

Scaffolding Children's Communication in Investigative Interviews

Miss Alex Jane Smethurst

A thesis submitted in partial fulfilment of the requirements of Teesside University for the

degree of

Doctor of Philosophy

December 2022

Acknowledgements

First and foremost, I would like to thank my Director of Studies, Dr Kimberly Collins. Her input and support have been invaluable in completing this thesis, and without her I very much doubt my PhD journey would have been nearly as enjoyable or rewarding. I would also like to thank the rest of my supervisors: Dr Nikki Carthy, Professor Becky Milne, and Dr Karri Gillespie-Smith. I genuinely could not have asked for a better team. Everyone has been amazing, providing their own unique contribution. I left every supervision meeting feeling truly inspired.

Next, I would like to thank my fellow PhD students. It has been brilliant to share and grow our ideas together throughout this process. I would also like to thank the research students that helped with my second study. I am really grateful for the time you all invested in the project. I would also like to extend my gratitude to the school and police forces involved in my research for being so accommodating with my data collection.

Finally, I would like to thank all my friends and family for your support. In particular I would like to thank my mum, dad, nana, and partner for always being by my side in what has not been the easiest of five years.

Publications and Conference Presentations

Smethurst, A., Collins, K., Carthy, N., Milne, B., & Gillespie-Smith, K. (2021). Can a communication assessment provide a reliable indication of a child's communication at interview? *Psychology, Crime & Law*.

<https://doi.org/10.1080/1068316X.2021.1962864>

Smethurst, A., Collins, K., Carthy, N., Milne, B., & Gillespie-Smith, K. (2019, May).

Communication Assessments: Informing Interviewers and Empowering Children?

Vulnerability, Justice and Communication Conference, Portsmouth.

Smethurst, A., Collins, K., Carthy, N., Milne, B., & Gillespie-Smith, K. (2017, May).

Scaffolding Children's Communication in Investigative Interviews. Poster presented at Children, Justice and Communication Conference, Portsmouth.

Abstract

The Youth Justice and Criminal Evidence Act (YJCEA, 1999) introduced a number of ‘special measures’ to enable young witnesses to give their best evidence. One of the ‘special measures’ was the Registered Intermediary (RI) provision. In order to implement this ‘special measure’ the Witness Intermediary Scheme (WIS) was developed. Despite the Scheme having now been active for over a decade, relatively few studies have examined the work of RIs (Collins et al., 2017). The aim of this thesis was to expand upon the limited body of research regarding the RI provision and pre-interview communication assessments – a central aspect of the RI role (but which can also be conducted by police officers). The thesis comprises of four studies designed to examine different aspects of the RI role. Study one examined RIs’ beliefs regarding children’s memory and investigative interview practice, and compared these to the beliefs of lay people. Significant gaps in the RIs’ knowledge were identified, with the RIs having an increased propensity to express uncertainty relative to lay people. Study two examined the efficacy of pre-interview communication assessments using the ‘Unpacking the Box’ framework (Triangle, 2015). The framework is used by RIs and police officers, prior to investigative interviews, to assess children’s communication abilities. Pre-interview assessments, using the framework, provided a good indication of children’s abilities in all areas of cognition, included in the study, other than resistance to suggestion. The pre-interview assessment was also found to be superior to professional judgement in determining children’s use of ground rules, responsiveness, and drawing ability. Study three examined the demographic and cognitive variables thought to be associated with children’s recall and use of the ground rules. Unlike the cognitive variables (i.e., visuospatial ability, expressive language, receptive language, and attention) age did not enter as a significant predictor in any of the regression models. This suggests that cognitive factors may be more important than age in predicting children’s event recall and use of the ‘you got it wrong’ rule. Study four examined

the impact of the RI provision on practice in real-world investigative interviews with children. Significant differences were found, between the RI and no RI interviews, in relation to the prevalence of multiple questions and the use of certain communication aids (e.g., dolls, models, and figures). Although the RI provision appeared to have some impact upon interview practice, it did not lead to child witnesses providing more detailed accounts. The findings of this fourth study suggest that there may be other benefits of the RI provision but that additional training relating to the ABE guidance (Ministry of Justice, 2011) is required. Overall, this thesis has provided an insight into the RI provision and the efficacy of pre-interview communication assessments using the 'Unpacking the Box' framework (Triangle, 2015). However, further research is needed in order to fully understand how effective each is in scaffolding children's communication in an investigative interview context.

Table of Contents

Acknowledgements	2
Publications and Conference Presentations.....	3
Abstract	4
Table of Contents	6
List of Tables.....	9
List of Figures	12
List of Acronyms.....	13
Overview and Aims of Thesis	15
Chapter One: Real-World Contextualisation of Thesis.....	19
1.1 Achieving Best Evidence Guidance	23
Chapter Two: Psychological Framework for Scaffolding Children’s Communication	26
2.1 Child Development	26
2.1.1 Memory	26
2.1.2 Suggestibility.....	32
2.1.3 Working memory	39
2.1.4 Attention.....	42
2.1.5 Language	42
2.2 Scaffolding Communication	49
2.2.1 Social Support and Rapport.....	49
2.2.2 Question Types.....	63
2.2.3 Ground Rules.....	73
2.2.4 Communication Aids.....	86
2.2.5 What Happens in Practice	94
2.2.6 Registered Intermediaries and Pre-Interview Communication Assessments	103
Chapter Three: Registered Intermediaries’ Beliefs about Children’s Memory and Investigative Interview Practice (Study 1).....	124
3.1 Method	130
3.1.1 Design.....	130
3.1.2 Participants	130
3.1.3 Materials.....	131
3.1.4 Procedure.....	132
3.2 Results	136
3.3 Discussion	144

Chapter Four: The Ability of Pre-Interview Assessments to Inform Interviewers (and Registered Intermediaries) about Children’s Communication (Study 2).....	150
4.1 Method	153
4.1.1 Design.....	153
4.1.2 Participants	153
4.1.3 Materials	154
4.1.4 Procedure.....	164
4.1.5 Predictions and Coding	166
4.1.6 Inter-Rater Reliability	168
4.2 Results	169
4.3 Discussion	177
Chapter Five: Individual Differences in Children’s Recall and Use of Ground Rules in an Investigative Interview Context (Study 3)	186
5.1 Method	193
5.2 Results	194
5.3 Discussion	203
Chapter Six: The Role of Registered Intermediaries in Scaffolding Children’s Communication in Investigative Interviews (Study 4)	211
6.1 Method	214
6.1.1 Sample	214
6.1.2 Procedure.....	216
6.1.3 Coding	217
6.1.4 Inter-Rater Reliability	225
6.2 Results	226
6.3 Discussion	238
Chapter Seven: General Discussion	257
7.1 Summary of Main Findings.....	257
7.2 Theoretical Implications.....	259
7.3 Practical Implications	266
7.4 Methodological Implications.....	272
7.5 Areas for Future Research.....	275
7.6 Conclusion.....	278
References	280
Appendices	329
Appendix A – Empirical Evidence Supporting Questionnaire Items	329
Appendix B – Reflection.....	332
Appendix C - Rule Cards	333
Appendix D - Interview.....	334

Appendix E – Empirical Evidence Supporting Predictions	339
Appendix F – Date of Real-World Interviews	342

List of Tables

	Page
<hr/>	
Table 2.1	
Language Developmental Milestones	45
Table 3.1	
Questionnaire Items	134
Table 3.2	
Registered Intermediary and Layperson Beliefs Regarding the 20 Item Questionnaire	142
Table 4.1	
Number of Correct and Incorrect Predictions Pertaining to Interview Performance	176
Table 5.1	
Correlations between Predictor Variables and the Outcome Variables	196
Table 5.2	
Stepwise Multiple Regression Predicting Amount of Correct Investigation Relevant Information	198

Table 5.3

Stepwise Multiple Regression Predicting Amount of Incorrect Investigation Relevant Information	199
--	-----

Table 5.4

Stepwise Multiple Regression Predicting Amount of Confabulated Investigation Relevant Information	200
---	-----

Table 5.5

Stepwise Multiple Regression Predicting Accuracy of Investigation Relevant Information	201
--	-----

Table 5.6

Stepwise Multiple Regression Predicting Children's Use of the 'You Got It Wrong' Rule	202
---	-----

Table 6.1

Descriptions and Exemplars of the Different Question Types	218
--	-----

Table 6.2

Descriptions of the Different Interventions	222
---	-----

Table 6.3

Investigation Relevant Information Coding Scheme	224
--	-----

Table 6.4

Means and Standard Deviations of Question Types in Interviews With and Without a Registered Intermediary	232
---	-----

Table 6.5

Median and Range of Rapport Scores in Interviews With and Without a Registered Intermediary	233
--	-----

Table 6.6

Number of Registered Intermediary and No Registered Intermediary Interviews Using Each Communication Aid	235
---	-----

Table 6.7

The Purpose for Which Communication Aids were Used in the Registered Intermediary and No Registered Intermediary Interviews	236
--	-----

List of Figures

	Page
Figure 4.1	
Assessment Plan	159

List of Acronyms

	Acronym
Achieving Best Evidence Language Screen	ABELS
Achieving Best Evidence	ABE
Autism Spectrum Disorder	ASD
Child Sexual Abuse	CSA
Cognitive Interview	CI
Criminal Justice System	CJS
Crown Prosecution Service	CPS
Her Majesty's Crown Prosecution Service Inspectorate	HMCPSI
Her Majesty's Inspectorate of Constabulary	HMIC
Investigation Relevant Information	IRI
Ministry of Justice	MoJ
National Institute of Child Health and Human Development Protocol	NICHD
Non-Registered Intermediary	NRI
Registered Intermediary	RI
Theory of Mind	ToM
Typically Developing	TD

Witness Intermediary Scheme

WIS

Youth Justice and Criminal Evidence Act

YJCEA

Overview and Aims of Thesis

“the age of a witness is not determinative on his or her ability to give truthful and accurate evidence” (R. v. Barker, 2010, para. 40).

Above is a quote from an Appeal Court judgement in which a conviction for rape was upheld based on the evidence of a child aged three at interview and four at trial, recalling events that happened when she was two. This case demonstrates how the evidence of young children can be instrumental to justice and emphasises why it is so important that these children are given a voice. However, some children, particularly those that are very young or have a concurrent vulnerability, may require additional scaffolding and support in order for their voice to be heard. Understanding how to scaffold children’s communication thus has major theoretical, research, and real-world implications; and is focus of this thesis. Below is a brief overview of the thesis structure and aims.

Chapter One: Real-World Contextualisation of Thesis. The thesis begins by briefly discussing the challenges associated with interviewing children about suspected abuse and the development of current best practice guidance. Although the chapter acknowledges a number of different guidance documents and interview protocols, the main emphasis is upon Achieving Best Evidence (ABE; Ministry of Justice [MoJ], 2011) as this is the guidance which currently governs practice in England and Wales. The chapter also introduces the reader to ‘special measures’. ‘Special measures’ are adjustments designed to make the Criminal Justice System more accessible to vulnerable witnesses and includes the Registered Intermediary (RI) provision. RIs are professionals who facilitate two-way communication between vulnerable witnesses (e.g., children) and criminal justice practitioners (e.g., police officers, barristers). Despite the RI provision having been available in England and Wales since 2004, there has been a paucity of research into the work of RIs. As such, the knowledge and practice of RIs is the central focus of this thesis.

Chapter Two: Psychological Framework for Scaffolding Children's

Communication. Chapter two discusses the main areas of child development, namely memory, attention, and language, which can impact upon a child's ability to recall events at interview. It also considers ways in which children's developmental difficulties can be overcome through scaffolding and reviews empirical research on rapport, question types, ground rules, and communication aids. Adherence to the current best practice guidance is then discussed, along with potential barriers to its implementation. Finally, the chapter reviews the current body of literature regarding RIs and discusses the importance of pre-interview communication assessments, which are central component of the RIs' work. The purpose of a pre-interview assessment is to determine whether a witness has the ability to give evidence, whether RI assistance will improve the quality of the witness' evidence, how best to communicate questions to the witness, and if there are any adjustments that can be made to enable the witness to communicate more effectively (MoJ, 2020a).

Chapter Three: Registered Intermediaries' Beliefs about Children's Memory and Investigative Interview Practice (Study One). Study one involved 32 RIs and 61 lay people completing an online questionnaire regarding their beliefs about children's memory and investigative interview practice. The aim of the study was to compare the beliefs of the two groups in relation to one another, and in the context of current research findings. The study also sought to determine whether there are any gaps in the knowledge of RIs which can be used to inform the content of future training and CPD opportunities.

Chapter Four: The Ability of Pre-Interview Assessments to Inform Interviewers (and Registered Intermediaries) about Children's Communication (Study Two). Study two was an experimental study in which 51 children, aged four to nine, took part in a staged event and were interviewed a week later about their experiences. Prior to the interviews the researcher made a number of predictions regarding the children's interview performance.

Some of the children had a pre-interview communication assessment which formed the basis of these predictions, for the other children the predictions were based solely on professional judgement. The aim of the research was to examine the efficacy of pre-interview communication assessments using the 'Unpacking the Box' framework (which is utilised by both RIs and police officers; Triangle, 2015). The research sought to determine the ability of these assessments to provide a reliable indication of a child's communication abilities. The analysis thus looked at how closely the predictions, based on the communication assessment versus professional judgement, reflected interview performance.

Chapter Five: Individual Differences in Children's Recall and Use of Ground Rules in an Investigative Interview Context (Study Three). Study three sought to analyse the relationship between individual difference variables (i.e., demographic and cognitive) and children's recall and use of the ground rules. The aim of the research was to identify predictors of children's recall (i.e., detail and accuracy) and predictors of children's understanding and use of ground rules (i.e., 'I don't know', 'I don't understand'). An awareness of these factors could prove helpful for police officers in identifying when RI assistance is required, and helpful for RIs in determining which areas to prioritise during their pre-interview communication assessments.

Chapter Six: The Role of RIs in Scaffolding Children's Communication in Investigative Interviews (Study Four). Study four involved the analysis of real-world child investigative interviews. Interviews with and without an RI were compared in respect to question types, communication aids, levels of rapport, and the amount of information provided by the children. The aims of the research were to identify what works when interviewing child witnesses, and to determine the benefits of the WIS along with any areas that may require improvement.

