Once you have gained a favourable outcome for your research, submit it to the departmental administrative office on the 1st floor of King Henry Building – no further action required.

Similarly, research involving animals must be assessed by the University Animal Ethics Committee to confirm that it is meeting Home Office Guidelines (see Moodle for form), just send a copy to the departmental administrative office on the 1st floor of King Henry Building once approved, no further action required.

Otherwise, complete the following procedure to undergo Departmental ethical review:

1. Complete all the questions in Part A, B and C below and save as a WORD document. As part of this process you will also be required to draft a short research proposal, recruitment documents, informed consent forms and debriefing forms, which again should all be saved electronically IN ONE WORD DOCUMENT PLEASE!! There is a fact sheet on Moodle that will help you along with deadlines for submission to the departmental ethics committee (DEC). Finally, complete the Science Faculty Ethics Committee (SFEC) submission document – copies available on Moodle with guidance inside the document.
   - PhD students should make an appointment with their supervisors to go through their responses and associated forms (usually good to email these to your supervisor in advance). Your supervisor will give feedback on your draft submission and associated documents which you must address to their satisfaction before submission.

2. It is University policy that all staff and postgraduate research must go for full review even if it has no substantive ethical issues. It is Science Faculty policy that all staff and PhD student ethical submissions should have a presence on the SFEC Moodle site, even if they are ‘light-touch’ (i.e. dealt with in house by your departmental committee).

Please complete before submission to the administrative office:

Once you have completed Part B:
XX Declaration A (substantive ethical issues) Declaration B (no substantive ethical issues)
Date: December 19th, 2013
To: Amy Leach (PI), Lucy Akeshurst (Co-PI, University of Portsmouth), and Hannah Lawrence (Co-PI, University of Portsmouth)
From: Bill Goodman, REB Chair
REB File #: 13-066
Project Title: Judging credibility in children: Helping adults get it right!
DECISION: APPROVED
START DATE: December 19th, 2013 EXPIRY: December 19th, 2014

The University of Ontario Institute of Technology Research Ethics Board (REB) has reviewed and approved the above research proposal. This application has been reviewed to ensure compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2) and the UOIT Research Ethics Policy and Procedures.

Please note that the (REB) requires that you adhere to the protocol as last reviewed and approved by the REB.

Always quote your REB file number on all future correspondence.

Please familiarize yourself with the following forms as they may become of use to you.

- Change Request Form: Any changes or modifications (e.g., adding a Co-PI or a change in methodology) must be approved by the REB through the completion of a change request form before implementation.
- Adverse or Unexpected Events Form: Events must be reported to the REB within 24 hours after the events occurred with an indication of how these events affect (in the view of the Principal Investigator) the safety of the participants and the continuation of the protocol. (i.e. unanticipated or un-natural physical, social or psychological harm to a participant)
- Research Project Completion Form: must be completed when the research study has completed.
- Renewal Request Form: Any project that exceeds the original approval period must receive approval by the REB through the completion of a Renewal Request Form before the expiry date has passed.

All Forms can be found at http://research.uoit.ca/faculty/policies-procedures/forms.php.

<table>
<thead>
<tr>
<th>REB Chair</th>
<th>Ethics and Compliance Officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Bill Goodman, PBIT</td>
<td><a href="mailto:uncompliance@uoit.ca">uncompliance@uoit.ca</a></td>
</tr>
<tr>
<td><a href="mailto:bill.goodman@uoit.ca">bill.goodman@uoit.ca</a></td>
<td></td>
</tr>
</tbody>
</table>

University of Ontario, Institute of Technology
2000 Simcoe Street North, Oshawa ON, L1H 7K4
PHON: (905) 721-8600, ext. 3693
Science Faculty Ethics Committee

Protocol Title: SFEC 2014-078, How well can you judge the credibility of children?
Date received PI response from 1st review letter: 14/11/14
Date Reviewed: 17/11/14

FAVOURABLE OPINION

Dear Ms Lawrence,

Thank you for your submission for ethical review. Having completed their review, members of the Science Faculty Ethics Committee have reached a Favourable opinion of your proposed research.

Please notify the committee of any substantial amendments to the proposed procedures, send an annual report to the committee regarding study progress and a final study report once the study has concluded. Please send these to sci_fac@port.ac.uk. Thank you and the committee wishes you well with your study.

Dr Simon Kalsbe - Vice Chair SFEC

CC:-
Heidi Stowyer - Faculty Administrator

If you would like to offer any feedback on the Science Faculty Ethics Committee process please email sci_fac@port.ac.uk, to be forwarded to the Chair.
Appendix B: Favourable Ethical Opinion (Chapter 3)

Hannah.Lawrence@port.ac.uk
11/11/14

Science Faculty Ethics Committee

Protocol Title: SPEC 2014-071, Helping adults tell the difference between children's truth and lies
Date received PI response from 1st review letter: 31/10/14
Date Reviewed: 11/11/14

FAVOURABLE OPINION

Dear Miss Lawrence,

Thank you for your submission for ethical review. Having completed their review, members of the Science Faculty Ethics Committee have reached a Favourable opinion of your proposed research.

Please notify the committee of any substantial amendments to the proposed procedures, send an annual report to the committee regarding study progress and a final study report once the study has concluded. Please send these to sci.fac@port.ac.uk. Thank you and the committee wishes you well with your study.

Dr Simon Kolstoe – Vice Chair SPEC

CC -
Holly Shawyer – Faculty Administrator

If you would like to offer any feedback on the Science Faculty Ethics Committee process please email sci.fac@port.ac.uk, to be forwarded to the Chair
Appendix C: Favourable Ethical Opinion (Chapter 4)

Favourable Opinion – SFEC 2015-021

Protocol Title: Can you tell the difference between children’s truths and lies?
SFEC Code: SFEC 2015-021
Date Resubmitted: 29 April 2015
Date reviewed: 6 May 2015

Thank you for resubmitting your application to the Science Faculty Ethics Committee (SFEC) for ethical review following the 1st SFEC review, in accordance with current procedures1. Thank you for the clear and comprehensive clarifications provided and the changes you have made. I am pleased to inform you that your application has been given a favourable opinion by SFEC.

Please notify us in the future of any substantial amendments that may be required to this study, by making an application for protocol amendment2.

Please also submit to ethics-sci@port.ac.uk an annual report on the progress of the study, and a final study report / publication when the study has concluded.

SFEC wishes you well with your study3.

Dr Jim House
Vice-Chair Science Faculty Ethics Committee
Review Chair

Information:
Holly Showyer – Faculty Administrator
Dr Lucy Akehurst – PhD 1st Supervisor – lucy.akehurst@port.ac.uk

1 Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 [to be updated].
2 Using the SFEC protocol amendment form.
3 If you would like to offer any feedback on the SFEC process please email ethics-sci@port.ac.uk, to be forwarded to the Chair.
Appendix D: Favourable Ethical Opinion (Chapter 5)

Favourable Ethical Opinion

Study Title: Helping adults detect children’s lies
Reference Number: SFEC 2015-064
Date Submitted: 24 November 2015

Thank you for resubmitting your application to the Science Faculty Ethics Committee (SEFC) for ethical review in accordance with current procedures¹, and for making the requested changes following the first SFEC review, and for the clarifications provided, particularly the explanations about the need for the children to lie.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A, and subject to standard general conditions (See Annex B).

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

Dr Jim House
Vice-Chair Science Faculty Ethics Committee

Annexes
A - Documents reviewed
B - After ethical review - Guidance for researchers

¹ Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).
Appendix E: Favourable Ethical Opinion (Chapter 6)

Faculty of Science

Hannah Lawrence
Department of Psychology
University of Portsmouth
Hannah.Lawrence@port.ac.uk

Science Faculty Ethics Committee
Science Faculty Office
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

T: 023 9284 3379
ethics-sci@port.ac.uk

Date 19 October 2015

FAVOURABLE ETHICAL OPINION

Study Title: Assessing child credibility in investigative interviews: Current practice and new interview techniques.
Reference Number: SFEC 2015-075 (Please quote this in any correspondence)

Thank you for submitting your application to the Science Faculty Ethics Committee (SEFC) for ethical review dated on 1 October 2015 in accordance with current procedures¹.