Chapter Seven: General Discussion. This chapter summarises the main findings of the four studies included within the thesis. This is followed by a discussion of the theoretical, practical, and methodological implications of the research. Suggestions are then made regarding areas for future research and further exploration. Finally, the thesis concludes by providing the reader with a holistic appraisal of the thesis and its findings.

Chapter One: Real-World Contextualisation of Thesis

Each year millions of children are victims of abuse and / or neglect (Lytle et al., 2019). A statistic briefing by the NSPCC (2021) estimated that 1 in 20 children in the UK have been sexually abused, with the number of child sex offences reported to the police increasing rapidly in recent years. In 2014/15 there were approximately 38,000 child sex offences reported to the police, compared to over 65,000 in 2019/20 (NSPCC, 2021)¹. This has led to more children coming into contact with the British Criminal Justice System (CJS). However, the system, as with other justice systems internationally, was not designed with children in mind. In order to meet the needs of child witnesses the CJS has had to adapt. Legislative changes (e.g., Youth Justice and Criminal Evidence Act [YJCEA], 1999) have been introduced in an attempt to make the legal process more developmentally appropriate. These changes, in England and Wales, have included the introduction of ‘special measures’ including providing evidence via live link, video recorded evidence in chief, being screened from the accused, removal of wigs and gowns, and aids to communication such as the assistance of an RI (Burton et al., 2006). Detailed best practice documents and investigative interview protocols (e.g., Memorandum of Good Practice [MOGP], Home Office, 1992; ABE, MoJ, 2011) have also been developed.

The primary purpose of an investigative interview is to elicit an accurate and complete account (Vrij et al., 2014), thus ascertaining what happened and who did it (Milne & Bull, 2006). Interviews are instrumental in any criminal investigation (Ernberg, 2018). The information gathered may not only provide important leads in the investigation but may also serve as evidence in subsequent legal proceedings (Vrij et al., 2014). Section 27 of the YJCEA (1999) permits a child’s ABE interview (MoJ, 2011) to be used in court as their

¹ These figures are unlikely to reflect the true extent of child abuse in the UK. Many children do not disclose (e.g., 88%, Lahtinen et al., 2018).

direct-examination. This has led to the ABE interview (MoJ, 2011) constituting almost all of the direct-examination in the majority of contemporary cases in England (Henderson & Lamb, 2017). Other countries including Australia (Office of the Director of Public Prosecutions and Australian Federal Police, 2005), New Zealand (Evidence Act, 2006), and the United States of America (Porter, 2018) have similar provisions in place.

The investigative interview can prove even more crucial when there are allegations of child abuse or maltreatment. Child abuse / neglect is considered one of the most difficult crimes to detect, investigate, and prosecute as children are often reluctant to talk about abuse (Lytle et al., 2019). Children can be reluctant to disclose abuse for a number of reasons. Some children may feel culpable or complicit in the abuse (Lyon & Ahern, 2011; Orbach et al., 2007), other children may have been threatened or sworn to secrecy (Hershkowitz, Lanes, & Lamb, 2007), and others may not disclose through fear of disrupting the household order and through a desire to protect family members and friends (Plastock, 2018). These challenges are further compounded by the fact that external evidence (e.g., eyewitness testimony or medical evidence) is seldom available (Hritz et al., 2015). The child's interview thus forms the crux of the prosecution's case and is pivotal to attaining a conviction (Willcock et al., 2006).

However, a child, especially one who is very young, may struggle with the demands of a traditional investigative interview - they may not understand the purpose or what is required. This is due to children typically conversing with knowledgeable adults such as parents and teachers. Children therefore become accustomed to the adult knowing the answer (Lyon, 2014). However, in an investigative interview these roles are reversed. The adult (i.e., interviewer) is naive to the truth, whilst the child is considered to be the 'expert'. As such the child is expected to do the majority of the talking (Malloy et al., 2015), again something children are not familiar with. Guidance documents and interview protocols have thus been developed to overcome these difficulties. The documents provide interviewers with the

appropriate tools to create an environment in which a child is able to understand their role and in turn provide their best evidence.

An abundance of guidance documents and interview protocols have been developed internationally to enable children to give their best evidence. Although not an exhaustive list, these include the MOGP (Home Office, 1992), ABE (MoJ, 2011), the National Institute of Child Health and Human Development (NICHD) Protocol (Lamb et al., 2007), the American Professional Society on the Abuse of Children (APSAC) guidelines (Saywitz et al., 2011), Step-Wise Interview (Yuille et al., 1993), Scenario Model (Rispen & Van der Sleen, 2017), and the CornerHouse Forensic Interview Protocol (Anderson, 2013). The overarching principles of these documents are very similar. All emphasise the importance of building rapport and eliciting a free narrative account. Despite these similarities there is disagreement between the guidelines as to how this is best achieved. One of the greatest differences relates to the rigidity of the protocols. For example, the NICHD Protocol (Lamb et al., 2007) is highly structured, whilst both the CornerHouse Forensic Interview Protocol (Anderson, 2013) and ABE (MoJ, 2011) offer far more flexibility. There is also a disparity in the extent to which these protocols have been subject to empirical research, with the NICHD Protocol (Lamb et al., 2007) receiving the most attention (Anderson et al., 2014). In comparison the ABE guidelines (MoJ, 2011) have undergone little exploration (different aspects of the guidance have received differing amounts of attention), despite informing best practice across England and Wales since 2002. Given that the current research will analyse interviews conducted in England and Wales, ABE (MoJ, 2011) will be at the forefront of further discussions.

In 2002, ABE (Home Office, 2002) replaced the Memorandum of Good Practice (Home Office, 1992). This was in order to accommodate the reforms set out in the YJCEA (1999). Compared to its predecessor, the ABE guidance (Home Office, 2002) placed a greater

emphasis on planning and witness care. Although ABE (Home Office, 2002) endorsed the same four-phase approach as the MOGP (Home Office, 1992), there was an increased emphasis upon the use of open-ended questions. The guidance also placed a greater importance upon individual differences (e.g., age, disability) and the utility of alternative interviewing procedures (Davies et al., 2016). The first revision of ABE took place in 2007 (Office for Criminal Justice Reform, 2007). The revision took into account the findings of a Home Office survey of vulnerable and intimidated witnesses (Hamlyn et al., 2004), incorporating an entirely new chapter dedicated to intimidated adults. There was a further revision of the guidance in 2011 (MoJ, 2011) which took into account the changes brought about by the Coroners and Justice Act (2009) and the advice of the Association of Chief Police Officers (2010). The revision maintained the traditional role of an investigative interview, namely eliciting the witness' account. However, an additional component was added - case-specific questions arising from other evidence – which would assist the investigation. The revised guidance also provided additional advice on the use of communication aids and extended coverage of the Cognitive Interview (CI; Davies et al., 2016).

The CI (Fisher & Geiselman, 1992) is one of the most widely researched and used methods to interview witnesses (Paulo et al., 2017). It is argued that the CI is part of a therapeutic jurisprudence approach and, as such, may enhance the psychological well-being of individuals interviewed using this method (for a more in-depth discussion see Dodier & Otgaar, 2019). The CI (Fisher & Geiselman, 1992) consists of four cognitive mnemonics designed to enhance recall: report everything, mental reinstatement of context, change order, and change perspective. It also includes several social and communicative components deemed necessary for conducting appropriate investigative interviews. These include rapport building, transferring control of the interview to the witness, and witness-compatible

questioning (Fisher & Geiselman, 1992). Many studies have shown that the CI can increase the amount of information recalled by witnesses, without compromising accuracy (e.g., Memon et al., 2010; Paulo et al., 2013). This has become known as the CI superiority effect (Paulo et al., 2017). However, not all components of the CI (Fisher & Geiselman, 1992) contribute equally to this effect. This is alluded to in the ABE guidance (MoJ, 2011). The guidance cautions against the use of certain components (e.g., change perspective) with young children. It also states that the CI should only be used by interviewers who are adequately trained. Given that the focus of the current research is on younger children, this thesis will not cover the CI in extensive detail (for further information see Geiselman & Fisher, 2014). However, occasionally comparisons will be made between the CI components (Fisher & Geiselman, 1992) and aspects of ABE (MoJ, 2011).

1.1 Achieving Best Evidence Guidance

ABE (MoJ, 2011) recommends that a standard investigative interview should consist of four stages: rapport, free narrative account, questioning, and closure. Below is a very brief summary of each of the phases as outlined in the guidance.

Rapport. The rapport phase of an interview should include preliminary introductions. Preliminary introductions should include who is present in the interview and a statement about the day, date, time, and place (note: this can take place without the interviewee present). Ground rules (i.e., instructions associated with the communicative expectations of the interview) should also be covered, along with an exploration of truth and lies. If necessary, there may also be a discussion of neutral topics (for witnesses who may benefit from a lengthy discussion of neutral topics this can take place prior to the interview during witness preparation). However, this should be kept to a minimum, as a lengthy discussion about irrelevant matters risks tiring the witness and distracting the court given that the interview is played to the jury at trial.

Free Narrative. The second phase of the interview should involve initiating and supporting a free narrative account. This is achieved through the use of an open-ended invitation (e.g., 'tell me what happened') and non-specific prompts² (e.g., 'tell me more about that'). Active listening is also instrumental during this phase (e.g., making the witness aware that what they have said has been received by the interviewer).

Questioning. Following the free narrative, it may be necessary for the interviewer to pose questions to the witness in order to expand and clarify their account. The interviewer should separate the witness' account into manageable topics and use open-ended and specific-closed questions to probe for further information.

Closure. Closure is the final stage of the investigative interview and should involve a summary of what the witness has said which can lead to further retrieval of information. The witness should also be thanked for their time, asked if they have any questions, and given information about what will happen next.

ABE (MoJ, 2011) also provides guidance on using the 'special measures' introduced by the YJCEA (1999). One of the 'special measures' was the intermediary provision. In order to implement this 'special measure', the Witness Intermediary Scheme (WIS) was created. Through the WIS, RIs are made available to certain categories of vulnerable witnesses³ (see section 2.2.6). The ABE guidance states that:

An intermediary may be able to help improve the quality of evidence of any vulnerable adult or child witness (as defined in Section 16 Youth Justice and Criminal

² The term 'prompts' is used here as opposed to 'questions' (the term used throughout the majority of the thesis) to reflect the wording used in the ABE guidance (MoJ, 2011).

³The YJCEA explicitly excludes defendants. However, a Non-Registered Intermediary (NRI) can be appointed to assist a vulnerable defendant at interview or court (see section 2.2.6).

Evidence Act 1999) who is unable to detect and cope with misunderstanding, or to clearly express their answers to questions, especially in the context of an interview or while giving evidence in court. (MoJ, 2011, p.59)

Thus, the role of an RI is to facilitate two-way communication between a vulnerable witness and criminal justice practitioners (e.g., police officers, barristers), to ensure that communication is as complete, coherent, and accurate as possible (MoJ, 2020a). However, due to a paucity of research (Collins et al., 2017) little is known about how RIs scaffold and support children's communication during investigative interviews. A psychological framework for understanding and scaffolding children's communication in investigative interviews will be presented in the next chapter.

Chapter Two: Psychological Framework for Scaffolding Children's Communication

2.1 Child Development

Current legal guidance does not place any restrictions upon the age with which a child is deemed a competent witness. With adequate support and appropriate questioning, even very young children can provide accurate and reliable accounts of past events (Brown & Lamb, 2015). In order to provide accurate and reliable accounts, children need to utilise multiple cognitive functions (Anderson et al., 2009). However, the cognitive abilities of children are not on par with those of adults. As such children, particularly those that are very young, may require additional scaffolding and support during an investigative interview (Oxburgh et al., 2010). In order to effectively tailor this support, interviewers need to consider the following: memory, attention, language comprehension, and other aspects of a child's development. As an investigative interviewer, it is important to be aware that the trauma associated with child abuse can also trigger numerous psychological and physical problems (Gruhn & Compas, 2020; Hobbs & Goodman, 2018). These may include insecure attachment, anxiety, depression, dissociative tendencies, post-traumatic stress disorder, and drug and alcohol misuse (for a review see Klika & Conte, 2017).

2.1.1 Memory

During an investigative interview, children will be utilising different aspects of memory. Memory involves a series of three stages: encoding, storage, and retrieval (Holliday et al., 2011). Encoding refers to the process of converting perceptual information into a form that can then be stored in the memory system and retrieval refers to the process of recovering the stored information from the system. It is now widely accepted that there are several long-term memory systems (Eysenck & Keane, 2015). The most important distinction between the different types of systems, is argued, to be between declarative and nondeclarative memory. Declarative memory involves conscious recollection as opposed to nondeclarative memory

which is unconscious (Squire & Zola, 1996). Children, as young as 2 years old, are considered to have a relatively good declarative memory for emotional or unique events (Eysenck, 2015). However, progressive improvements can be seen throughout childhood, into adolescence, through to adulthood (Ofen et al., 2007). Declarative memory includes semantic memory and episodic memory (Squire & Zola, 1998). Semantic memory consists of a general knowledge of the world. For example, facts, word meanings, or information about people, whereas episodic memory is concerned with the storage and retrieval of specific events (Squire & Zola, 1998). It is constructive in nature (Hemmer & Steyvers, 2009) and can involve free-recall or recognition memory (Eysenck & Keane, 2015). Free-recall involves an independent search of memory whereas recognition memory relies on external cues (La Rooy et al., 2011). Episodic memory is closely related to autobiographical memory, both involve memories of personal experiences (Pathman et al., 2011). However, there are distinct differences. Autobiographical memories are stored for longer and relate to more salient life events than episodic memories (Eysenck & Keane, 2015).

Autobiographical Memory. The success of an investigative interview relies upon a witness' ability to remember and accurately report past events (Anderson et al., 2009). However, it is important to recognise that the route to remembering is a dynamic one, with many factors capable of affecting the flow of information within the developing memory systems of children (Ornstein & Haden, 2002). This is exemplified in Ornstein and Haden's (2002) framework for children's memories of salient, personally experienced events. The framework identifies four principles related to the encoding, storage, and retrieval of children's memories: (1) not everything gets into memory, (2) what gets into memory may vary in strength, (3) the status of information in memory changes over time, and (4) retrieval is not perfect.