I am pleased to inform you that SFEC was content to grant a favourable ethical opinion of the above research on the basis described in the submitted documents listed at Annex A. Please include dates and version numbers on future submissions.

Please note that the favourable opinion of SFEC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including the University of Portsmouth or supervisor, prior to the start of the study.

Wishing you every success in your research

Yours sincerely,

Dr Jim House
Vice-Chair Science Faculty Ethics Committee

Information:

Dr Lucy Axehurst
Holly Shawyer - Faculty Administrator

¹ Procedures for Ethical Review, Science Faculty Ethics Committee, University of Portsmouth, October 2012 (to be updated).
Appendix F: UPR 16 form

**FORM UPR16**

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information)

<table>
<thead>
<tr>
<th>Postgraduate Research Student (PGRS) Information</th>
<th>Student ID: 569229</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGRS Name: Hannah Jayne Lawrence</td>
<td></td>
</tr>
<tr>
<td>Department: Psychology</td>
<td></td>
</tr>
<tr>
<td>First Supervisor: Dr Lucy Akehurst</td>
<td></td>
</tr>
<tr>
<td>Start Date: 1st February 2014</td>
<td></td>
</tr>
<tr>
<td>Study Mode and Route: Part-time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full-time</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title of Thesis:</th>
<th>'Mind the Gap': Exaggerating Verbal Differences between Children's True and False Reports using Cognitive Lie Detection Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thesis Word Count:</td>
<td>47,447 (excluding ancillary data)</td>
</tr>
</tbody>
</table>

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study.

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

**UKRIO Finished Research Checklist:**

(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukri.org/what-we-do/code-of-practice-for-research/)

<table>
<thead>
<tr>
<th>a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>b) Have all contributions to knowledge been acknowledged?</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>c) Have you complied with all agreements relating to intellectual property, publication and authorship?</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>e) Does your research comply with all legal, ethical, and contractual requirements?</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

**Candidate Statement:**

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

**Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCRIEC):** 2014-007 (#13-066), 2014-078, 2014-071, 2015-021, 2015-084, 2015-075

If you have not submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

N/A

UPR16 – August 2015

231
Appendix G: Journal article published from Chapter 2

Published online 20 December 2016 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/acp.3303

‘Look This Way’: Using Gaze Maintenance to Facilitate the Detection of Children’s False Reports

HANNAH LAWRENCE1*, LUCY AKEHURST1, AMY-MAY LEACH3, JULIE CHERRYMAN3, ALDERT VRIJ1, MEGAN ARATHOON1 and ZARAH VERNHAM1

1University of Portsmouth, Department of Psychology, Portsmouth, UK
2University of Ontario Institute of Technology, Toronto, Canada

Summary: In two experiments, we investigated whether imposing a secondary task in an effective technique for detecting child de- ceit. First, 83 children aged 8 to 11 years old provided either a true or false report of a recent school event. At interview, some children were asked to gaze towards either the interviewer’s face (IF) or a teddy bear’s face (TB), whereas some children were given no gaze instruction. In both the IF and TB conditions, liars told significantly fewer stories than truth-tellers. A total of 57% of children reported false stories under gaze conditions with and without guidance on level of detail. Evaluators discriminated truth from lies successfully when judging children instructed to look at the IF, but not when children were asked to gaze towards the TB. Evaluators who received guidance demonstrated better discrimination between true and false reports than evaluators who received no such information. Copyright © 2016 John Wiley & Sons, Ltd.

Child deception research has focused on both the developmental origins of children’s lie-telling behaviour, and the forensic implications of deceptive child testimonies going undetected (see Tabor & Crossman, 2012 for a review). Past research has painted a bleak picture: Children not only have the potential to lie in forensic interviews (Tye, Amato, Hoels, Devitt, & Peters, 1999), but, when the video- recordings of their statements are presented to legal professionals (e.g., police officers, judges), they experience great difficulty in uncovering false testimonies (Bala, Ramakrishnan, Lindsay, & Lee, 2014; Losh, Tabor, Lee, Bala, & Lindsay, 2004). Thus, if children do decide to provide deceptive reports, they could easily slip through the net resulting in miscarriages of justice that are damaging to both the victims and defendants (C’Donohue, Benotto, & Fanetti, 2010). Clearly, more effective deception detection strategies are needed.

Cognitive processing is an important factor in deception (Zuckerman, DePado, & Rosenthal, 1981), particularly for children whose growing cognitive abilities are closely related to their ability to maintain false reports (Tabor & Crossman, 2011). Indeed, children’s development of global executive functioning (Gordon, Lyon, & Lee, 2014) as well as their development of specific executive functions, such as inhibitory control, working memory, executive planning, and forward search planning, significantly contribute to their ability to conceal inculminating information when questioned (Alloway, McCullar, Alloway, & Hodda, 2015; Evans & Lee, 2011; Tabor & Lee, 2008; Williams, Leduc, Crossman, & Tabor, 2016). Furthermore, lie-telling proficiency follows the developmental patterns of cognitive processes, such as inhibitory control (Debey, De Schryver, Logan, Sachtzki, & Verschueren, 2015). This suggests that child liars, who are still developing certain cognitive skills that might facilitate their lie-telling, might be affected by any interview technique that impacts upon these skills.

*Correspondence to: Hannah Lawrence, Department of Psychology, University of Portsmouth, King Harry Building, King Henry I Street, Portsmouth, PO1 3DY, U.K.
E-mail: hannah.lawrence@port.ac.uk

Copyright © 2016 John Wiley & Sons, Ltd.
the tasks may arise. When working at cognitive capacity, performance will depend on a person’s ability to divide his or her attention in accordance with task demands. Attention can be flexibly allocated from moment to moment (Kahneman, 1973). As the secondary task becomes more difficult, additional resources can be allocated if the tasks share a particular pool of resources, then diverting resources from the primary task to the secondary task should result in a trade-off (i.e., decreasing performance for the primary task and increasing performance for the secondary task).

Imposing cognitive load in order to detect deception could be particularly effective with a younger population whose ability to manage their attentional resources has not yet fully matured. Before the age of 11 years, children find it difficult to differentially allocate their attention in dual-task processing (Irwin-Chase & Burns, 2000). Furthermore, research has shown that the increase in cognitive load experienced, when moving from single tasks to dual-tasks, is greater for children than it is for adults (Kane & Engle, 2000). Although 10-year-olds can allocate their attention similarly to adults, their control over attention management in response to task difficulty is not yet fully developed. In the context of the current study, this suggests that child liars may overcompensate for the rising demands of a secondary task, diverting too many resources away from the primary task of lying, thus decreasing their performance on this task. It is also possible that children may prioritise the primary task, sacrificing their performance on the secondary task.

To date, two studies have examined the effects of CLD techniques on children. First, Liu et al. (2010) asked unanticipated questions of children aged 10 to 12 years old about a non-experienced life event. They found that, compared to truthful-tellers, child liars were more likely to respond to unexpected questions. Second, Saykaly, Crossman, Morris, and Talwar (2014) imposed cognitive load by asking children to falsely allege or deny play with a certain toy using the ‘reverse order’ interview instruction. Their results revealed that reverse order recall made it harder for child liars to maintain their reports compared to child truthful-tellers, suggesting that telling a story backwards does increase cognitive demands. In summary, both these studies indicated that, when children have to perform a secondary task (i.e., answering a difficult question) at the same time as maintaining their false reports, their ability to maintain the lie is negatively affected. In the current experiment, the secondary task, introduced at interview, was an instruction to maintain gaze with either the interviewer’s face (IF) or a teddy bear’s face (TF). A secondary task that has yet to be investigated with children.

Gaze maintenance

Using a systematic approach, Glenberg, Schacter, and Robertson (1998) demonstrated that as the cognitive demands (i.e., cognitive difficulty) of a task increase, adults naturally avert their gaze. This cognitive strategy of gaze aversion is functional, as adults performed better on moderately difficult questions when they disengaged from (i.e. closed their eyes), rather than engaged with (i.e. looked at the interviewer’s nose), disruptive visual components in their environment. Looking towards a visual/social stimulus, therefore, interfered with their task performance when the cognitive demands of the task were moderate. This behavioural response to avoid cognitive overload has also been investigated with children. Dolbey-Stevenson, Bruce, Bowner, Longbotham, and Doyle (2012) compared gaze aversion behaviour in children aged 5 and 8 years old in response to easy (low cognitive load) and difficult (high cognitive load) questions. Results revealed that the older children averted their gaze away from the questioner’s face more frequently in response to rising question difficulty (i.e., cognitive effort), but that this gaze pattern was only observed for younger children and for certain types of questions. This suggests that gaze aversion is used as an overt response to cognitive effort more consistently with increasing age. In addition, there is evidence to support that the primary function of gaze aversion is to manage cognitive demands rather than as a response to social difficulty. Dolbey-Stevenson and Phelps (2005) measured gaze aversion in 8-year-old children who were questioned either face-to-face or via live video links. Results revealed that question difficulty strongly influenced gaze aversion in both interview conditions. In the current study, it was anticipated that, as children’s ages ranged from 8 to 11 years old, they would attempt to use gaze aversion to reduce cognitive effort more so in the ‘lie-telling’ condition where cognitive load is higher than in the ‘truth-telling’ condition. Requiring interviewees to maintain gaze during questioning, as was the case for this study, would disable this coping mechanism for liars and maintain the increased cognitive demands of providing a false report. Furthermore, as maintaining gaze is not a natural behaviour, it could be necessary for interviewees in this experiment to intentionally remind themselves to comply with our gaze instruction, creating additional cognitive load.

In a previous study, maintaining eye contact was used to impose cognitive load on adult interviewees (Vrij, Mann, Leal, & Peeters, 2010). The researchers found that requiring eye contact elicited two cognitive cues (out of 14 cues) that discriminated liars from truthful-tellers: namely, deceitful accounts contained fewer spatial details and were more chronological compared to truthful accounts. No significant differences were elicited between truthful-tellers and liars when interviewees were given no ‘eye contact’ instruction. In terms of detection accuracy, the small difference in elicited cues only improved lie detection accuracy from 44% in the ‘control’ condition to 53% in the ‘eye contact’ condition. As an improvement in accuracy rates is dependent on the exaggeration of behavioural differences between truthful-tellers and liars, eliciting two cognitive cues did not suffice. Vrij et al. (2010) suggest that these findings may be the product of anxiety for liars rather than increased cognitive load. Alternatively, previous research has shown that, even when adults find maintaining gaze with a person’s face to be more difficult than either looking at the floor or closing their eyes, this does not result in them performing worse in the former condition compared to the latter two gaze conditions (Dolbey-Stevenson, Bowner, & Bruce, 2011).

On the contrary, the difficulty that children experience when instructed to direct their gaze does translate into poorer
performance compared to a control condition involving no gaze instruction. In their first experiment, Doherty-Stanley et al. (2009) compared the effect of gaze instruction (look at the speaker vs. look at the floor vs. close your eyes) on both adult and 10-year-old children's task performance. Like adults, most children (83%) found looking at the floor or closing their eyes to be the easiest (least cognitively demanding) conditions. Results showed that, when children looked at the floor, this reported gaze translated into them performing significantly better than when they looked at the speaker. This difference in task performance was also found across Doherty-Stanley et al.'s subsequent experiments for different tasks and for a younger age group (4 years old). Children, therefore, experienced great difficulty in moderating the negative effects of gaze maintenance, which looked towards a face resulting not only in increased levels of cognitive demand for children, but also diminished task performance (Doherty-Stanley et al., 2003). Thus, it was anticipated, for the current study, that children's interview performance would be affected by gaze maintenance.

**EXPERIMENT 1**

The aim of Experiment 1 was to investigate whether an instruction to maintain gaze would exaggerate differences between children's true and false reports. With a view to the future clinical value of this research, it was important to consider how appropriate an instruction to maintain gaze would be with a child population. As maintaining gaze has already been linked to anxiety (Vrij et al., 2010), asking a child to look at an IF may intimidate some interviewees. In this study, we instructed some of the children to look at a face stimulus considered to be less intimidating: a TF. It should be noted that toys can be useful in child witness interviews (Wilson & Powell, 2001), and a teddy bear was chosen because it has a face and is non-threatening. In this experiment, we predicted that the children would experience more dual-task interference than truth-tellers when instructed to maintain gaze. That is, lie-tellers' ability to provide a detailed account would be more negatively associated with their level of gaze compliance, compared to truth-tellers (Hypothesis 1). Second, we anticipated that this dual-task interference would exaggerate subtle differences in level of detail between true and false reports. Thus, it was expected that children liars would provide reports that were significantly less detailed than those provided by child truth-tellers, and this difference in detail would be greater for children instructed to maintain gaze compared to children given no gaze instruction (Hypothesis 2).

**Method**

**Participants**

Eighty-five children (37 boys, 48 girls) aged 8 to 11 years old ($M = 10.46$ years, $SD = 0.81$ years) were recruited from four primary schools in the United Kingdom. Participant information sheets were sent home to children's legal guardians who returned a signed written consent form. The general procedure was outlined to the children to obtain their verbal assent to participation, but they were naive to the specific purpose of the study and to the anticipated effect of maintaining gaze. All children, who were asked to lie, complied with the request to lie. Verbalization was sought from teachers that they had not taken part in the event that they were interviewed about. All children received a certificate and a stationary set in exchange for taking part.

**Procedure**

The experiment took place in two quiet areas of each school and involved the Principal Investigator (PI) who ran the study and a Research Assistant who conducted all interviews and was blind to the aims and hypotheses of the study. All children were tested individually.

The PI invited each child to take part in a short interview about a recent event at their school: thus, events differed across schools. These events included a school sports day, a visit to the local cathedral, a school play, and a music concert. Children were randomly assigned to a Veracity condition within each year group in each school so that there were roughly equal numbers of multi-tellers and lie-tellers for each of the four events. Truth-tellers ($n=39$, $M_{age}=10.28$ years, $SD_{age}=0.73$ years) were interviewed after they had experienced the event and were asked to provide a truthful recollection of what happened. Lie-tellers ($n=46$, $M_{age}=10.42$ years, $SD_{age}=0.81$ years), on the other hand, were interviewed about an event that they had not experienced and were asked to convince the interviewer that they had already taken part in the event, when in fact they had not. This is similar to the veracity allocation carried out by other researchers interested in eliciting false allegations from children (e.g. Alchourrón, Künnki, & Höfer, 2001; Brunet et al., 2012; Lyon, Malloy, Quas, & Tulab, 2008).

Within their veracity groups, children were also randomly assigned to a Gaze Instruction condition. Look at the IF ($n=28$, $M_{age}=10.42$ years, $SD_{age}=0.81$ years) or Look at the TF ($n=29$, $M_{age}=10.52$, $SD_{age}=0.74$ years) or No gaze instruction (Control, $n=28$, $M_{age}=10.32$ years, $SD_{age}=0.91$ years). The teddy bear was seated on the interviewer's lap throughout all interviews (i.e. for all conditions). Prior to the interview, children in the IF and TF conditions were instructed by the PI to maintain gaze with the relevant face stimulus as much as they possibly could throughout the interview (i.e. to look at it as much as they could remember to do so). All children then received a sheet listing general themes that they could tell the interviewer about (e.g. talk about who was there, what happened, when it happened). This does not constitute coaching as neither multi-tellers nor lie-tellers were told exactly what they should say and they did not rehearse their story with the PI. Providing children with these themes was anticipated to elicit longer statements, allowing for more cues to detect to occur (Lesl, Vrij, Wammelink, Verheem, & Peizer, 2015; Vrij, 2015). All participants were given approximately three minutes to prepare themselves before the PI greeted them to the interview room. Before entering the interview room, children in the IF and TF condition were given a final reminder by the PI to maintain gaze with the relevant face stimulus. This was done out of concern of the interviewer so that the remained blind to the aims and hypotheses of the study.
The interview protocol reflected the initial stages of a Cognitive Interview (Fisher & Geseleman, 1992): A rapport-building phase (that took place off-camera) was followed by two open-ended questions. First, an invitation to provide a free, uninterrupted narrative (e.g., tell me everything that happened when you took part in your school sports day), and then, second, a request, to all interviewees, to provide one additional piece of information about an aspect of the event that they had not already mentioned. No other questions were asked. All children were video-recorded, and their interviews later transcribed. All interviewees were asked the following question, which served as a manipulation check: Where were you instructed to look during the interview? The response options were ‘interviewer’s face’, ‘teddy bear’s face’, or ‘no instruction given’.