Not Everything Gets into Memory. Not every detail of an event is encoded. This is because humans can only attend to and process a limited amount of information (Ornstein & Haden, 2002). Consequently, certain aspects of the event will be omitted during recall as a result of having never been represented in memory in the first place. This effect can be exacerbated by stress as demonstrated by Merritt et al. (1994). Their study examined 3- to 7-year-olds' memories for an invasive medical procedure (i.e., a voiding cystourethrogram [VCUG], a radiological procedure involving urinary bladder catheterisation). The amount of detail recalled about the VCUG was found to be negatively correlated with the level of stress experienced during the procedure. Ornstein and Haden (2002) suggest that the stressful nature of this procedure affected the children's attention and thus led to reduced encoding. However, this phenomenon is not a robust one. Under certain circumstances, increased stress can facilitate the encoding of information (Chae et al., 2018). A recent study by Chae et al. (2018) involved children, aged 3 to 5 years old, participating in a moderately distressing event (i.e., the Strange Situation Procedure) and being interviewed about their experiences. Greater distress during the event was found to result in increased memory performance. The disparate findings pertaining to the impact of stress upon memory has been attributed to methodological differences across studies (for an in-depth discussion see Marr et al., 2021). Unfortunately, due to these disparate findings it remains unclear as to whether, and under what conditions, stress acts as a facilitator or an inhibitor to memory.

Another factor that can influence the encoding of information is a child's understanding of the events (Ornstein & Haden, 2002). If a child is able to comprehend and make sense of what they are experiencing, their attention will be more focused upon the key features of the experience and thus the encoding of these features will be more complete (Ornstein et al., 1997). When asked to identify the main challenges in child sexual abuse (CSA) investigations involving young children, prosecutors were concerned that children may

have little understanding of their experiences and therefore could not properly retell them (Ernberg, 2018). Understanding can be driven by endogenous or exogenous influences. Endogenous influences are those that derive from within the individual, such as prior knowledge and expectations, of which older children will possess more. Exogenous influences, on the other hand, are external. Adults or older children can act as exogenous influences (Ornstein & Haden, 2002). As such interviewers need to be mindful that children's reports may be significantly influenced by previous conversations with and the interpretations of adults which may affect the children's understanding and thus impact upon their subsequent reports (Ornstein & Haden, 2002).

What Gets into Memory May Vary in Strength. Once information has been encoded the strength of the representation can vary (Ornstein & Haden, 2002). The strength of the representation can determine the types of questions needed to elicit the information. Representations with strong memory traces are more readily accessible. They can be obtained with minimal prompting and can easily be retrieved using open-ended questions⁴ (e.g., 'tell me everything that happened?'). Open-ended questions utilise recall memory processes. These processes enable a witness to conduct an independent search of memory (La Rooy et al., 2011) and are thought to elicit the most accurate information (Orbach & Lamb, 2001). Representations with weaker traces are more difficult to recover and can often be better retrieved using specific-closed questions (e.g., 'what was he wearing?'). Specific-closed questions specify the nature of information required and thus to some extent utilise recognition memory processes. Recognition memory processes rely upon specific memory cues and are thus more prone to error. This is because these cues may invite speculation from the child or introduce incorrect information that the child simply cedes too/selects (La Rooy et al., 2011).

⁴Psychological literature uses many different definitions / labels for question types. For continuity and ease of interpretation this thesis will use the definitions given in ABE unless otherwise stated (MoJ, 2011).

There is an abundance of research exploring the effect of different question types on the informativeness and accuracy of children's accounts. The research demonstrates clear developmental differences in how children respond to different question formats (see section 2.2.2 for further discussion). Hershkowitz et al. (2012) found that with age children become more proficient at responding to open-ended questions. This age-related effect can be attributed to developmental changes in mnemonic skills. Mnemonic skills (i.e., memory strategies such as rehearsal) are linked to information acquisition. As older children possess more developed mnemonic skills, they are able, when experiencing an event for a similar duration, to acquire more information about it. Thus, older children are expected to develop stronger memory traces than younger children (Ornstein & Haden, 2002). Given that stronger memory traces are more easily accessible, it is to be expected that older children are more apt than younger children at responding to open-ended questions. It is important that interviewers are aware of developmental differences so that they can implement strategies to mitigate them.

The Status of Information in Memory Changes Over Time. The strength of a memory trace usually decreases over time, with younger children experiencing the most rapid decay (Brainerd et al., 1985). A study by Flin et al. (1992) found that 6-year-olds recalled less information, after a delay of 6 months, than either 9-year-olds or adults. One reason memory decay is more pronounced in young children is because their initial memory trace is weaker. Although time can be considered a barrier to remembering, other influences, both internal and external, can also compromise the integrity of the initial memory representation. One such influence is the child's current knowledge. When memories begin to fade, knowledge is sometimes used to fill in the gaps. A study by Ornstein et al. (1998) found that as memory for a specific medical examination fades, children will incorporate their general knowledge of visiting the doctors into their accounts. Another influence that has the potential to jeopardise

the integrity of the initial memory representation is post-event information. Although there is a considerable body of research to suggest that exposure to inaccurate post-event information can alter / distort children's accounts, conflict has arisen regarding which underlying mechanisms are responsible for this (see section 2.1.2). In order to prevent children from incorporating incorrect post-event information from interviewers into their accounts, interviewers should rely predominantly on open-ended questions and avoid introducing information not previously mentioned by the child. Precautionary measures however should also be considered (e.g., familiarising the child with the ground rules) in case of any accidental intrusions.

Retrieval is Not Perfect. Retrieval is the final stage of remembering and is the central focus of an investigative interview. Although information may have been stored in memory it may not easily be retrieved (Ornstein & Haden, 2002). Many factors can affect the retrieval of information. One factor particularly pertinent in relation to allegations of sexual abuse is acute stress. Acute stress, at the point of retrieval, has been found to impair memory (Shields et al., 2017). Retrieval difficulties are more pronounced in younger children. This is because younger children do not possess as adequate retrieval strategies. Instead, they rely upon adults to provide retrieval cues (Lamb & Brown, 2006). Hence, why younger children are often less responsive to open-ended questions. However, it is also important to recognise that not everything that is remembered may be reported in an interview due to either fear or embarrassment (Ornstein & Haden, 2002). Furthermore, what is reported may not be representative of the initial memory representation. As memory is a constructive process, current knowledge may be used to fill in the gaps (Ornstein et al., 1998). Alternatively, the initial memory may have been altered due to post-event misinformation (Johnson et al., 1993). Again, this reiterates the importance of conducting a child-led interview, whereby the child largely dictates the flow of information. Given that children's narrative skills are not

fully developed, strategies and techniques can be used to scaffold communication. One strategy is to use information spontaneously provided by the child to cue for further information e.g., ‘you said you were in his bed, tell me more about that’ (Collins, 2012). Other techniques include building rapport, introducing the ground rules, the inclusion of a practice narrative, and providing alternative means of communication (e.g., drawing, dolls).

It is evident from Ornstein and Haden’s framework (2002) that the developing memory system is not only extremely complex, it is also extremely vulnerable to error. Despite this, young children (i.e., 3-year-olds; Baker-Ward et al., 1993) and even toddlers (Meltzoff, 1988) are capable of providing accounts of past events. Although their memories of past events are likely to decay quicker (La Rooy et al., 2011) and their accounts comprise of fewer details, the information they do recall can be just as accurate as that of older children or adults (Jack et al., 2014). However, accuracy can be compromised through inappropriate interview practices.

2.1.2 Suggestibility

Suggestibility is the extent to which social and psychological factors influence the encoding, storage, retrieval, and reporting of events (Ceci & Bruck, 1993). Four components of suggestibility have been identified: (1) interrogative suggestibility (acquiescing to misleading questions or agreeing to misinformation), (2) misinformation effects (incorporating false information into subsequent reports), (3) source misattribution (failing to remember the source of misinformation), and (4) false event creation (constructing a narrative of an event that never happened; Bruck & Melnyk, 2004). However, when considering the latter of these four components, one should be aware that there is no single cause of false memory. Different experimental paradigms (e.g., Deese-Roediger-McDermott paradigm, memory conformity) give rise to different types of erroneous recollections (Calado et al.,

2018). Although younger children are often seen as the most vulnerable to suggestion, suggestibility does not always decrease with age (Hritz et al., 2015).

Before discussing prior research in more detail, along with developmental trends, it is important to place the research into context. During the 1980s and early 1990s there were a series of highly publicised day care abuse cases (e.g., the McMartin case, the Kelly Michaels case; Lamb et al., 2011). The cases generally involved allegations supposedly arising from preschool children that they had been sexually abused by day-care workers. Many of the cases, it is argued, involved claims of Satanic or ritualistic abuse (Garven et al., 1998). These cases sparked an upsurge in research exploring children's suggestibility (Goodman et al., 2017). However, with this upsurge arose a great deal of debate and what appears to be the emergence of two factions - those who sought to demonstrate the vulnerability of children and those who sought to demonstrate their resistance to suggestion. Each faction has attempted to criticise and discredit the approach of the other in order to bolster their own position. Such critiques have centred on the ecological validity of the research, the selective interpretation of data (Lyon, 1998), and even the deliberate manipulation of interview transcripts (Cheit, 2014). Although it is beyond the scope of this thesis to delve further into such issues, it is important to recognise that they exist and to be mindful of them when interpreting the research findings. One of the greatest disparities in the early research is the methodologies. These appear to be dictated, at least in part, by how the researcher conceptualises a 'typical' investigative interview (Lyon, 1998). Some very early and influential studies adopted highly suggestive techniques (e.g., Bruck et al., 2002; Leichtman & Ceci, 1995), that it seems they believed reflected practice at the time.

It is indisputable that suggestive interview techniques can lead children to make false allegations. These techniques include asking predominantly yes/no and forced-choice questions (i.e., questions that request specific information and serve to restrict the witness'

response), repeating questions, incorporating undisclosed and potentially leading information into the interview, selective encouragement, peer pressure, and guided imagery (Bruck et al., 2002). Although in laboratory research these techniques have been shown to elicit high levels of false reports (Bruck et al., 2002; Leichtman & Ceci, 1995), Lyon (1998) contends that the techniques used in these studies are not representative of real-world interviews and go far beyond what would be considered 'typical' practice. This is not surprising given how far these techniques stray from best practice guidelines (e.g., ABE; MoJ, 2011). However, interviewers do sometimes ask leading questions (Lyon, 1998) and may misinterpret what children have said or recap their accounts incorrectly (Evans et al., 2010). If these suggestions go unchallenged or errors uncorrected, they may be incorporated into the children's future accounts. Multiple techniques are more effective in achieving acquiescence, than one technique in isolation (Bruck & Ceci, 1999; Garven et al., 1998). However, even when subjected to multiple suggestive techniques not all children will assent. This is because suggestibility can be influenced by individual differences.

Individual Differences. Despite a large body of research, it has proved difficult to draw firm conclusions regarding which characteristics make children particularly vulnerable to suggestive interview practices (Lehman et al., 2010). In 2004, Bruck and Melnyk (2004) conducted a review of the literature. They examined the relationship between children's suggestibility and demographic (socioeconomic status, gender), cognitive (intelligence, memory, Theory of Mind [ToM], distractibility, language, executive functioning, creativity), and psycho-social factors (self-concept, anxiety, social engagement, mental health, temperament, maternal attachment, parent-child relationship, parenting styles). The most promising predictors of suggestibility appeared to be language ability, creativity, maternal attachment, self-concept, and parent-child relationship. More specifically, children who are likely to be the most susceptible to suggestive interview practices are those that are more

imaginative, have less advanced language skills, poorer self-concept, less supportive parental relationships, and mothers that are insecurely attached in their romantic relationships. For the other variables either no relationship with suggestibility was found or the findings were inconsistent.

As noted above stress / anxiety was one of the variables in Bruck and Melnyk's (2004) review. The review identified a number of studies that explored the relationship between anxiety and suggestibility. The results, however, were deemed to be inconsistent. Some studies failed to show any significant relationship (e.g., Eisen et al., 2002), whilst the studies that found a significant relationship varied in terms of their direction. For example, Ridley et al. (2002) found a negative correlation between stress and suggestibility whereas Goodman et al. (1997) found a positive relationship. However, the relationship in Goodman et al. (1997) was no longer significant when additional variables were considered e.g., parental attachment. Given that an investigative interview can be an unpleasant and often stressful experience for children (Moston & Engelberg, 1992), anxiety is a topic that will be covered in more detail further on in this thesis (see section 2.2.1). Of interest is the relationship between social support, anxiety, and memory performance. The theory is that social support reduces anxiety. Anxiety is thought to impair cognitive functioning (Eysenck & Calvo, 1992) and thus lead to increased suggestibility (Almerigogna et al., 2007). This is partly because anxious thoughts expend children's limited and valuable cognitive resources (Eysenck & Calvo, 1992). Resources which could otherwise be allocated to enhanced retrieval and memory monitoring. It is likely that anxiety perpetuates at least some of the difficulties detailed below.

Cognitive Explanations. Suggestibility is closely associated with memory. Memory interference theories have been proposed to account for suggestibility effects. Memory interference theories assume "that post-event misinformation interferes with the storage and / or retrieval of event details" (Holliday et al., 2002, p.53). Holliday et al. (2002) considered

three theoretical models in their discussion of memory interference theories. The first of these is the trace-alteration model. The trace-alteration model asserts that post-event misinformation alters or overwrites the original memory trace, leaving only one trace at the point of retrieval (Loftus et al., 1985). In contrast, trace-strength and retrieval interference models propose that both the original and post-event memory traces co-exist at the time of retrieval (Lehman et al., 2010), with the post-event memory trace creating access competition. According to trace-strength models the magnitude of the suggestibility effect relates to the relative strength of the post-event and original memory trace (Holliday et al., 2002). Retrieval interference models, on the other hand, explain suggestibility purely in terms of retrieval failure.

One of the most influential versions of this model, Headed Records (Morton et al., 1985), asserts that original and post-event misinformation are represented in memory by two discrete unalterable headed records. Only one of these records can be retrieved at a time. The record with the heading that matches the given retrieval cues or is the most recent will be chosen. A further explanation of suggestibility effects is the source monitoring hypothesis. It claims that post-event misinformation is reported as a result of source misattribution errors, namely attributing the source of misinformation to the original event (Johnson et al., 1993).