Coding for detail
Two independent coders rated the children’s interview transcripts for number of details included. To make the coding more precise, all transcripts were coded for five different types of details: visual details (e.g., ‘white clay head’ contains three visual details), auditory details (e.g., ‘the teacher told us to take deep breaths’ contains one auditory detail), spatial details (e.g., ‘the stood behind the curtains’ contains one spatial detail), temporal details (e.g., ‘at the end we left’ contains one temporal detail), and action details (e.g., ‘we played football’ contains one action detail). One coder coded all of the transcripts for the current study, while the second coder coded a random sample of 20 transcripts. Considering that general level of detail is a reliable indicator of veracity (DePalma et al., 2003), total number of details was calculated for each interviewee, adding together the scores for all five detail types. Inter-class correlation coefficients (ICCs) were calculated for the two coders. Inter-rater reliability was high, with all ICCs demonstrating high levels of agreement between coders (visual details, ICC = .96; auditory details, ICC = .98; spatial details, ICC = .94; temporal details, ICC = .96; action details, ICC = .92; and total number of details, ICC = .98).

Coding for gaze maintenance
To provide an objective measure of gaze behaviour, two different independent judges, using INTERACT 14.0 software (Mangold, 2015), coded all interviews (from start to end) for the amount of time (in seconds) that the child interviewees gazed towards the IF and the TF. The duration of these gaze patterns for both face stimuli were then added together to give the total number of seconds spent gazing at the IF and the TF for each child. Percentage of time spent gazing towards both the IF and the TF were calculated by taking the total number of seconds spent gazing towards each face stimulus, dividing it by the total length of the interview in seconds and multiplying it by 100. Percentage of time spent gazing elsewhere was calculated by adding together the percentages for IF and TF and subtracting this total from 100. First, both raters coded 17 interviews (20% of the total) to check for inter-rater reliability. Inter-rater reliability was high for time spent looking at the IF (ICC = .90) and at the TF (ICC = .91). Rater 1 then coded the next 40% of the video recordings (n = 34) and Rater 2 coded the remaining 40% of the video recordings (n = 34). Percentage of time spent gazing at each face stimulus was calculated for each child by dividing the time spent gazing at the stimulus (in seconds) by the total duration of the interview (in seconds) and multiplying the result by 100.

Results
Manipulation checks
All 85 children correctly indicated where they had been asked to look during the interview. To test level of compliance more objectively, two-way ANOVAs were performed with Veracity and Gaze Instruction as the between-subjects factors. These were conducted to investigate differences in percentage of time spent gazing at (i) the interviewer’s face, (ii) the teddy bear’s face, and (iii) elsewhere (i.e. towards neither face stimulus). Figure 1 displays the distribution of gaze behaviour across ‘veracity’ conditions and Figure 2 across ‘gaze instruction’ conditions.

In terms of gazing towards the IF, there was a significant main effect of Veracity, F(1, 79) = 5.78, p = .019. Children providing a false report (M = 45.80%, SD = 22.41) spent a higher percentage of their interviews looking at the IF than children providing a true report (M = 35.34%, SD = 23.72), d = .46, 95% CI [.03, .89]. There was also a significant main effect of Gaze Instruction, F(2, 79) = 10.50, p < .001. Pairwise comparisons using Bonferroni adjustment showed that children instructed to look at the IF (M = 55.93%, SD = 24.97) spent a greater portion of the interview gazing at the IF than children instructed to look at the TF (M = 31.61%, SD = 23.41), p < .001, d = .98, 95% CI [.42, 1.52], or given no gaze instruction (M = 28.71%, SD = 17.49, p = .001, d = 1.04, 95% CI [.47, 1.59]). There was no difference between these later conditions, p = 1.00. There was no significant interaction effect, F(2, 79) = 1.16, p = .34.

In terms of gazing towards the TF, face was no significant main effect of Veracity, F(1, 79) = .32, p = .57. There was, however, a significant main effect of Gaze Instruction, F(2, 79) = 9.50, p < .001. Pairwise comparisons using Bonferroni
adjustment showed that instructing children to gaze at the TF (M = 16.77%, SD = 1.87%) resulted in a higher percentage of time looking at the TF than instructing children to look at the IF (M = 5.22%, SD = 5.21%, p < .001, d = .83, 95% CI [.29, 1.37]), or giving no gaze instruction (M = 4.66%, SD = 2.66, p < .001, d = .90, 95% CI [.35, 1.44]). There was no difference between these latter conditions, p = 1.00. There was no significant Vencity x Gaze Instruction interaction effect, F(2, 79) = .28, p = .76.

Finally, in terms of gazing elsewhere, there was a significant main effect of Vencity, F(1, 79) = 7.15, p < .009. Truth-tellers (M = 86.90%, SD = 34.22) spent a higher proportion of the interview looking elsewhere compared to liars-tellers (M = 44.48%, SD = 21.37, d = .54, 95% CI [.10, .97]). There was also a significant main effect of Gaze Instruction, F(1, 79) = 7.59, p < .001. Pairwise comparisons using Bonferroni adjustment showed that children given no gaze instruction (M = 61.73%, SD = 17.74) spent more time looking elsewhere compared to children instructed to look at the IF (M = 38.85%, SD = 24.21, p < .001, d = 1.08, 95% CI [.51, 1.54]). Percentage of time looking elsewhere did, however, not differ between children in the ‘control’ condition and those in the TF condition (M = 49.63%, SD = 22.61, p = .10). There was also no significant difference in percentage of time spent gazing elsewhere between children in the ‘IF’ condition and children in the TF condition, p = .28. There was no significant interaction effect, F(2, 79) = .80, p = .45.

In sum, children were able to comply with the instruction to look at the IF or the TF. That said, although our instructions did increase time spent gazing towards a specific face stimulus, overall compliance was relatively poor as the average participant complied with their gaze instruction for less than 50% of their interview. Furthermore, children in the ‘TF’ condition only spent 16% of the time looking at their specified stimuli and just as much time looking at the IF and elsewhere as children in the ‘control’ condition. This lack of compliance may be because gazing at a static toy when responding to a person is an unnatural behaviour. It could also be because the location of the teddy bear was problematic; staring at the interviewer’s lap may have seemed strange.

**Hypothesis testing**

Preliminary analyses revealed no significant effects of child age, child gender, or specific activity reported (e.g. sports day, school trip) during the interview, on any of the dependent variables. The data for all participants were, therefore, combined for subsequent analyses.

**Dual-task interference.** We investigated whether liars-tellers experienced more dual-task interference than truth-tellers, when given the secondary task of maintaining gaze with either the IF or the TF whilst being questioned. The ‘performance operating characteristic’ (POC, Norman & Bobrow, 1975) of truth-tellers and liars-tellers was calculated separately for children in both ‘gaze instruction’ conditions. By calculating Pearson’s correlations between the total number of details included in the interviewee’s account (i.e. level of detail) and the time they spent gazing towards either the interviewee’s or the TF (i.e. level of gaze compliance), we were able to examine to what extent the two tasks interfered with one another. High levels of interference would be characterised by a strong negative correlation between performances on both tasks (i.e. increasing compliance with the gaze instruction resulting in decreasing level of detail in responses).

First, when the secondary task required interviewees to look at the IF, findings revealed a weak, negative correlation for truth-tellers, r = -.28, p = .40, and a small to moderate, positive correlation for liars-tellers, r = .39, p = .16. Although these correlations are not significant, this may be because of the effect of a limited sample size. Following the suggestion of Ferguson (2009), we therefore looked at the effect size of these correlations as well as effect sizes are resistant to sample size influence, and thus provide a clearer measure of the magnitude of effect between variables (p. 552). Interpreting these r values as effect sizes (Field, 2013), the data showed that there was a small effect for truth-tellers and a medium effect for liars-tellers. This suggests that there was mild interference between truth-tellers’ ability to provide detailed answers and their compliance with the gaze instruction. However, it also shows that there was no interference for liars-tellers, whose level of detail in fact increased with their level of compliance with the gaze instruction. Second, when interviewees were instructed to look at the TF, there was no correlation between level of detail and compliance with the gaze instruction for truth-tellers, r = -.04, p = .91, nor for liars-tellers, r = .08, p = .78.

**Level of detail.** Preliminary analyses showed that true reports (M = 750.79, SD = 670.31) contained significantly more words than false reports (M = 503.33, SD = 500.88), t(83) = 1.82, p = .037, d = .40, 95% CI [.00, .82]. As longer reports allow for more details to occur, length of statement would have an effect on our analysis of total words. To take this effect into account, length of statement (in words) was entered as a covariate in our analyses. This is similar to previous work by Strömwall and Granhag (2005) when analysing reality monitoring scores.
First, a 2 (Veracity) × 3 (Gaze Instruction) ANCOVA was performed with total number of details as the dependent variable. There was a significant main effect of Veracity, $F(1, 78) = 8.44, p = 0.004$, a significant main effect of Gaze Instruction, $F(2, 78) = 3.16, p = 0.048$, and a significant Veracity × Gaze Instruction interaction effect, $F(2, 78) = 4.22, p = 0.018$. Descriptive statistics for each of the experimental cells are displayed in Table 1.

Of interest for the hypotheses is the Veracity × Gaze Instruction interaction effect. Separate ANCOVAs were conducted: first, for each of the Gaze Instruction conditions with Veracity as the independent variable, and second, for each of the Veracity conditions with Gaze Instruction as the independent variable. When children were instructed to gaze at the IF, truth-tellers provided significantly more details compared to lie-tellers, $F(1, 25) = 8.35, p = 0.007, d = 0.92$ (95% CI [1.13, 1.70]). Similarly, when children were instructed to look at the TF, truth-tellers provided more details in their statements than lie-tellers, $F(1, 26) = 5.88, p = 0.023, d = 0.83$ (95% CI [0.58, 1.59]). Veracity did not have a significant effect on the number of details provided by children who were given no gaze instruction, $F(1, 25) = 0.24, p = 0.63$. Irrespective of whether they were providing a true report or a false report, children in the control condition included the same amount of detail.

For children who provided a truthful account, there was a significant effect of Gaze Instruction condition, $F(2, 35) = 4.04, p = 0.026$. Post-hoc testing using Bonferroni adjustment revealed that truth-tellers who looked at the IF provided more details than truth-tellers who were given no gaze instruction, $p < 0.01, d = 0.80$ (95% CI [0.33, 1.29]). There was no difference in quantity of detail between truth-tellers looking at the IF and those looking at the TF, $p = 1.00$, and no difference between truth-tellers looking at the TF and those in the control condition, $p = 0.14$. For children who provided a fabricated account, there was no significant effect of Gaze Instruction, $F(2, 35) = 0.21, p = 0.81$.

**Discussion**

The analysis of the association between providing a detailed account and complying with the gaze instruction revealed a small positive effect for lie-tellers in the IF condition. That is, the more lie-tellers looked at the IF, the more details they gave. This was contrary to Hypothesis 1. Furthermore, the instruction to look at the TF did not elicit dual task interference for the liars but for the truth-tellers. Our theoretical assumption posited that lie-tellers, who have a more cognitively demanding primary task compared to truth-tellers, would reach the limit of their resources when a secondary task was imposed (Kornreich, 2006), and therefore, experience a high level of dual-task interference (Kornreich, 1983). However, our analysis of liars-tellers' dual-task interference does not support this theoretical assumption. Indeed, the positive relationship between level of detail and gaze compliance for lie-tellers instructed to look at the IF fully contradicts our hypothesis. This could be because of the cognitive resources required for each task originating from separate (limited) resources. Multiple resource theory (Wickens, 2002) posits that tasks that are structurally dissimilar, such as answering interview questions (verbal) and maintaining gaze (visual/aesthetic), will interfere less. This may explain why imposing cognitive load through constructing/maintaining a lie (verbal) and telling the lie in reverse order (verbal) had greater success in previous studies (Szykula et al., 2016; Venj et al., 2018) because the two tasks use similar cognitive processes.

An alternative explanation could be that the effect of gaze maintenance on task performance can vary dependent on the relevance of the visual stimulus to the primary task (Doherty-Sneddon et al., 2004). It could be the case, in the current study, that child lie-tellers instructed to look towards the IF found the information communicated by her face more task-relevant than truth-tellers. Liar-tellers, who are more concerned with appearing honest than truth-tellers (Venj et al., 2015), might have monitored the IF for feedback on how their deception was being received and used this to modify their responses (e.g., to say more to appear honest). However, this tactic works to their disadvantage, as longer statements are more likely to contain cues to deceit (Venj et al., 2015). This would particularly be the case for child interviewees who tend to reveal their deceit verbally (Tabwar & Lee, 2002). Furthermore, these unanticipated findings might be explained by differences in children’s developing cognitive abilities that are associated with lying ability, such as executive functioning (Tabwar & Crossman, 2011). Child lie-tellers in our study may have had good working memory skills that allowed them to look at the IF whilst telling their false report. Future research should investigate whether the effects of imposing cognitive load are moderated by children’s growing cognitive development.

Interestingly, truth-tellers instructed to look at the IF did experience some dual-task interference. This unexpected finding requires further investigation. As memory can be dual-limited (i.e., limited by a person’s ability to recall a past experience), it could be that factors other than gaze compliance influenced our child truth-tellers’ ability to provide a detailed account. Finally, the absence of dual-task interference for children instructed to look at the TF could be explained by the IF not being as cognitively effortful to look at as the IF. As the TF did not provide any relevant feedback, it was not necessary for the interviewees to monitor it for deception. Nevertheless, we suggest caution in interpreting these correlations because of their non-significant nature.

Irrespective of the findings for dual-task interference, significant differences in level of detail between child truth-tellers and child lie-tellers were only elicited when a secondary task was imposed. For children instructed to look at the IF, these findings are in line with previous work with
Maintaining gaze to detect children’s lies

adults (Vrij et al., 2010), which has also found exaggerated behavioural differences between truth and lies when gaze was maintained. For children instructed to look at the TF, these findings extend current knowledge and demonstrate that gaze towards a non-human stimulus could act as a less threatening, but still effective, substitute in practice. Although exaggerated differences occurred when a dual-task was imposed, it remains unclear from a theoretical standpoint if this was the case. The dual-task processes involved in providing a narrative and maintaining gaze require further examination to understand the theory behind this effect. Indeed, further probing of the significant interaction suggests that using different gaze instructions does not have an effect on false reports but rather has an effect on true reports. Thus, these exaggerated differences could be because of gaze maintenance facilitating longer truthful accounts rather than inhibiting false accounts. Our findings suggest that the request to look at the TF elicited more reports that were significantly more detailed than when no gaze instruction was provided. This may be because of the demeanour of our interviewer; supportive interviewers have been shown to dictate longer reports (Vrij, 2015). However, it is not within the scope of this research to draw any firm conclusions regarding these results. Furthermore, these findings should be interpreted with caution. Because of small experimental cell sizes, there is a risk of Type I error. This study, therefore, requires replication with a larger sample size to verify that the interaction effect remains significant.