Cognitive explanations have also been offered to account for the widely held belief that young children are more suggestible than older children or adults (Ceci & Bruck, 1993). Younger children possess poorer encoding skills which result in weaker memory traces (Ornstein & Haden, 2002). Given that weaker memory traces are more difficult to access, it is likely that younger children will be exposed to more questions which utilise recognition memory processes than older children. This immediately places younger children at a disadvantage as questions that utilise recognition memory are associated with higher error rates (La Rooy et al., 2011). According to trace-alteration models weaker memory traces are also problematic in that they are more susceptible to being overwritten. The contrasting view,

held by proponents of trace-strength and retrieval interference models, is that weaker traces provide a more hospitable encoding context for post-event misinformation to be admitted into memory as a co-existing trace (Ceci & Bruck, 1993). In sum, all of the models suggest that the strength of a memory trace is inversely related to the likelihood of reporting misinformation. This is in line with the findings of Marche (1999), Fivush and Hammond (1989), and Powell et al. (1999). Marche (1999), for example, found that children who had experienced an event multiple times were more resistant to suggestion, than those who had experienced the event only once, as they possessed a stronger memory trace. However, experiencing an event multiple times can prove problematic if each time the event changes slightly. The more similar an event, the more difficult it is to differentiate between each individual episode (Holliday et al., 2002). As older children possess superior source-monitoring skills (Lindsay et al., 1991), it would be expected that they would be more successful at achieving this. However, under certain circumstances this is not the case.

Sometimes young children are in fact less suggestible than older children or adults. This is known as a reverse developmental trend (Calado et al., 2018) and is exemplified by the Deese-Roediger-McDermott task. The task involves selecting a target word (e.g., sleep) along with 15 closely related words (e.g., bed, pillow, night). The 15 words are presented, the target word is not. This is then followed by a test of recognition or free-recall. A review by Brainard and Reyna (2012) found that false memory for the target word was more prevalent among older children. Older children can therefore be considered as more suggestible when undertaking this task. This can be explained using fuzzy-trace theory (Brainard & Reyna, 2012). The theory discusses two types of memory traces: verbatim and gist traces. Verbatim traces contain information about specific details, whereas gist traces contain semantic information about past events. Both verbatim and gist memories improve during childhood. According to the fuzzy-trace theory the superior gist memory of older children increases the

amount of accurate information elicited but also the likelihood of false memories, as demonstrated in Brainerd and Reyna's (2012) review. Although, it could be argued that these conclusions are not representative of a criminal event, the findings have been replicated in more naturalistic studies (e.g., Odegard et al., 2009). As demonstrated by both developmental and reverse developmental trends, cognition has a significant role in susceptibility to suggestion. However, it is not the only explanation that can account for suggestibility effects.

Social Explanations. The social characteristics of an interview can also serve to increase children's suggestibility. The perceived authority of the interviewer can perpetuate the effects of misinformation. Young children, in particular, are likely to perceive adults as trustworthy conversational partners. They may therefore fail to challenge the interviewer's suggestions and may simply acquiesce, despite knowing the information to be incorrect (Bruck & Ceci, 1999). These effects can be further exacerbated when an interviewer adopts an intimidating demeanour or exerts social pressure on a child (Garven et al., 1998). Children may also assent to the interviewer's suggestions out of a fear of getting in trouble. This may stem from negative encounters with the police in the past and / or the connotations of punishment associated with the role (Collins, 2012). However, it is not only the perceived authority of the interviewer that can influence children's accounts. If the interviewer mentions another witness' statement, the status of that witness can be influential, with young children more likely to acquiesce to incorrect suggestions when they believe the source of the information to be an adult as opposed to a peer (Carol & Compo, 2017).

Overall, interviewing officers need to be mindful of how they present themselves to child witnesses. Where possible they want to try to address the power imbalance that exists within adult-child interactions and dispel any preconceived notions or fears the child may have regarding their role. Also, in order to protect children's initial memory representations, officers need to avoid introducing misinformation into the interview and be extremely careful

when interviewing children about repeated abuse. When undertaking this difficult task officers need to encourage children to describe and focus on each individual episode in isolation. This is because officers need children to use their verbatim memory as opposed to their gist. Although gist memory can increase accurate recall, it can also result in more memory intrusions (Brainerd & Reyna, 2012). Consideration of these factors may be even more important in investigative interviews involving children who have experienced abuse or maltreatment. Although the famous day care cases (e.g., McMartin Preschool) occurred in the 1980s, similar cases have arisen more recently. One example is the Jakarta International School case, which took place in 2014. In this particular case, teachers and cleaning staff were convicted of sexual abuse solely on the basis of children's eyewitness reports. It later transpired that the allegations of abuse were most likely the result of suggestive questioning (Calado et al., 2018). This indicates that some investigative interviewers are still not applying the findings and recommendations of psychological research to their own practice. Interviews continue to fall short – a topic discussed in more detail in section 2.2.5.

2.1.3 Working memory

During an investigative interview, children will be utilising not only long-term memory (i.e., recalling the alleged incident) but also working memory. Working memory refers to the process of maintaining and manipulating (i.e., processing) information (Baddeley et al., 2014). Working memory can only store a limited amount of information for a short period of time. It can therefore be thought of as both capacity and time limited (Cowan, 2001). It can give an indication of the length of sentence or, more pertinent to an investigative interview, length of question an individual is able to process. There are different theoretical models of working memory (Baddeley et al., 2014). Baddeley's multiple component model is the most extensively studied and best-attested with young children (Boyle et al., 2013). Baddeley's model identifies four central components: the phonological loop, visuospatial

sketchpad, episodic buffer, and central executive. The phonological loop is responsible for holding, for a short period, information in a speech-based form. It is also responsible for rehearsing information to prevent it from fading. The visuospatial sketchpad involves the storage and manipulation of visual and spatial information. The episodic buffer is responsible for temporarily storing chunks of information. It integrates and holds information from the phonological loop, visuospatial sketchpad, and long-term memory. Finally, the role of the central executive is to control and direct attention, and allocate working memory resources (Baddeley et al., 2014; Eysenck & Keane, 2015).

Despite Baddeley's multiple component model being the best-attested with young children (Boyle et al., 2013), a limitation of the model is that it fails to consider how information from all the five senses is processed. The model only contains two modality-specific subsystems, related to the processing of visual (i.e., visuospatial sketchpad) and auditory (i.e., phonological loop) information. Baddeley's model does not contain equivalent subsystems responsible for the processing of olfactory, tactile, or taste information. This is despite empirical evidence having emerged in support of these additional subsystems (e.g., olfactory [Andrade & Donaldson, 2007], taste [Daniel & Katz, 2018], and tactile [Katus et al., 2015]). Given the important role of working memory in transferring information from a sensory input to long-term memory (Van der Linden, 1998) the omission of these subsystems is a major limitation of the model. In the context of child investigative interviewing, the inclusion of sensory information, including smell and taste details, in a child's statement are thought to be a good indication of credibility (Niveau, 2020) and thus may impact the progression of a case. As such it is recommended that future research examines how children process information from all five senses and potentially looks to expand upon Baddeley's multiple component model - the model is believed to have scope for additional subsystems (Andrade & Donaldson, 2007).

As with autobiographical memory and suggestibility it is important, when conducting an investigative interview, to be aware of how working memory is affected by age. Research has found that the capacity of the phonological loop, visuospatial sketchpad, and central executive all increase during childhood (Gathercole et al., 2004) leading to improved performance on working memory tasks. This, however, is not the only plausible explanation as to why working memory is affected by age (Cowan, 2001). One explanation for the poorer performance of young children on working memory tasks is that they possess less knowledge of the world. Knowledge can aid working memory as it enables individuals to link items together into meaningful patterns (also known as ‘chunking’; Miller, 1956) thus reducing the effective memory load. Another explanation, closely related to that of knowledge, is the use of mnemonic strategies including rehearsal. Research has found that younger children are less likely to employ these strategies which could potentially enhance recall (e.g., Flavell et al., 1966). An additional explanation is that of brain maturation. Brain maturation, specifically completion of the myelin layer, is thought to increase speed of processing (Hale & Jansen, 1994). Given that working memory is time-limited, the faster the information can be processed the more likely that the necessary processing will be completed before the information becomes unavailable. Further explanations are associated with developmental changes in attention. It has been proposed that younger children may be able to attend to less information at one time, be less able to maintain attention, and less proficient at inhibiting irrelevant stimuli (Cowan, 2001) – all offer plausible accounts of their poorer working memory. At present, there are lots of possibilities but no definitive explanation to account for working memory development. Regardless of which underlying mechanism/s is responsible, working memory capacity is an important consideration when conducting a child investigative interview. If a child is asked a question that exceeds their capacity, it is likely that they will be unable to fully process the question which, at best, could lead to no response and, at worst, an

unreliable or incorrect response. Given the seriousness of these repercussions, the relationship between working memory capacity and investigative interview performance is an area that needs to be explored further. To date, this area of research appears to have been neglected.

2.1.4 Attention

For information to be processed it must first be attended to (Smith et al., 2015). Attention is a cognitive function closely related to memory. It is the ability to selectively focus and sustain concentration, whilst ignoring irrelevant stimuli. As with memory, it is an acquired ability that develops with age (Anderson et al., 2009). The general standard is that children will be able to attend for 3 to 5 minutes per year of age (e.g., a 4-year-old should be able to attend for between 12 and 20 minutes; Schmitt, 1999, as cited in Anderson et al., 2009). This principal is largely reflected in the CornerHouse Forensic Interview Protocol (Anderson, 2013). Engagement is estimated to be 15 minutes for 3-year-olds, 20 to 25 minutes for 4- to 5-year-olds, 30 to 45 minutes for 6- to 10-year-olds, and up to an hour for 10- to 12-year-olds (Anderson et al., 2009). Despite offering these guidelines the protocol does stress that these are only to be used for reference and that attention span varies from child to child. Although ABE (MoJ, 2011) does not offer similar time references it does emphasise the importance of going at the witness' pace and taking breaks when required. It also discusses the permissibility of multiple interviews. There appears to be an awareness, amongst researchers and practitioners, regarding children's attentional abilities. However, how best to utilise a child's limited attentional resources is a highly contentious issue. There is debate as to whether the available time is better spent preparing the child for the interview or eliciting relevant case specific information (Anderson et al., 2009).

2.1.5 Language

Language is instrumental within an investigative interview as it enables witnesses to convey information about their past experiences. Language can be defined as a systematic

method of communicating feelings or ideas. Communication can be via sounds, signs, marks, or gestures that have an understood and defined meaning (Houwen et al., 2016).

Communication can be separated into two categories: receptive and expressive. Receptive refers to how language is understood, whereas expressive refers to how language is used (Houwen et al., 2016). Four main areas of language competence have been identified: phonology, semantics, syntax, and pragmatics. Phonology is concerned with how particular sounds (i.e., phonemes) are used in different languages, semantics refers to the meaning encoded in language, syntax is the way in which words or phrases are arranged in order to create sentences, and pragmatics is a knowledge of how language is used in different contexts (Smith et al., 2015). At present it is not fully understood how children acquire these four areas of language competence (Smith et al., 2015). Some theorists argue that children are born with an innate universal grammar that structures and constrains language development (Roeper, 2007), whereas others assert that language competence develops through cultural interactions (Tomasello, 2003). Despite disagreement as to how children acquire competence with language it is generally accepted that the ability to understand and use language improves with age (Lamb, et al., 2011), with the most rapid improvements in early childhood.

Given that the current thesis is concerned with children's communication in child investigative interviews the remainder of this section will focus predominantly on the development of language from 2 years onwards. With extensive planning and appropriate questioning, children as young as two are thought to be able to provide accurate and reliable accounts of their experiences (Marchant, 2013).

2 to 5 Years. Table 2.1, provided as part of some of Triangle's (i.e., an independent organisation, based in the UK, that provides communicative assistance to children and young people; Triangle, n.d.) training courses and taken from the Derbyshire Language Scheme (Knowles & Masidlover, 1982) gives details of language development milestones from age 2

to 5 years old. It can be seen from the table that both receptive and expressive language abilities develop rapidly over this period, with children typically more proficient at understanding as opposed to using language. At 2 years old, children's speech will often include grammatical errors (e.g., 'Mouses gone away') whereby the child is applying grammatical rules of syntax indiscriminately. At this age children also tend to use a lot of idiosyncratic language (i.e., words that they have invented), begin to use prepositions (e.g., 'in', 'on', 'under'), apply irregular verb endings (e.g., 'fought'), use 'Wh' questions, and reorder sentences in order to create questions (e.g., 'John is swimming?' becomes 'Is John swimming?') or negative statements (Smith et al., 2015). At 3 years old children are able to maintain reasonable conversations. However, these tend to be rooted in the immediate present. It is also at this age that they begin perfecting their use of pronouns (e.g., 'I' and 'we'), auxiliary verbs (e.g., 'am' in 'I am dancing'), passive verbs (e.g., 'the door was opened'), and irregular verbs (e.g., 'I knew; Smith et al., 2015). By 5 years old children can understand and produce more complex sentences and adjust their speech to the audience. However, they continue to make some logical errors and have difficulty with certain aspects of syntax (Smith et al., 2015).

Table 2.1*Language Developmental Milestones*

Age	Intelligibility	Comprehension	Receptive Vocabulary	Expressive Vocabulary	Mean Length of Utterance**
2 years	25%	1ICW*	300	50-100	1.0-2.0
2 years 6 months	60%	2ICW	50-1,000	300	2.0-2.5
3 years	75%	2-3ICW	850-1,000	600	2.5-3.0
3 years 6 months	80%	3ICW	1,200-2,000	720-1,200	3.0-3.75
4 years	90%	4ICW	2,100	850 -1,700	3.75-4.5
5 years	100%	Complex utterances	2,800	900-2,000	4.5+

*Information carrying words – the words the child needs to process in order to understand a sentence.

** The number of words the child will combine together.

5 Years +. Between the ages of 5 and 10 years old, children acquire approximately 3,000-5,000 new words each year. They begin to understand abstract vocabulary (e.g., ‘welfare’, ‘democracy’) and non-literal language. They also gain a greater appreciation of pragmatics, in that they can take the perspective of others and understand the consequences of their utterances (The PEARL Project, 2002, as cited in NHS, n.d). Between the ages of 10 and 12 children start to use meta-linguistic and meta-cognitive verbs (e.g., ‘infer’, ‘conclude’). They also become aware of the difference between factive verbs that imply certainty (e.g., ‘I know’) and non-factive verbs that imply uncertainty (e.g., ‘I believe’). At this age children also begin to understand and use sarcasm and ambiguity (The PEARL Project, 2002, as cited

in NHS, n.d.). After the age of 12, children understand and use idioms (e.g., ‘over the moon’) and develop sophisticated language abilities that can aid persuasion and negotiations (The PEARL Project, 2002 as cited in NHS n.d.).