In this study, we were able to examine the memory accuracy of the truth-tellers’ detailed reports. Based on the information provided by the schools, we were only able to establish whether the children had taken part in the events or not, but, because of the scope of the events, we were unable to capture all the information regarding the events to correctly and incorrectly details. Future research is required to explore the relevance and accuracy of the reports provided by truth-tellers in the ‘gaze instruction’ conditions to understand the specific benefits of eliciting more details in true reports.

In the current study, the interview protocol was short and non-experimental. Using open-ended questions did allow us to go beyond the majority of past research, which has primarily focused on forced-choice questions using temptation resistance paradigms, to examine how gaze maintenance would affect children’s longer narratives. However, this does not reflect interview protocols in real-life police investigations with child witnesses, where a variety of question types are used. We can, therefore, not generalise these findings to a whole police interview, but only to the beginning of the police interview where an uninterrupted free narrative is requested. Finally, our study represents a ‘best case scenario’ in which a child provides a long narrative. As we reduced our interview protocol to focus on two open-ended questions, it was important to facilitate long responses by providing all of the children with examples of the type of information they could provide and some time to prepare. Child witnesses typically provide shorter statements than both their adolescent and adult counterparts (Jack, Leow, & Zajac, 2014); this may be because of them not knowing what level of detail is required at interview (Lamb, Ostbak, Herschkowitz, Ezplin, & Horowitz, 2007). Future research should continue to test the generalisability of these findings by using a procedure where no examples are provided.

Despite the exaggerated difference in level of detail elicited between child truth-tellers and child lies-tellers in the dual-task gaze condition (compared to the single-task control condition), the major concern still remained whether evaluators would be able to discriminate between lies-tellers and truth-tellers more effectively when child interviewees were instructed to maintain gaze compared to when no gaze instructions were given. We investigated this issue in Experiment 2.

EXPERIMENT 2

In Experiment 2, we tested the prediction that evaluators would discriminate better between truth-tellers and lies-tellers instructed to maintain gaze, than truth-tellers and lies-tellers who were given no gaze instruction (Hypothesis 3).

We also examined whether telling evaluators that truth-tellers provide more detail in their reports than lies-tellers would improve discrimination accuracy. Previous research into training to improve lie detection has shown that informing evaluators about empirically supported verbal cues to deceit has the largest effect on their detection accuracy (Haush, Soporta, Michael, & Melmour, 2014). Overall, level of detail has been found to be a key indicator of veracity (DePaulo et al., 2003). It is also one of the general characteristics coded for in Criteria-Based Content Analysis (Stein & Kohn, 1989) that has received the most support for distinguishing between child truth-tellers and child lies-tellers in the predicted direction (Vrij, 2005). It was, therefore, anticipated that evaluators who received this guidance regarding detail would demonstrate better discrimination than evaluators who received no guidance (Hypothesis 4). It was further predicted that an improvement in discrimination, as a result of guidance, would be most pronounced when judging the credibility of children instructed to maintain gaze, because of a greater difference in detail being elicited in these conditions in Experiment 1 (Hypothesis 5).

Successful discrimination depends on whether evaluators can interpret behavioural cues correctly. It was, therefore, important to recognise that gaze inversion can be perceived as a strong indicator of deception (Global Deception Research Team, 2006), even though this cue is non-diagnostic (DePaulo et al., 2003). We could not rule out the possibility that gaze behaviour perceived to be somewhat ‘strange’ might impact on evaluators’ judgments of credibility. Half of the evaluators were, therefore, played visual-audio clips of the children’s interviews, and the other half were played audio-only clips. We anticipated that evaluators who watched the visual-audio presentations displaying the gaze maintenance behaviour would demonstrate a truth bias because gaze maintenance might be interpreted as a sign of truthfulness (Vrij et al., 2010) (Hypothesis 6).

Method

Participants

A sample of 192 adult evaluators (89 males, 103 females) with an age range of 18 to 76 years (M = 27.14 years,

Copyright © 2015 John Wiley & Sons, Ltd.
240

SD = 11.71 years) was recruited. One hundred and ten participants (52% of the total sample) were undergraduate students who received 0.5 course credit for their participation. The further 82 participants were members of the general public recruited via convenience sampling. The non-animate participants were not compensated for their participation.

**Interview clips**

A total of 30 interview clips were selected from the sample of 85 children in Experiment 1. There were 10 clips per ‘Gaze Instruction’ condition; within each of those three sets of 10 clips, there were five truth-tellers and five lie-tellers. In the first round of the interview clip selection process, all recordings that contained noise interference (e.g. school bell, road traffic) were excluded (n = 20). Second, clips in which the first false recall lasted longer than 30 s were removed (n = 7). This decision was made to limit the total duration of the study (50 min maximum), reducing potential fatigue effects on evaluators’ performance. The remaining 28 clips were divided by Gaze Instruction condition (IF, n = 22; TF, n = 16; CONTROL, n = 20), and five truth-tellers and five lie-tellers were randomly selected for each condition. The final 30 clips were edited down so that they only contained the child interviewee’s first false recall. This selection process resulted in an even distribution of gender (three boys to two girls, or two boys to three girls) in each Veracity × Gaze Instruction cell, except for the false reports in the ‘Control’ condition, which was all provided by boys. It was not anticipated that this would bias results as no response bias has been previously found for adults judging boys’ credibility (Talwar, Crossman, Gulman, Romanciuk, & Williams, 2009). Interview clips lasted from 53 s to 239 s (M = 135.68 s, SD = 56.16 s). A 2 (Veracity) × 3 (Gaze Instruction) ANOVA was performed to ensure that there were no significant differences in lengths of clip across conditions. There was no significant main effect of Veracity. F(1, 24) = 13.3, p = .02, no significant main effect of Gaze Instruction, F(2, 24) = .05, p = .99, and there was no significant Veracity × Gaze Instruction interaction effect, F(2, 24) = .02, p = .98. For each ‘gaze instruction’ condition, four random rotations of the 10 clips were created to reduce order effects.

**Guidance on detail**

Evaluators who received guidance were provided with a sheet stating that truth-tellers provided more detail overall in their accounts compared to liars, as has been reported in previous deception research (DePaulo et al., 2003) and was also found in Experiment 1. To help evaluate understand what the experiment meant by the term ‘detail’, five different types of detail were presented in a table. For each type of detail, a description, and an example of that detail were provided (i.e., ‘visual detail refers to what the interviewer said that they saw. For example, a red hat contains two visual details’). Participants were advised to refer back to the guidance sheet as much as they found useful when watching/listening to the interview clips and were able to ask the experimenter for clarification on these types of detail before and during the experiment.

**Procedure**

The study took place in a quiet environment with few distractions. In order to prevent evaluators from working on the assumption that they would be presented with equal numbers of truth-tellers and lie-tellers, two steps were taken. First, participants were informed that they would be asked to evaluate the veracity of twelve children interviews in turn (actually they only evaluated ten clips in total). Second, they were told that it was just as likely for a child to be telling the truth as it was for them to be telling a lie.

First, evaluators were randomly assigned to a Gaze Instruction condition. That is, they judged the credibility of 10 interview clips (five truth-tellers and five lie-tellers) from only one of the Gaze Instruction conditions in Experiment 1 (IF vs. TF vs. Control). Evaluators who were provided with guidance on detail received this at the beginning of the experiment. Half of the evaluators watched all of the interview clips in visual–audio format, whilst the other half watched all interview clips in audio–only format. Participants who watched visual–audio presentations of the interviewees in the ‘IF’ and ‘TF’ conditions were informed that the child interviewees had been asked by the experimenter to direct their gaze during the interviews. Evaluators who watched and/or listened to the clips, one at a time, via a computer. Headphones were provided to record their credibility judgments, evaluators were given a hard copy answer booklet. Following each interview clip, evaluators were asked to decide if the child interviewee was lying or telling the truth.

Participants’ dichotomous judgments (truth or lie) for each clip were used to measure hits (proportion of deceitful clips correctly identified as deceitful) and false alarms (proportion of truthful clips incorrectly identified as deceitful) for subsequent signal detection analysis.