Although developmental milestones can provide a helpful basis upon which to consider a child’s ability to communicate within an investigative interview, it is important to recognise that there is considerable variation in the rate at which children develop language competence (Adams et al., 1999). One must also be aware that children progress at different rates across the milestones. For example, a child’s receptive language may be in line with that expected for their age, yet their expressive vocabulary may be severely delayed. Hence, the importance of communication assessments (see section 2.2.6). Communication assessments can help identify whether children have any of the difficulties outlined, in the bullet points below, which have the potential to impact upon the accuracy, clarity, and coherence of the account they provide during an investigative interview:

- As mentioned previously, children do not possess as extensive a vocabulary as adults. Their vocabulary is also less descriptive and more idiosyncratic (Lamb et al., 2011). The implications of this are two-fold. For one, children may be unfamiliar with the terminology frequently used in an interview e.g., ‘allegation’, ‘witness’, ‘social worker’, ‘crime’ (Aldridge et al., 1997). Language therefore might have to be adapted in line with children’s developmental capabilities. The second implication is that children may not have the vocabulary to describe their experiences and may require additional methods of communication (e.g., drawing, body diagrams) in order to do this (Morgan et al., 2013).
- Many young children will not have acquired anatomical terms for sexual body parts (Kenny & Wurtele, 2008) and may instead use ambiguous terms such as ‘butterfly’, ‘downstairs’, and ‘minnie’ (Burrows et al., 2017). The use of these terms can make it difficult to ascertain the nature of an alleged offence. Determining the specific body parts involved in an offence

is often essential in securing a successful prosecution (Burrows et al., 2017). Interviewers therefore need to seek clarification of meaning. In order to do this, interviewers frequently ask children for an alternative term - a task beyond the capabilities of many children (Burrows et al., 2017). Research has found that questions regarding body function are far more apt at providing clarification (Burrows et al., 2017).

- Young children have difficulty understanding abstract concepts (Bruck, 2009). For example, if a child has not fully grasped the concept of 'touch' they may deny that any 'touching' has occurred but acknowledge that there was 'tickling' or 'licking' as they do not realise tickling and licking fall under the bracket of touching. This has serious implications for the child's credibility as it may appear as if the child is contradicting themselves.
- Young children have a tendency to interpret information literally (Anderson et al., 2009). As such indirect speech acts (e.g., 'do you know what happened?') can prove problematic for children. These questions directly ask if the child knows, whilst indirectly asking what they know (Evans et al., 2014). Young children may not recognise the indirect question and may reply literally and in the affirmative (i.e., 'yes').
- Young children can experience confusion when reference is made to prepositions such as 'inside', 'outside', 'under', 'behind', 'above', and 'below'. However, children are often able to understand prepositions before they are able to use them accurately in their own narratives or descriptions (Marchant, 2013). Prepositions can be particularly problematic in relation to clothing. Clothing can be neither on nor off (e.g., trousers around ankles). This is known as intermediate placement and can prove challenging for children to describe (Stolzenberg et al., 2017). Again, children may require an additional means of doing this.
- Young children can have difficulty locating events in time. They may not fully understand temporal terms such as 'before', 'after', 'first', and 'last' (Lamb, et al., 2011). Once children have grasped these concepts and possess an understanding of the number and pattern of

days in the week they can still have difficulty judging the temporal distance of autobiographical events (Hudson & Mayhew, 2011). This creates difficulties in terms of particularising and identifying individual incidents. However, research has found that children are more accurate when asked temporal distance (e.g., which did you learn a longer time ago, x or y?) as opposed to temporal location questions (e.g., which did you learn before today, x or y?; Tang et al., 2017).

- Young children have difficulty producing narrative accounts that are both well-structured and sufficiently detailed (Westcott & Kynan, 2004). Given that it is a child's ability to "tell their story" that can determine the success of a case (De Jong & Rose, 1991) interviewers need to establish a means to scaffold children's accounts.
- Young children have a tendency to drift off topic. In daily conversation, a 3.5-year-old can typically stay on topic for two conversational turns and a 5-year-old for five turns (Gotzke & Gosse, 2009). Hence, regular signposting can be crucial.

Although this is by no means an exhaustive list it gives an indication of how complex interviewing young children can be (for a further discussion see La Rooy et al., 2015). The complexity of conducting a developmentally appropriate interview is further compounded by the fact that many of the children the police interview will potentially have been subject to maltreatment. Research has shown that language and social skills are generally poorer amongst this group (Lum et al., 2018). A pre-interview communication assessment, is thus arguably, even more imperative with maltreated children as they may be operating significantly below what would be expected for a child of their age. Police officers have reported incidences of both under- and over-estimating children's abilities as a function of age (McCullough, 2017). A pre-interview assessment can overcome this as it can give an indication of an individual child's language abilities, along with an insight into their memory,

suggestibility, and attention. It can also offer an opportunity to practise scaffolding strategies / techniques.

2.2 Scaffolding Communication

Children's developmental difficulties in terms of memory, attention, and language can in part be overcome. Vygotsky's zone of proximal development (1978) discusses the concept of scaffolding. Children's cognitive abilities can be scaffolded and developed through support and interaction with adults. Scaffolds can be verbal (e.g., brief summaries of the information the child has previously provided) or non-verbal (e.g., drawings, body diagrams) and can be used to support the thinking, recall, and communication of young children (Marchant, 2013). Some researchers assert that very young children, may in fact, be incapable of producing an account without this additional support (Oxburgh et al., 2010). Nonetheless there needs to be careful planning as the inappropriate use of scaffolds can lead to the contamination of memory (Marchant, 2013; Oxburgh et al., 2010). The contamination of memory is a major concern for both researchers and practitioners. These concerns have driven over 2 decades of research into the fallibility of memory (Saywitz et al., 2015). The research has sought to identify strategies to elicit from children as complete, coherent, and accurate accounts as possible (Waterman et al., 2004). Researchers have examined the contribution of many different factors in an attempt to fulfil this objective including rapport, question types, ground rules, and communication aids.

2.2.1 Social Support and Rapport

The vast majority of best practice guidelines, including ABE (MoJ, 2011), consider rapport to be an essential component of an investigative interview (NICHD Protocol, Lamb et al., 2007; CI, Memon et al., 1997). Good rapport is believed to lower anxiety, reduce suggestibility, and improve communication (Saywitz et al., 2015). Despite these positive assertions, some researchers have made claims to the contrary. They have argued that certain

types of rapport can increase children's suggestibility (Hershkowitz, 2011) and that a protracted rapport phase can be counterproductive (Davies et al., 2000; Teoh & Lamb, 2010). It is argued that a protracted rapport phase may exhaust the limited attentional resources of children and therefore reduce their productivity in the substantive phase of the interview (this, however, could be overcome by having a short break). Despite this conflict, little research has emerged to resolve the debate. Although many studies have examined rapport in forensic interviews with adult witnesses and suspects (e.g., Alison et al., 2013; Collins et al., 2002; Collins & Carthy, 2019; Holmberg & Christianson, 2002; Huang & Teoh, 2019; Leahy-Harland & Bull, 2017; Walsh & Bull, 2012), far fewer have explored the independent effects of rapport on interview outcomes with children (Saywitz et al., 2015). As the focus of this thesis is upon scaffolding children's communication, the extensive literature pertaining to rapport in adult populations will not be considered further. However, before reviewing the relevant studies related to rapport, research concerning the related concept of social support is considered.

Social Support. Rapport needs to be considered in light of social support as, despite being distinct concepts, the two are often conflated (Saywitz et al., 2016). Some researchers see being supportive as part of developing good rapport, whilst others see developing good rapport as part of being supportive (e.g., Davis & Bottoms, 2002; Hershkowitz et al., 2006). A recent study by Hershkowitz et al. (2021) examined this relationship across repeated interviews, with higher levels of support during an initial interview found to be associated with greater rapport in a second interview. Davis and Bottoms (2002) define social support, within an interview context, as "a form of social interaction or communication that fosters a feeling of well-being in the target" (p. 186). The following behaviours are thought to be indicative of a supportive interviewer: smiling, eye contact, friendliness, using the interviewee's name, open-body posture, provision of warmth, and positive feedback (Saywitz

et al., 2016). According to Vygotsky (1978) social support can help children function at a higher level than would otherwise be possible. Research has found that social support can increase children's competency in a variety of cognitive tasks (Fischer et al., 1993). More specifically it has been shown to bolster children's memory performance in an interview context (Blasbalg et al., 2018, 2019; Brubacher, Poole, et al., 2019; Carter et al., 1996; Goodman et al., 1991; Hershkowitz et al., 2014). Social support has been found to increase both children's accuracy and resistance to suggestion (Saywitz et al., 2016). A number of explanations have been suggested to account for these findings.

The first explanation is related to anxiety reduction. It has been proposed that a supportive interviewer may have a calming effect. This in turn may reduce negative emotional states that, according to attentional control and processing efficiency theories, compete for and interfere with mental resources (Eysenck & Calvo, 1992; Eysenck et al., 2007). Although few studies have examined the complex relationship between interviewer support, anxiety, and memory performance, there is some preliminary evidence to support the theory.

Almerigogna et al. (2007) found that a supportive interviewing style not only reduced children's level of state anxiety but also increased their resistance to suggestion. Quas et al. (2014) found similar results based upon self-report measures. However, in their study the children's self-reported anxiety ratings did not correspond with their cortisol levels - a physiological measure of stress arousal. As such, the children's cortisol levels did not provide any evidence to suggest that anxiety is a mediating factor in the relationship between social support and memory performance. An earlier study by Quas and Lench (2007) also opted to use a physiological measure of anxiety (i.e., heart rate) in order to explore this relationship. They found that an increased heart rate at retrieval (indicative of a heightened stress response) was associated with poorer memory, but only in the non-supportive condition. In the supportive condition, heart rate was unrelated to memory performance. This highlights the

necessity for further research not only into the effects of a supportive interview approach but also into the effects of a non-supportive one. Davis and Bottoms' (2002) study also warrants further exploration given their somewhat conflicting findings. They found social support to decrease anxiety. However, they found no link between anxiety and report accuracy. Despite this they do not rule out anxiety as a mediating factor. Instead, they suggest that the mediation of anxiety may not have been tested very accurately due to the children having not been very anxious to begin with. This is likely to have been a difficulty inherent in all the aforementioned studies due to their experimental nature and ethical constraints.

An alternative explanation is that social support may empower children, increasing their self-efficacy and reducing the perceived power differential between themselves and the interviewer. Thus, providing the children with the confidence to contradict adult suggestions. This is known as resistance efficacy (Davis & Bottoms, 2002). Davis and Bottoms (2002) conducted a study to explore this phenomenon. They found that children interviewed in a supportive environment were more resistant to suggestion than those interviewed in a non-supportive environment. For the older children (i.e., 79 to 92 months old) in their study the effects of social support were found to be mediated by resistance efficacy. However, no evidence emerged to suggest a similar mediating effect amongst the younger children (i.e., 72 to 78 months old). This could indicate that there is another psychological mechanism underlying the effects of social support on younger children's resistance to suggestion. The authors, however, suggest that this is unlikely and instead attribute the results to measurement issues - measuring complex psychological constructs, such as resistance efficacy, can be very difficult in young children. As such, the younger children in their study may have experienced difficulties with accurately quantifying their feelings on the scale, understanding the construct of self-efficacy, and recognising the impact of the interviewer's behaviour on their own

feelings. Due to measurement difficulties, it would be premature to discount resistance efficacy as a mediating factor.

Rapport. It is really important to make the distinction between ‘psychological rapport’ and the rapport phase of an investigative interview (Collins et al., 2014). The former, ‘psychological rapport’, can be thought of as a “state of communicative alliance” (Abbe & Brandon, 2013, p. 238). It characterises the smoothness of an interaction and thus only possesses meaning as a description of a dyad or group. Tickle-Degnen and Rosenthal (1990) developed a theoretical model of rapport. The model consists of three components: mutual attentiveness, positivity, and coordination. Mutual attentiveness refers to the degree of involvement in the interaction, positivity to feelings of mutual friendliness, and co-ordination to the level of synchrony between parties. Although the three components are interrelated, they are also distinct, and independently fluctuate and can decline throughout an interaction (Tickle-Degnen & Rosenthal, 1990). The status of a relationship is thought to determine the degree to which the different components are present (Tickle-Degnen & Rosenthal, 1990). For example, positivity is likely to be present during early encounters, as individuals want to make a positive impression and develop a relationship. Positivity is then likely to decline once the relationship is established. Coordination, on the other hand, is unlikely to occur early on in the interaction, as the participants are not familiar with one another and have not adapted to each other’s communicative style. Coordination therefore would occur further on in the interaction once the participants become more accustomed to one other. In contrast, mutual attentiveness is likely to be present throughout the entire interaction. A lack of mutual attention could indicate disinterest which could potentially create a negative impression and compromise the “communicative alliance” that may have been forged. Given rapport's dynamic nature, it is argued that it should not be seen as exclusive to the rapport phase of an interview (Abbe & Brandon, 2013).

The rapport phase is the first phase of an investigative interview and may, or may, not include 'psychological rapport' (Collins et al., 2014). One purpose of the rapport phase is to help the witness understand what is expected during the interview. It can also help the witness to feel at ease with the interviewer and the interview process. According to best practice guidelines rapport should be established during the following components: an explanation of the interview's purpose, exploration of truth and lies, discussion of ground rules, and a brief conversation about neutral topics (MoJ, 2011). Interviewers, at times, fail to recognise the benefits of rapport building. This can lead to the rapport phase being conducted in a hurried and superficial manner whereby all the rapport components identified in ABE (MoJ, 2011) are present but 'psychological rapport' fails to be established (Yarbrough et al., 2013).