**Results**

**Accuracy**

Overall accuracy (M = 51.72%, SD = 16.23) was not significantly different from chance, r(104) = 1.47, p = .14, but truth accuracy (M = 60.62%, SD = 20.35) was significantly above chance, r(91) = 7.16, p < .001, d = .52 (95% CI [37, 85]), and lie accuracy (M = 42.81%, SD = 21.23) was significantly below chance, r(91) = –4.59, p < .001, d = –.44 (95% CI [19, 48]). When evaluators judged the credibility of children instructed to look at the IF (M = 58.91%, SD = 16.44), they performed significantly better than chance, t(60) = 4.33, p < .001, d = 0.54 (95% CI [28, 80]). When judging children instructed to look at the TF (M = 47.97%, SD = 15.45) or children given no gaze instruction (M = 48.29%, SD = 14.54), they were no better than chance (p > .05). Moreover, when evaluators were guided to look for differences in detail (M = 33.96%, SD = 14.7), they were better than chance, t(65) = 2.27, p = .028, d = .33 (95% CI [0.1, .51]), but not when no guidance was provided (M = 40.48%, SD = 15.1), r(95) = –.34, p = .74.

**Signal detection analysis**

The application of signal detection theory to deception detection research has been largely recommended because it provides an opportunity to measure two conceptually different
parameters of accuracy (Meisser & Kassin, 2002; discrimination accuracy) — ability to discriminate liars from truth-tellers (in this experiment, referred to as d'). And, response bias — tendencies to favor a particular response (truth or lie) in this experiment, referred to as b. Means and standard deviations for discrimination accuracy and response bias across all conditions are displayed in Table 2.

**Discrimination accuracy.** A 3 (Gaze Instruction) x 2 (Guidance Provision) ANOVA was performed with participants' sensitivity scores (d') as the dependent variable to examine their ability to discriminate between truth- and liars.

First, there was a significant main effect of Gaze Instruction, F(2, 180) = 10.84, p < .001. Post-hoc analyses using Bonferroni adjustment revealed that evaluators discriminated better between children's truthful and deceptive accounts when the interviewees were instructed to look at the interviewer than when they were instructed to look at the TA, p < .001, d = .66 (95% CI [.30, 1.02]).

Second, there was a significant main effect of Guidance Provision, F(1, 180) = 4.20, p = .042. Pairwise comparisons using Bonferroni adjustment showed that evaluators who received guidance discriminated better between truth- and liars than those who received no guidance, d = .27 (95% CI [.04, .55]).

Finally, there was a significant Gaze Instructions x Guidance Provision interaction effect, F(2, 180) = 4.88, p = .009. We performed univariate analyses to test the effect of providing guidance within each gaze instruction condition. There was a significant main effect of Guidance Provision for evaluators judging the credibility of children instructed to look at the interviewer, F(1, 62) = 12.10, p < .001. For evaluators in the ‘TP’ condition, those who received guidance (M = .22, SD = .76) were able to discriminate better than those who received no guidance (M = -.38, SD = .63), d = .87 (95% CI [.35, 1.38]).

Table 2. Discrimination accuracy (d') and response bias (b) as a function of gaze instruction, guidance provision and modality of presentation

<table>
<thead>
<tr>
<th>Condition</th>
<th>d'</th>
<th>SD</th>
<th>b</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gaze instruction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Look at interviewer’s face</td>
<td>0.43***</td>
<td>.06</td>
<td>1.21**</td>
<td>.49</td>
</tr>
<tr>
<td>Look at Teddy bear’s face</td>
<td>-0.38</td>
<td>.75</td>
<td>1.02</td>
<td>.57</td>
</tr>
<tr>
<td>No instruction (control)</td>
<td>-0.08</td>
<td>.71</td>
<td>1.04</td>
<td>.58</td>
</tr>
<tr>
<td><strong>Guidance provision</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.20*</td>
<td>.44</td>
<td>1.12</td>
<td>.48</td>
</tr>
<tr>
<td>No</td>
<td>-0.32</td>
<td>.73</td>
<td>1.05</td>
<td>.56</td>
</tr>
<tr>
<td><strong>Modality of presentation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video audio</td>
<td>-0.64</td>
<td>.77</td>
<td>1.01</td>
<td>.55</td>
</tr>
<tr>
<td>Audio only</td>
<td>0.22**</td>
<td>.86</td>
<td>1.16</td>
<td>.48</td>
</tr>
</tbody>
</table>

*p < .05
**p < .01
***p < .001

There was no significant main effect of Guidance Provision for evaluators assigned to the ‘TP’ condition, F(1, 62) = 1.27, p = .26, or the ‘control’ condition, F(1, 62) = 1.15, p = .29. There were no other significant interaction effects (p-values > .05).

In a second level of analysis, d' values were compared to 0 (no ability to differentiate between children’s truths and lies) using one-sample t-tests. With regard to Gaze Instruction, evaluators could reliably discriminate child-truth-tellers from child-liars in the ‘TP’ condition, t(63) = 4.32, p < .001, d = .54 (95% CI [.28, .80]), but not in the ‘TP’ condition, t(63) = -.81, p = .42, nor the ‘no gaze instruction’ condition, t(63) = -.87, p = .38. For Guidance Provision, evaluators were able to discriminate reliably when provided with guidance, t(63) = 2.30, p = .024, d = .21 (95% CI [.03, .44]), but not when guidance was withheld, t(63) = -.20, p = .84.

Finally, we compared d' scores to 0 for the significant interaction between Gaze Instruction and Guidance Provision. When evaluators judged the credibility of children instructed to look at the interviewer, they were able to discriminate liars from truth-tellers when guidance was provided (M = .55, SD = .85), t(31) = 3.63, p < .001, d = .64 (95% CI [.26, 1.02]), or not (M = -.32, SD = .75), t(31) = -.43, p = .69, d = .43 (95% CI [.06, .93]).

For children instructed to look at the TA, evaluators were not able to discriminate between children’s truths and lies when provided with guidance (M = .22, SD = .76), t(31) = 1.65, p = .12, nor when there was no guidance provision (M = -.38, SD = .63), t(31) = 3.46, p = .002, d = .61 (95% CI [.23, .93]). That is, evaluators labelled the groups incorrectly (i.e. they tended to label liars as truth-tellers and truth-tellers as deceptive). Finally, when children were given no gaze instructions, evaluators were not able to discriminate truthful from fabricated reports, with guidance provision, (M = .17, SD = .78), t(31) = 1.25, p = .22, or without guidance provision, (M = .018, SD = .63), t(31) = .52, p = .61.

**Response bias.** Participants’ response bias (b scores) was investigated to see whether they tended to identify children as liars or truth-tellers in any particular condition. A three-way ANOVA, with Gaze Instruction, Guidance Provision and Modality of Presentation of the clips as between-subjects factors, revealed significant main effects of Gaze Instruction, F(2, 180) = 5.03, p = .007, and Modality of Presentation, F(1, 180) = 6.55, p = .01. First, responses were more biased when judging the credibility of children instructed to look at the interviewer (M = 1.21, SD = .49) compared to children instructed to look at the TA (M = 1.02, SD = .37), p = .02, d = .46 (95% CI [.10, .81]), and children given no particular gaze instruction (M = 1.01, SD = .38, 95% CI [.32, 1.11]), p = .019, d = .45 (95% CI [.10, .80]).

Response bias did not significantly differ between evaluators judging child credibility in the latter two gaze conditions (p = 1.00). Second, evaluators demonstrated more bias in the ‘audio-only’ condition (M = 1.16, SD = .48) than in the ‘video-audio’ condition (M = 1.01, SD = .35), d = .36 (95% CI [.07, .64]). There was no significant main effect of Guidance Provision, and there were no significant interaction effects (p-values > .10).

Using one-sample t-tests, each β was compared to 1 (no bias). In signal detection theory, β values below 1 signify a tendency to respond ‘yes’ (or ‘lie’ in the current study), whereas values above 1 signify a tendency to respond ‘no’ (or ‘truth’ in the current study; Stanislav & Todorov, 1999). Therefore, the subsequent analyses examined the existence and magnitude of the bias. With regard to Gaze Instruction, evaluation who judged the credibility of children instructed to look at the TF were significantly biased to respond ‘truth’, t(63) = 3.46, p < .01, t(64) = 4.3 (95% CI [1.18, .6]), whereas no significant response bias was found for evaluators who judged children instructed to look at the TF, t(63) = .35, p = .73, nor for evaluators who judged children in the ‘no gaze instruction’ condition, t(63) = .30, p = .77. In terms of Modality of Presentation, evaluators in the ‘audio only’ condition displayed a significant truth bias, t(95) = 3.18, p = .002, d = .33 (95% CI [1.12, .53]), whereas evaluators in the ‘video-audio’ condition showed no bias, t(95) = .17, p = .87.

Discussion

Instructing child interviewees to maintain gaze with the TF enabled evaluators to discriminate between truth-tellers and liars to a better degree than when no instruction was given, in spite of a significant truth bias. However, discrimination accuracy was not affected when child interviewees were instructed to gaze towards the teller’s face. Thus, Hypothesis 3 was partially supported. The ability to accurately detect deception for evaluators rating children instructed to gaze at the TF may be because of differences in details provided by child truth-tellers and child lie-tellers. The CLD approach posits that the ability to discriminate between truth and lies should increase with the activation and exaggeration of cognitive behavioural differences (Vrij, 2015). Considering that significant behavioural differences were elicited for both children instructed to look at the TF and children instructed to look at the TF, it is possible that the exaggeration of these cues might need to reach a certain threshold, beyond which they become more apparent to an evaluator. It is possible that this threshold was only reached when child interviewees were instructed to look at the TF, in turn, facilitating evaluators’ credibility judgments; but the threshold was not met when the children were asked to look at the TF.

Informing evaluators that truth-tellers provide more detailed reports compared to lie-tellers did improve their ability to detect deception, thus supporting Hypothesis 4. However, it is difficult to conclude about the extent to which evaluators applied this guidance to the interview clips. Although training in verbal content cues is recommended because it leads to the highest training effects, it is also important to note that false information regarding cues to deceit can work as effectively as true information (Huscher et al., 2014). To encourage evaluators to engage more with the guidance and base their final credibility judgments on this specific information, it would be better to use methods such as the Psychologically Based Credibility Assessment Tool (Evans, Michael, Messaux, & Brandov, 2013) that include the rating of diagnostic cues in the final credibility assessment.

Contrary to Hypothesis 5, the provision of guidance was not more beneficial when judging children who were instructed to maintain gaze compared to those in the ‘neutral’ condition. Indeed, the only benefit of providing guidance was that it protected evaluators in the ‘TF condition from incorrectly labelling child veracity. As children in this condition were neither maintaining eye contact, nor free to look where they wished, their ‘strange’ gaze behaviour of looking at the interviewer’s lap might have been interpreted incorrectly as suspicious. Directing evaluators’ attention towards what the child was saying, through the use of our guidance, and encouraging them to trust their own impression judgments on the child’s verbal behaviour, may have detracted from the misinterpretation of their ‘strange’ gazing towards the teller’s head.

Finally, although we predicted in Hypothesis 6 that evaluators who watched the visual-audio presentations displaying the gaze maintenance behaviour would demonstrate a truth bias, this was not the case. This lack of truth bias might be because of evaluators interpreting gazing maintenance behaviour differently than that suggested by the general deception literature. On the one hand, gaze aversion is believed to be a cue to deceit (Global Deception Research Team, 2006), but, on the other hand, nonverbal behaviour that deviates from the expected norm, such as staring, can also be perceived to be ‘gaze’ (Bond et al., 1992). It is not known to what extent gaze behaviour influenced evaluators’ judgments, or how much suspicion evaluators attached to this nonverbal cue; however, the lack of bias might suggest that opposing interpretations may have cancelled each other out. Alternatively, informing evaluators that children had been instructed to divert their gaze may have made them more aware of their own bias.

For the current study, evaluators were exposed to 10 interview clips. This may have led to evaluators comparing cues and information across instances. In real police investigations and court proceedings, it is likely that these comparisons will occur between children’s statements, adult’s statements, and physical evidence. Future research should try to replicate this scenario to understand how a police officer or judge might judge the credibility of a child both in isolation and in comparison to other sources.

GENERAL DISCUSSION

We conducted the first empirical investigation exploring the use of gaze maintenance to detect deception in child witnesses during investigative interviews. Similar to Vrij et al. (2014), we predicted that the interview strategy would magnify differences in level of detail between children’s true and false reports. We also expected that the exaggeration of this cue would facilitate evaluators’ ability to discriminate children’s lies from truths.

The present findings show that gaze maintenance can be effective for discriminating the credibility of child witnesses. In Experiment 1, lie-tellers provided significantly fewer details in their reports compared to truth-tellers but only when they were instructed to look towards the TF or a TF. No significant difference was elicited when a secondary task was absent. In Experiment 2, we found that the exaggeration of this diagnostic cue facilitated evaluators’
discrimination accuracy, but this was only when children were instructed to look at the IF.

Theoretically, the effect of imposing a secondary task on interviewee performance remains unclear. The findings of Experiment 2 make it difficult to discern whether the secondary task had any negative impact on interviewers’ memory or whether interviewers experienced any additional cognitive load. The latter issue may be because of the nature of the secondary task in this study and the difficulty in pinning down the exact cognitive mechanisms involved. As previously mentioned, the development of certain cognitive skills is closely linked to children’s propensity to tell and maintain lies (Tulvar & Cosmossom, 2011). It may therefore be wise, in future, to provide cognitive measures of the specific executive functions that the imposed secondary task aims to affect to be able to establish whether (i) there is a link between these cognitive skills and the performance on the tasks, and (ii) whether children’s ability to perform these cognitive skills predicts the effectiveness of imposing cognitive load. When testing dual-task methodologies, it would also be beneficial to obtain baseline measures of an individual’s performance on single tasks (Task A only and Task B only) to which their performance on a dual-task (Tasks A and B simultaneously) could be compared.

Our findings provide further support for the practical value of manipulating cognitive load as a potential means for discriminating between children’s true and false reports. In particular, the results demonstrate that the effects of imposing cognitive load are not limited to asking children to tell their stories backwards. This is beneficial because Skykula et al. (2016) found that reverse order recall can adversely affect the accuracy of both truthful and deceptive statements, suggesting that it might not be helpful in real police investigations. In our study, requiring child interviewees to perform the secondary task of maintaining gaze had a positive effect on truthful-tellers, eliciting more information from them than when no gaze instruction was given. This finding is in line with the primary goal of any investigative interview, which is to extract as much information as possible from the interviewee. This finding could be because of the interviewee’s supportive demeanor, which has been found with adults to elicit more details from truth-tellers than lie-tellers (Mann et al., 2013). Further investigation is required to determine whether it is the combined effect of a gaze maintenance instruction to witnesses and supportive interviewer behaviour that helps truth-tellers but not lietellers, rather than the technique on its own.

A practical limitation of using gaze maintenance with child interviewees may be its appropriateness in certain contexts. Maintaining gaze with an authoritative figure, such as a police officer, might be an intimidating task for children. Although none of the children instructed to look at the IF reported any discomfort, the average child did maintain gaze for more than half of their interviews. A recent school event is far less traumatic to talk about than incidents of physical and/or sexual abuse, which can be the main focus of police investigations involving child witnesses. Future research must examine the scope of the beneficial effects elicited in this study and balance them with potential discomfort in certain contexts. As such, the preliminary findings relating to an instruction to concentrate on the less intimidating teddy bear (or similar) should be extended.

Maintaining gaze, particularly with a child, is an effective strategy for judging the credibility of children. Future research should continue to explore the application of dual-task processing to child interviews by examining strategies that target children’s under-developed executive functioning, with a view to creating more appropriate secondary tasks for this potentially sensitive context.

REFERENCES


Copyright © 2018 John Wiley & Sons Ltd.