A number of studies have examined the effects of different rapport-building strategies on interview outcomes. The first of these studies was conducted by Roberts et al. (2004). They compared two types of rapport-building style: open-ended rapport-building which comprised of open-ended questions and direct rapport-building which comprised of specific-closed and yes/no questions. Half of the children were asked open-ended questions (e.g., 'Tell me about yourself'), whereas the other children were asked specific-closed (e.g., 'How old are you?') and yes/no questions (e.g., 'Do you have any pets at home?') about neutral topics during the rapport phase. They found that an open-ended rapport-building style enhanced the accuracy but not the informativeness of children's accounts. However, the length of the rapport-building phase may have confounded the results - on average, direct rapport-building lasted 6 minutes whilst open-ended rapport-building lasted 16 minutes. A more recent study, by Brown et al. (2013), attempted to control for this confounding variable. Their study held the length of the rapport phase constant. Brown et al. (2013) again found an open-ended rapport style to be superior. They found that an open-ended rapport style increased children's responsiveness without jeopardising their accuracy.

However, when considering the above studies, it is really important to be aware of their limitations. Both studies treat rapport as part of the first phase of the interview. However, the concept of limiting rapport to the rapport phase is very outdated. Prior to an ABE (MoJ, 2011) interview, police officers will often visit the child at home for an initial discussion. In addition, all cases with an RI will involve a pre-interview communication assessment in which the police officer should also be present (MoJ, 2020a). The assessment is considered within the Registered Intermediary Procedural Guidance Manual (2020a) as an additional opportunity to build rapport with the child. Thus, regardless of whether an RI is involved, or not, the officer will likely have spent time and established rapport with the child prior to the onset of the interview.

Importantly, the rapport that has been built should be maintained for the duration of the interview (not just during the rapport phase). Yet the aforementioned studies made no attempt to explore how rapport fluctuated. Given the changing nature of the environment, rapport behaviours should be measured at multiple time-points / intervals throughout the interview (Johnston et al., 2019). Furthermore, both studies used informativeness and accuracy as outcome measures. Although these outcome measures are essential if the results are to be applied to a forensic context, it is impossible to determine, with any certainty, whether rapport is in fact responsible for any of the observed effects. The enhanced accuracy (Roberts et al., 2004) and informativeness (Brown et al., 2013) reported in the studies above, may simply be the result of practise in retrieving information from memory and responding to open-ended questions. In order to determine whether rapport does impact upon communication a direct measure of rapport needs to be used. Direct measures include non-verbal indicators of 'psychological rapport' (e.g., expressivity, synchronicity, and interpersonal distance; Bernieri et al., 1996) or the participant's subjective experience.

Practice Narrative. A practice narrative is a structured discussion regarding a non-allegation related event (e.g., birthday, recent fun activity) prior to the substantive phase of the interview (Price et al., 2013). It has a number of benefits. A practice narrative is an opportunity to increase rapport and establish a desired conversational pattern, whereby the child becomes accustomed to what is expected from their communication and understands their role as expert (Anderson et al., 2014). It also enables children to practise remembering, retrieving specific details, and responding to open-ended questions (Roberts et al., 2011). Furthermore, it can offer another context in which to teach and practise the ground rules (Brubacher et al., 2015). However, the beneficial effects are not exclusive to the child or witness, a practice narrative can also prove beneficial for the investigative interviewer. It offers the interviewer an opportunity to assess the witness' communication, practise asking open-ended questions, and eliciting episodic information (Brubacher et al., 2011). Despite the many perceived benefits of a practice narrative, McCullough (2017) found that police officers do not always appreciate the value of practice narratives. Prior to specialist training in pre-interview assessments, police officers described the practice narrative as a general chat as opposed to an in-depth exploration of a single event. Concerns have also been raised regarding its use. As with the implementation of ground rules, one of the biggest concerns is that of time. It is feared that following a lengthy practice narrative children will be too fatigued to make a full disclosure. However, research has shown that it only takes approximately 2 minutes of practise to see benefits (Whiting & Price, 2017). Concerns have also been raised, by practitioners, regarding identifying a suitable target event (Roberts et al., 2011). Although Roberts et al. (2011) acknowledge these issues, they assert that the benefits of a practice narrative far outweigh any concerns.

There is limited research on the efficacy of practice narratives. However, the research that has been conducted has generally yielded positive findings (Anderson et al., 2014; Brown

et al., 2013; Brubacher et al., 2011; Price et al., 2013; Whiting & Price, 2017). Price et al. (2013) found that the inclusion of a practice narrative can potentially improve the quality of an investigative interview both in terms of children's communicative performance and the interviewers' behaviour. When the substantive phase was preceded by a practice narrative interviewers asked fewer questions, of which a greater proportion were open-ended. Children responded to these questions in more detail than when no practice narrative was conducted. These beneficial effects were enhanced when the practice narrative was conducted in accordance with best practice guidelines. This is in line with the findings of Anderson et al. (2014) who found that when interviewers asked more open-ended questions during the practice narrative, children aged 3 to 18 years old provided more detailed accounts of abuse at interview.

However, research by Brubacher et al. (2011) suggests that children of different ages may benefit from different types of practice narrative, particularly if there is a suspicion that they have been the victim of multiple incidents of abuse. Brubacher et al. (2011) compared three types of practice narrative: generic recall (i.e., identified a repeated event and children described what usually happens), incident specific recall (i.e., identified a repeated event and children described the time they remembered the best), and novel recall (i.e., identified a unique event). They found that incident specific recall was most beneficial for 5- and 6-year-old children, who had experienced a repeated event. These children were more likely to spontaneously disclose that the event had occurred on more than one occasion, provide more target information, and identify more differences across occurrences. Although such benefits were not observed for older children (7- and 8-year-olds), children of both age groups were found to use more episodic language during the substantive phase if they had practised episodic recall (i.e., incident specific and novel). Furthermore, no negative effects of incident specific recall were observed for children who had only experienced a single event. These

findings have significant implications for practice. Approximately 50% of child complainants allege repeated abuse (Roberts et al., 2011). In order to attain a conviction of repeated abuse, individual offences need to be 'particularised' - each act with which the suspect is charged needs to be identified with reasonable precision. Events can be 'particularised' by reference to a time or unique contextual detail (Guadagno et al., 2006). Thus, a technique such as incident specific practice, which has been found to increase the disclosure of such details, could prove highly beneficial in cases where repeated abuse is suspected. This is even more salient given that, irrespective of accuracy, reports of repeated-events are often perceived as less credible than reports of single-events (Connolly et al., 2008). Anything that can be done to enhance perceived credibility (where justified) warrants further investigation.

Further research, by Danby, Brubacher et al. (2017), has explored whether children benefit from practise narrating two as opposed to one episode of a repeated event. This was achieved through engaging children (5- to 9-year-olds) in four similar classroom-based activities and later interviewing them about their experiences. Prior to the interviews, the children provided a practice narrative about either one or two episodes of an autobiographical repeated event. The older children who practised recalling two episodes were found to report a greater number of episodic details during the interview than those who only recalled one episode. However, no benefits of a second practice narrative were observed for younger children. This was attributed to younger children possessing less advanced cognitive abilities, such as source monitoring (the ability to accurately identify where information, including memories and knowledge, originated; Johnson et al., 1993), and therefore being less receptive to the subtle differences between the practice conditions. Danby, Brubacher et al.'s (2017) study suggests that a practice narrative about two episodes of a repeated event may enable older children to recall more substantive information. The authors conclude that should two-episode practice narratives be unfeasible or the child too young, practise narrating one episode

of a repeated event should be sufficient, given that it primes many of the same cognitive skills.

Despite a number of studies demonstrating the utility of practice narratives, there is evidence to the contrary. For example, Hardy and Van Leeuwen (2004) found young children, aged 3 to 5.5 years old, to be less accurate when the rapport phase included a discussion of a specific past event (i.e., the child's last birthday), as opposed to a general conversation (i.e., how to play the child's favourite game). However, the authors, themselves, acknowledge that there is reason to question the validity of this finding. By opting to use a story as the to-be-remembered event, the children who talked about a general event during the rapport phase may have been at an advantage. Although the to-be-remembered event was unique, the children's recall may have been enhanced by activation for the schema of story scripts (Hardy & Van Leeuwen, 2004). The authors suggest replicating the study with a to-be-remembered event that does not follow a narrative sequence reminiscent of a story.

A more recent study by Otgaar et al. (2016) similarly yielded disappointing results. They found that practice narratives reduced the completeness of young children's (6- to 7-year-olds) accounts, with no impact upon accuracy. Otgaar et al. (2016) acknowledge that the absence of any beneficial effects may have been due to the nature of practice. It may have been that the instruction was not well-selected. The children were asked questions about an unrelated neutral event, the example given is 'their last vacation'. From this information, it is impossible to determine whether children were recalling novel or repeated events. If the events were repeated, it may have been that the children were relying on a script and thus providing generic information. As the finding from Brubacher et al.'s (2011) study suggests this does not elicit the most detailed information from children. Although the mechanisms underlying the discrepancies between studies is unknown, the type of practice and nature of the to-be-remembered event both offer plausible explanations. Overall, there does appear to

be a number of beneficial effects of a practice narrative. As such, ABE (MoJ, 2011) suggests that, particularly with younger children and witnesses with learning difficulties, it can be helpful to conduct a practice narrative prior to the interview.

Incorporating Social Support and Rapport Throughout an Interview. It is evident from the aforementioned studies that the bulk of research in this area has focused upon establishing rapport at the outset of interviews. Little attention has been given to maintaining rapport throughout the interviews' duration for research with children. However, this was addressed in a recent study by Sauerland et al., (2018) that looked at the impact of rapport on the quality and quantity of children's, adolescents', and adults' memory reports. First, the participants watched a video of a mock crime and were allocated to one of three rapport conditions (i.e., none, minimal, extensive). Interviewer behaviours in the no rapport condition included a formal / neutral greeting, a closed posture, no verbal feedback, and no personal questions. In the minimal rapport condition, the interviewer did adopt a neutral / open posture and asked a number of personal questions, but the interviewer did not introduce themselves or address the interviewee by name. In the extensive rapport condition, the interviewer adopted an open / engaging posture, introduced themselves, addressed interviewees by their name, provided non-verbal feedback, and asked personal questions. Following the rapport building the participants were asked to provide a narrative account of what they had seen. This was followed by 18 questions (specific-closed, yes/no, and forced-choice). Overall, rapport was found to have very little impact upon memory performance. Only the adolescents were found to reap any benefit. Adolescents who received extensive rapport provided more accurate details during the free report phase than those in the no rapport condition. The authors attributed these disappointing results to the choice of control group. Had the control group been negative rapport as opposed to neutral, a greater impact upon memory performance may have been observed.

It is perhaps unsurprising that the majority of research has focused upon establishing, as opposed to maintaining rapport, given that ABE (MoJ, 2011) refers to rapport as a ‘phase’. However, in recent years interview protocols have begun to adapt and recognise the importance of continued support, particularly with children who are reluctant to disclose. Child abuse victims often fall under this remit (Blasbalg et al., 2018). For example, the standard NICHD Protocol (Lamb et al., 2007) has now been revised to incorporate more social support throughout the interview. Within the Revised NICHD Protocol (Hershkowitz et al., 2014) interviewers are encouraged to address children by their name; express interest and care in children’s experiences; echo, acknowledge, and explore children’s feelings; encourage children, both verbally and nonverbally, to describe experienced events; provide thanks, appreciation, and positive reinforcement of children’s efforts throughout the interview; and empathise with children’s difficulties associated with the interview experience. Further interviews (two – three) may be conducted if the interviewer believes that better rapport can be built with the child and as a consequence of this more information elicited to help determine whether abuse has occurred. Additional interviews should continue to follow the Protocol with reference to the former interview/s to build upon previously established rapport.

A number of studies have examined the efficacy of the Revised NICHD Protocol (Hershkowitz et al., 2014). A recent study by Blasbalg et al. (2018) examined the impact of interviewer support on children’s reluctance and production of information. The study involved 200 interviews with suspected victims (6 to 14 years old) of physical abuse. All interviews followed the Revised NICHD Protocol (Hershkowitz et al., 2014), and all cases were corroborated by external evidence. Interviewer support was found to reduce child reluctance and increase informativeness. Positive findings were also reported by Blasbalg et al. (2019) and Karni-Visel et al. (2019). Blasbalg et al. (2019) compared 166 interviews using the Revised NICHD Protocol (Hershkowitz et al., 2014) and 88 using the Standard NICHD

Protocol (Lamb et al., 2007). Blasbalg et al. (2019) found that interviews following the Revised Protocol involved better questioning and interviewer support along with decreased reluctance and increased informativeness on behalf of the children. Karni-Visel et al. (2019) similarly compared interviews using the two versions of the NICHD Protocol. Children interviewed using the Revised NICHD Protocol (Hershkowitz et al., 2014) were more emotionally expressive than those interviewed using the Standard Protocol (Lamb et al., 2007). Greater emotional expressivity was found to be associated with increased informativeness. Further research has also shown that children interviewed using the Revised as opposed to Standard Protocol are more likely to make allegations of abuse (Hershkowitz et al., 2014; Hershkowitz & Lamb, 2020). A 14.3% increase in the odds of a child making an allegation were reported by Hershkowitz and Lamb (2020). Taken together these findings suggest that the Revised NICHD Protocol (Hershkowitz et al., 2014) may encourage reluctant children to disclose abuse.

Despite the Revised NICHD Protocol (Hershkowitz et al., 2014) possessing a sound theoretical basis and some evidence to support its use, ultimately its utility is reliant upon the competency of those employing it. For example, Ahern et al. (2014) found that when using the Revised NICHD Protocol (Hershkowitz et al., 2014) interviewers provided proportionally more support than when using the Standard NICHD Protocol (Lamb et al., 2007). However, the support provided was indiscriminate. The interviewers were not receptive to children's reluctance and did not always provide additional support in response to this. This is in line with previous research that has found that interviewers often respond less supportively as opposed to more supportively when children are uncooperative, heightening the children's resistance further (Hershkowitz et al., 2006). Nevertheless, Ahern et al. (2014) found that when reluctant utterances were met with appropriate support, in the Revised Protocol interviews (Hershkowitz et al., 2014), immediate cooperation often followed. The findings of

Ahern et al. (2014) demonstrate the utility of social support with children who are uncooperative or reluctant to disclose.

2.2.2 Question Types

How children are questioned has a profound impact upon the detail, accuracy, and coherence of their narrative accounts (Brown & Lamb, 2015). ABE (MoJ, 2011) identifies five question types: open-ended (i.e., allows for an unrestricted response), specific-closed (i.e., specifies the type of information required), forced-choice (i.e., provides a small number of response options to choose from), multiple (i.e., asks for multiple pieces of information at once), and leading (i.e., implies desired response). It advocates commencing the interview with open-ended questions and then proceeding, if necessary, to specific-closed questions. Forced-choice, multiple, and leading questions, if possible, should be avoided. The five question types vary in their ability to elicit accurate information from children. Their effectiveness is thought to be correlated to the type of memory process they utilise. Open-ended questions utilise free recall processes. These processes enable an interviewee to conduct an independent search of memory (La Rooy et al., 2011) and are thought to elicit the most accurate information (Orbach & Lamb, 2001). Forced-choice questions, on the other hand, utilise recognition memory. Recognition memory processes rely upon specific memory cues and are more prone to error (La Rooy et al., 2011). Specific-closed questions lie somewhere between free recall and recognition memory processes on the memory continuum (La Rooy et al., 2011).

Open-Ended Questions. Given their relative position on this continuum, it is perhaps unsurprising that open-ended questions have been deemed the best for use within an investigative interview (MoJ, 2011). Open-ended questions can be further categorised into open-ended invitations, open-ended depth, open-ended breadth, and facilitators (Triangle, 2015). Open-ended invitations initiate new topics by inviting a comprehensive account. These

questions do not specify what information is required (e.g., 'Tell me everything that happened'). Open-ended breadth questions prompt a witness to recall additional parts of an event (e.g., 'Then what happened?'). In contrast open-ended depth questions prompt a witness to provide more information about part of an event that has previously been mentioned (e.g., 'Tell me more about the part where...'). Facilitators are neutral and demonstrate active listening (e.g., 'uhuh', 'go on', and 'tell me more'; Triangle, 2015).

Despite this detailed categorisation of open-ended questions, the majority of studies have analysed open-ended questions as a homogenous group. Research into open-ended questions has found them to elicit more accurate (Brown et al., 2013) and detailed accounts (Sternberg et al., 1996), than either specific-closed or forced-choice questions. They have also been shown to create fewer inconsistencies (Lamb & Fauchier, 2001) and enable children to provide more coherent narratives (Feltis et al., 2010). Children are also less likely to guess when presented with an open-ended question that they do not know the answer to, compared to a forced-choice question (Waterman et al., 2004). A further benefit over forced-choice questions is that open-ended questions make children (i.e., 7- to 12-year-olds) feel more listened to and better able to provide their stories (Brubacher, Timms et al., 2019). Although open-ended questions have been found to be beneficial with children as young as 4 years old (Sternberg, Lamb, Orbach, et al., 2001), the ability of children to respond competently to these questions has been found to improve with age (Hershowitz et al., 2012). This is because older children tend to have stronger memory traces and possess more effective retrieval strategies than younger children. Weaker traces are more difficult to access and often require specific memory cues (Ornstein & Haden, 2002).

A recent study, by Danby, Sharman et al. (2017) explored the differential effects of two subtypes of open-ended questions (i.e., open-ended breadth and open-ended depth) on children's (5- to 9-year-olds) recall of individual episodes of a repeated event. Open-ended

breadth questions tended to elicit from children episodic information that was common across all episodes. Open-ended depth questions, on the other hand, were found to be more effective in eliciting specific details about individual episodes. Open-ended depth questions were also found to be superior in terms of source accuracy (i.e., correctly linking co-occurring details together). This is attributed to open-ended depth questions containing pre-disclosed details that serve to scaffold children's recall by focusing their attention (Orbach & Lamb, 2000). This is supported by the research of Gagnon and Cyr (2017) which found that open-ended questions which utilise cues previously mentioned by the child (i.e., open-ended depth questions) elicit more informative responses than open-ended questions that are absent such cues.

Specific-Closed Questions. Specific-closed questions are deemed the “second-best type of question” (MoJ, 2011, p.78). These questions specify the nature of information required and should be used only when a free narrative account is exhausted. Examples of specific-closed questions include: ‘what’, ‘where’, ‘when’, ‘why’, and ‘who’ questions (MoJ, 2011). These are commonly known as the 5WH questions. Although ABE (MoJ, 2011) classifies these as specific-closed questions, this is not a universally held opinion (Oxburgh et al., 2010). Some researchers argue that, under certain circumstances, these questions are in fact open-ended (Phillips et al., 2012). Regardless of their classification these questions are appropriate for use within an investigative interview (Oxburgh et al., 2010) and have been found, in some instances, to be superior to very broad open-ended questions in eliciting a detailed narrative account (Hershkowitz et al., 2012). A study by Hershkowitz et al. (2012) found that young children (3- to 4-year-olds) responded more informatively to the 5WH questions than to more open-ended questions. This is attributed to these questions seeking more specific information and thus demanding less retrieval effort. This may suit younger children as they tend to employ less effective retrieval strategies.

Forced-Choice Questions. Forced-choice questions should “only be used as a last resort” (MoJ, 2011, p. 80). These questions present the witness with a small number of alternatives to choose from and may not include the correct response (MoJ, 2011). The guidance therefore suggests that when employing such questions, children should be reminded of the permissibility of an alternative or a ‘don’t know’ response (MoJ, 2011; Rocha et al., 2013). For example, ‘was it winter, or spring, or something else?’ – these are known as open-choice questions.

When presented with a forced-choice question, children may exhibit a response bias. Some studies have shown that when asked a two-option forced-choice question (e.g., ‘Was it winter or spring?’) children have a recency tendency, whereby they have a propensity to choose the last option (Mehrani & Peterson, 2015; Mehrani & Peterson, 2017; Rocha et al., 2013). This recency tendency has been found to be more pronounced in younger children (i.e., children younger than 5 years old) and when questions pertain to unfamiliar objects (Mehrani & Peterson, 2017). Nevertheless, with any child these questions should be used only with extreme caution.

Open-choice questions have undergone far less exploration than two-option forced-choice questions. To the author’s knowledge, only two studies to date have looked at the effect of open-choice questions, on the recall of young children (e.g., London et al., 2017; Stolzenberg et al., 2017). A study, by London et al. (2017), compared the accuracy of children’s (3- to 5-year-olds) responses to standard two-option forced-choice questions and open-choice questions (i.e., the same questions with an additional ‘something else’ alternative). The children were presented with three question types: false (i.e., no correct alternative present), true (i.e., correct alternative present), and unanswerable questions. Although the inclusion of the ‘something else’ alternative was found to have no impact upon the accuracy of children’s responses to either true (61% standard vs. 54% open-choice) or

unanswerable questions (23% standard vs. 30% open-choice), it was shown to increase the number of accurate responses to false questions (15% standard vs. 31% open-choice). Regardless of the ‘something else’ alternative, children provided high rates of incorrect responses to both false and unanswerable questions. Furthermore, despite performance improving with age, accuracy remained low amongst all age groups. Thus, suggesting that open-choice questions do not offer a viable alternative or solution to forced-choice questions and should therefore be employed with similar caution.

Yes/No Questions. Although not referred to specifically in the ABE guidance (MoJ, 2011) it is pertinent to draw attention here to another type of question - yes/no. Evidence has emerged to suggest that a similar bias exists with yes/no questions. Research has found that very young children (2- to 3-year-olds) can exhibit a ‘yes’ bias (Fritzley & Lee, 2003; Fritzley et al., 2013; Moriguchi et al., 2008; Okanda & Itakura, 2010; Peterson & Grant, 2001). Children’s ‘yes’ bias has been found to become weaker with age (Okanda & Itakura, 2010), suggesting that developmental factors may be responsible. Evidence has emerged to suggest that children’s cognitive development could account for a ‘yes’ bias. There are two cognitive abilities associated with this bias: verbal ability and inhibitory control (i.e., capacity to inhibit thought processes / actions that are irrelevant to the current task; Moriguchi et al., 2008). Moriguchi et al. (2008) found that children who had not developed inhibitory control were more likely to display a ‘yes’ bias. This was found to be compounded by a low verbal ability. Although the ‘yes’ bias becomes weaker with age, other biases can develop (Fritzley & Lee, 2003; Fritzley et al., 2013). Four- to Five-year-olds have been found to possess a nay-saying bias to incomprehensible questions (e.g., ‘Did I twireno the ball?’; Fritzley & Lee, 2003; Fritzley et al., 2013). However, Fritzley et al. (2013) acknowledge that this may not be a genuine bias but that the children, in their study, may have realised that the incomprehensible questions did not contain real words. Thus, they rejected the question by providing a ‘no’

response. Nevertheless, this has important implications for practice. Investigative interviews may contain words that children have not come across in everyday discourse and are therefore unfamiliar with – ‘no’ may not reflect an appropriate response (i.e., ‘Do you want to make an allegation?’).

The aforementioned studies suggest that children may possess some form of response bias. However, some researchers have suggested that this pattern of responses is instead the result of a compliance tendency (Mehrani, 2011; Mehrani & Peterson, 2017). Research has found that children are in fact influenced by the syntactic properties of questions and will respond in the direction implied by the question - ‘yes’ to positively worded questions and ‘no’ to negatively worded questions. As such the types of questions asked in the previous studies (e.g., Fritzley & Lee, 2003; Moriguchi et al., 2008) may have fostered an affirmation bias, rather than the bias having been inherent within the children themselves. In line with the previous research into response biases, Mehrani and Peterson (2017) found that children’s compliance tendency grows weaker with age. When considering the utility of yes/no questions it is immaterial whether children possess a response bias or a compliance tendency each demonstrates the potential risks of posing such questions to young children.

Multiple Questions. Multiple questions should be avoided with child witnesses. Multiple questions request multiple pieces of information at once. These questions can be explicit, for example “Did you see him? Where was he? What was he wearing?” (MoJ, 2011, p. 80). They can also be much more subtle, for example “What did they look like?” (MoJ, 2011, p. 80). Multiple questions have been found to compromise the accuracy (Carter et al., 1996; Perry et al., 1995), richness, and quality of children’s accounts (Katz & Hershkowitz, 2012). Katz and Hershkowitz (2012) found that 24% of children’s responses to multiple questions were unintelligible. They also found that when substantive answers were provided, children tended to only respond to the final part of the question. Difficulties with multiple

questions are not exclusive to young children. Perry et al. (1995) found that children, adolescents, and young adults alike, struggle to provide correct answers to multiple questions with correct answers never exceeding 35%. The authors deemed multiple questions more problematic than other complex questions (e.g., questions containing negatives, double negatives, or difficult vocabulary). However, it is important to note that questions containing negatives, double negatives, and difficult vocabulary also presented significant problems.

Leading Questions. Leading questions should also be used “only as a last resort” (MoJ, 2011, p. 78). These questions imply answers and assume facts that are potentially in dispute. A question can be leading based on the question’s structure (e.g., tag questions), the tone in which the question is asked, or the question's context (e.g., the question may incorporate previously undisclosed information; MoJ, 2011). In line with this definition leading questions can be formulated as open-ended questions, specific-closed questions, or forced-choice questions (Brown et al., 2013). A tag question is a declarative statement followed by a question tag inviting confirmation. There are two types of tag questions: positive and negative tags. A positive tag question comprises of a positive declarative statement followed by a negative tag (e.g., ‘It is for reading, isn't it?’). In contrast a negative tag question comprises of a negative declarative statement followed by a positive tag (e.g., ‘It isn’t for reading, is it?’). Both types of tag questions require either a ‘yes’ or ‘no’ response (Behzadnia & Mehrani, 2017). As with yes/no questions young children have been found to display a strong ‘yes’ bias in response to positive tag questions. Again, this bias has been shown to become weaker with age (Behzadnia & Mehrani, 2017). However, to date no research has been conducted examining whether children possess a response bias to negative tag questions. Further research into leading questions has found that they are more likely to elicit contradictions than non-leading questions (Andrews et al., 2015) and have the potential to foster seemingly credible false accounts. These false accounts can contain detailed

information and descriptions of objects, subjects, actions, and locations (Hughes-Scholes & Powell, 2008). Thus, extreme caution needs to be given when asking leading questions, particularly with young children (Marchant, 2016).

Repeating Questions. ABE (MoJ, 2011) discourages the use of repeated questions. However, question repetition is prevalent in investigative interviews (Krähenbühl et al., 2010). Krähenbühl et al., (2010) found question repetition in almost all (98%) of the MOGP (Home Office, 1992) interviews analysed in their study, with over 25% of all the questions asked repeated. Interviewers may repeat questions in order to clarify information previously given by the child, to make clear requests, or to encourage children who are reluctant to disclose. Of concern is that questions are most often repeated in a forced-choice, yes/no, or leading format (Andrews & Lamb, 2014). Whilst repeating open-ended questions can lead to children providing additional information (Memon & Vartoukian, 1996), repeating leading questions can be considered as problematic as children may alter aspects of their account (Andrews & Lamb, 2014).

Research has found that children are susceptible to shifting (i.e., changing their response) following repeated questioning (Howie et al., 2004; Krähenbühl & Blades, 2006, 2009; Krähenbühl et al., 2010). Krähenbühl et al. (2010) found that children, aged 4 to 9 years old, changed over a quarter of their responses when asked repeated questions (about a staged event). Although the youngest children (aged 4 to 5 years old) were the most vulnerable to shifting, changing their response to 40.5% of the repeated questions, shifting was still high amongst the oldest children (aged 8 to 9 years old) at 17.8%. Very few of the shifts were desirable (i.e., 1.4%, from an inaccurate to an accurate response), most shifts were classified as either undesirable (i.e., 16.2%, from an accurate to an inaccurate response) or novel inaccurate (i.e., 82.4%, from an inaccurate to a different inaccurate response), demonstrating the potentially damaging effects of question repetition. Studies have found that repeating

unanswerable questions (i.e., where the correct response is ‘I don’t know’) can have a particularly deleterious effect on accuracy (Krähenbühl & Blades, 2006; Krähenbühl et al., 2010). This can be further exacerbated when repeating an unanswerable specific-closed question in a yes/no or forced choice format (Krähenbühl & Blades, 2006). It has been proposed that children may alter their initial response in the belief that it was incorrect or as an attempt to please the interviewer (Howie et al., 2004). This can be exacerbated by children’s limited understanding of the interviewer’s role, by children’s inflated beliefs regarding the interviewer’s knowledge, and by an unequal power balance between the interviewer and child (Howie et al., 2004).

The Utility of Different Question Types. Interpreting the research pertaining to questions types can be somewhat challenging given the different coding systems and interchangeable labels used. For example, open-ended questions are also referred to, in the literature, as free report (e.g., Aldridge & Cameron, 1999) or invitation questions (e.g., Brown et al., 2013); specific-closed questions are referred to as directive (e.g., Lamb & Fauchier, 2001) or probing questions (e.g., Korkman et al., 2006); forced-choice questions are referred to as option-posing (e.g., Lamb et al., 2000) or specific questions (e.g., Aldridge & Cameron, 1999); leading questions are referred to as suggestive questions (e.g., Cederborg et al., 2000); and multiple questions are referred to as marathon questions (e.g., Shepherd, 2007 as cited in Oxburgh et al., 2010).

Further difficulties arise in that some studies appear to have coded indirect speech acts as open-ended questions (Westcott et al., 2006). Examples of indirect speech acts include ‘Can you tell me what happened?’, ‘Do you know what happened?’ These questions directly ask if the child knows, whilst indirectly asking what they know (Evans et al., 2014). These speech acts can prove problematic for young children. Young children may not recognise the indirect question and may simply reply in the affirmative (i.e., ‘yes’). Thus, this cannot be

considered a well-constructed open-ended question, if it can in fact be considered an open-ended question at all. Other examples of poorly constructed questions include those that are unnecessarily wordy, contain difficult vocabulary, or include overly complex concepts (Powell & Guadagno, 2008).

There is a tendency amongst researchers to code questions based purely on their typology, as opposed to their utility and appropriateness. This narrow approach raises a number of issues when evaluating the performance of investigative interviewers and the utility of training packages (Powell & Guadagno, 2008) - interviewers may be asking open-ended yet developmentally inappropriate questions. Conversely interviewers may ask a number of forced-choice questions. Although in isolation these questions might be seen as problematic their use might be entirely appropriate within a particular context. Although context (i.e., what the witness has previously disclosed; Dodier & Denault, 2018) is paramount in assessing question usage (Griffiths & Milne, 2006), many studies neglect to consider this factor and tend to evaluate interviews by simply counting each type of question. This approach can provide a somewhat misleading representation of interview quality. This was exemplified in a recent study by Waterhouse et al. (2019). The researchers sequentially mapped two 'good' and two 'poor' child investigative interviews (based on interviewer question type). Both the 'good' and 'poor' interviews included practices, such as using forced-choice questions early in the interview, which are discouraged in best practice guidelines. These practices would not have been identified had the proportion of questions been considered irrespective of context.

A final consideration when interpreting the above findings is that the researchers often coded transcripts, without the accompanying DVD footage. This is problematic as the transcripts can only be interpreted at a very literal level, non-verbal behaviour and tone of voice cannot be analysed (Phillips et al., 2012). For example, a question could, on paper,

appear as non-leading, yet through the strategic use of intonation an interviewer may convey the desired response to the witness (e.g., 'Was it *winter*, spring, or don't you know?'). Despite these challenges there appears to be a general consensus amongst researchers and practitioners regarding the superiority of open-ended questions. These questions, along with specific-closed questions are deemed as appropriate for use within investigative interviews (Oxburgh et al., 2010).

2.2.3 Ground Rules

The purpose of an investigative interview is to elicit as accurate and detailed an account as possible. This relies not only on the competency of interviewers but also on the ability of children to draw attention to any misunderstandings (Malloy et al., 2015). Failure to do this could potentially jeopardise the credibility and accuracy of children's accounts. Unfortunately, this failure on the behalf of children is widely documented. Research has shown that children will often answer questions that they do not understand (Waterman et al., 2000), have a propensity to guess when they do not know the answer (Rohwer et al., 2012), and seldom correct interviewers' errors (Evans et al., 2010). This has led to many interview protocols and best practice guidelines incorporating ground rules (e.g., ABE, MoJ, 2011; CornerHouse Forensic Interview Protocol, Anderson, 2013; NICHD Protocol, Lamb et al., 2007) - instructions associated with the communication expectations of the interview (Danby et al., 2015). Although there is a general consensus that ground rules should be mentioned there is little agreement, in the guidance, as to the number, placement, and format of the rules (Brubacher et al., 2015). ABE (MoJ, 2011) recommends the early introduction of ground rules. The guidance states that witnesses should be made aware of the permissibility of giving an 'I don't know' response. Witnesses should also be encouraged to say if they do not understand and correct the interviewer if the interviewer makes a mistake (i.e., misunderstands what has been said or summarises information incorrectly). In addition, with

vulnerable witnesses the guidance recommends that they are informed of the interviewer's naivety (i.e., that the interviewer was not present at the incident and therefore does not know what happened) and thus encouraged to provide as much information as possible. Despite international guidance advocating the use of ground rules there are sceptics. The sceptics argue that ground rules are too abstract and developmentally inappropriate for very young children (Geddie et al., 2001).

Ground Rules and Child Development. Theoretically, ground rules may be less effective with younger children due to younger children possessing less advanced cognitive skills. A number of cognitive skills have been identified that could potentially impact upon the extent to which children benefit from ground rules. The first is ToM. ToM is the ability to understand the mental states of the self and others (Wellman & Liu, 2004). Brubacher et al. (2015) identified two components of ToM which are thought to be instrumental in the comprehension and acquisition of ground rules. The first is knowledge access. This refers to an understanding of how individuals acquire knowledge and who has access to that knowledge. The second is false belief. This is the ability to understand that others can hold beliefs that are incorrect and contrary to reality. These two skills typically develop between the ages of 4 and 6 years old, with knowledge access often emerging first (Wellman & Liu, 2004). Without an understanding of knowledge access, children may be unable to comprehend the rule associated with interviewer naivety. Thus, eliminating any beneficial effects. The rule associated with correcting interviewer errors may be similarly redundant, with children who have not developed an awareness that others can hold false beliefs (Brubacher et al., 2015). Preliminary evidence has emerged in support of this theory. Dickinson et al. (2015) found improvements in children's comprehension of ground rules between the ages of 4 and 6 years old. This corresponds with the age at which children develop ToM, thus indicating that ToM may potentially be intrinsic to the understanding of

ground rules. Comprehension of these rules is also thought to be dependent upon the development of complex metacognitive skills (Brubacher et al., 2015). Although, the comprehension of a rule is essential in its utility, the beneficial effects of ground rules go beyond comprehension. For example, some children may be able to understand ground rules but may be unable to implement them during an investigative interview due to insufficient executive skills (i.e., working memory and inhibitory control; Brubacher et al., 2015). Overall, developmental theories suggest that there are a number of potential challenges associated with the acquisition of ground rules. A recent study, by Brown et al. (2019), found that children's ability to apply the ground rules correctly during an interview improved with developmental level. Age differences therefore may account for why research studies have yielded mixed findings relating to the efficacy of ground rules.

'I Don't Understand'. Both field and experimental studies have shown that children rarely make requests for clarification (Carter et al., 1996; Malloy et al., 2015). A number of explanations have been proposed to account for these findings. The first is that children do not seek clarification as they do not experience any comprehension difficulties (Malloy et al., 2015). Malloy et al. (2015) acknowledge that this is highly unlikely given the complex questions frequently used in investigative interviews (Evans et al., 2010). The credibility of this explanation is further jeopardised by the finding that requests for clarification increase with age (Malloy et al., 2015). If the absence of requests for clarification was indicative of complete comprehension, older children should have made fewer requests than younger children, due to their superior communication abilities. Thus, perhaps a more plausible explanation for children's reluctance to seek clarification could be that of comprehension failure. Due to limited metacognitive abilities some children, particularly those that are very young, may lack the awareness to recognise when they do not understand (Malloy et al., 2015). However, research has found that this explanation alone is insufficient in explaining

children's failure to draw attention to misunderstandings. Waterman et al. (2000) found that even when children have this awareness, they frequently fail to make requests. It has therefore been suggested that children may realise their limitations but may not wish to acknowledge them. Children may be succumbing to demand characteristics (Malloy et al., 2015). They may perceive it as necessary to provide a response to every question, as this is what they have become accustomed to during everyday discourse (Brubacher et al., 2015). Both comprehension failure and demand characteristics offer compelling explanations as to why children fail to make requests for clarification during investigative interviews. It is likely that both of these explanations contribute to this failure. Theoretically, demand characteristics should be easier to overcome with the use of ground rules. However, studies have examined whether children can be taught effective comprehension monitoring strategies and can thus effectively implement the 'I don't understand' rule (Peters & Nunez, 1999; Saywitz et al., 1999).

Very few studies have examined the efficacy of the 'I don't understand' rule. The studies that have been conducted have mixed findings. Danby et al. (2015) examined the ability of 5- to 9-year-olds to use the 'I don't understand' rule. All of the children heard the rule at the start of the interview. Yet, only half of the children practised it. The children had a maximum of three attempts to answer the practise question correctly. Danby et al. (2015) found that, even with practise, children use the 'I don't understand' rule infrequently and are likely to acquiesce to misleading questions. However, the practise strategies in Danby et al.'s (2015) study were very brief.

Studies that have employed more intensive training protocols have yielded more positive findings (e.g., Peters & Nunez, 1999; Saywitz et al., 1999). In Peters and Nunez's (1999) study children (preschool, kindergarten, and second-grade) received either Task Demand Training (TDT; The training involved teaching the children about the

role of a witness including the importance of alerting the interviewer when a question is not understood) or TDT and Comprehension Monitoring Training (CMT). The children in the latter group learnt about strategies to recognise and respond appropriately to incomprehensible questions (CMT), along with their role as a witness (TDT). The training involved three 10-minute sessions. The children who received the combined training (TDT and CMT) were more likely to ask for complex questions to be rephrased. This resulted in more correct responses as children were more likely to answer rephrased questions correctly. Although the preschool children gave fewer correct responses than the older children, they were equally likely to ask for complex questions to be rephrased, indicating that children, as young as 5 years old, can benefit from training in comprehension monitoring. However, it is important to note that although beneficial effects of training were found in Peters and Nunez's (1999) study, the effect was modest. Overall, the children who had received training only asked for 25% of the complex questions to be rephrased.

In contrast, Saywitz et al. (1999) found that children (6 and 8 years old) who received extensive training (totalling 35 to 45 minutes) in comprehension monitoring indicated either that they did not understand or asked for the question to be rephrased on 73% of occasions. Saywitz et al. (1999) found that children who had received the training were significantly more accurate and less inaccurate than those who had either received rephrasing or motivational instructions. Rephrasing instructions involved the children being told to inform the interviewer of any misunderstandings. They did not practise utilising the strategies they were taught. Motivational instructions involved the children being told to do their best. They were made aware that some questions might be difficult to understand. It is important to note that the children who had received rephrasing instructions performed better than those who were given the motivational instructions. This indicates that rephrasing instructions alone improve baseline performance but are far more effective when incorporated with practise,

feedback, and reinforcement. Despite the positive effects observed by Peters and Nunez (1999) and Saywitz et al. (1999), neither training strategy would be feasible in practice as both were very time intensive. Danby et al. (2015) thus highlight the necessity for future research. The research should seek to determine the least intensive practise regime that can still reliably assist children to employ the 'I don't understand' rule.

'Tell Me if I Get It Wrong'. Interview guidelines tend to discourage leading questions (ABE, MoJ, 2011). However, interviewers may unintentionally incorporate incorrect information into investigative interviews. Research has found that interviewers sometimes paraphrase children's statements incorrectly (Evans et al., 2010), put incorrect versions of events to children, and confuse details across occurrences (Pichler, 2018). Worryingly children often neglect to correct these errors (Evans et al., 2010). To overcome this compliance, some guidelines include a statement about correcting the interviewer's mistakes (e.g., 'tell me if I get it wrong'). Again, the research findings regarding the efficacy of this ground rule are very mixed. Some studies have found the rule to have little or no beneficial effects (Ellis et al., 2003; Geddie et al., 2001). Ellis et al. (2003) examined the impact of the rule on children's (3- to 5-year-olds) recall of a staged event. Half of the children received a brief instruction (this included informing the interviewer if something was untrue) at the beginning of the interview. The instruction / rule had a negligible effect on the children's accuracy. This could be attributed to the children not being given the opportunity to practise the rule. Contrary to this explanation, Danby et al. (2015) found no beneficial effects of practise on the utilisation of this rule. However, as noted above, practise opportunities within Danby et al.'s (2015) study were limited and not intensive enough to allow for the acquisition of the more cognitively complex rules. Although children were afforded more intensive practise opportunities (i.e., interviewers worked with each child until it was felt the child comprehended the rule) in Geddie et al.'s (2001) study, no beneficial

effects of ground rules were found on 3- to 6-year-olds recall. However, a different research assistant to the one conducting the interview delivered the ground rules. The rationale for this was to control for interviewer bias, however this may have made it difficult for the children to generalise the rules from the training to the interview.

Nevertheless, there has been some evidence that the ‘tell me if I get it wrong’ rule is effective with young children (Gee et al., 1999; Krackow & Lynn, 2010). Krackow and Lynn (2010) tested the efficacy of Event Report Training for 4- to 8-year-olds. A component of this training is suggestibility reduction. This component sought to inform children of the interviewer’s naivety and the permissibility of correcting the interviewer if a mistake was made. The training was intensive and involved modelling, practise, and feedback of the rules. Krackow and Lynn (2010) found that the training reduced suggestibility in 4- to 5-year-old children. An earlier study by Gee et al. (1999) also found pre-interview training to be beneficial. As in Krackow and Lynn’s (2010) study, children were encouraged to correct the interviewer’s errors. Gee et al. (1999) found that the training reduced commission errors (i.e., falsely reporting something that did not happen) to misleading questions in children aged 9 to 13 years old. However, this did come at a cost. The training also reduced correct responses to non-misleading questions. Children were overgeneralising the rule and had become wary of providing available answers. In order to overcome this, Gee et al. (1999) adapted the original training package. In addition, to the original ground rules, the modified training encouraged and reinforced correct answers. This had a positive impact upon children’s responses and the modified training was found to reduce commission errors without compromising accuracy. However, there are difficulties in interpreting the findings from Krackow and Lynn (2010) and Gee et al.’s (1999) study. Neither study solely examined the instruction to correct the interviewer. For example, in Krackow and Lynn’s (2010) study the children were informed of the interviewer’s naivety and in Gee et al.’s (1999) study children were instructed not to guess

and subsequently praised for ‘don’t know’ responses. It is therefore very difficult to determine the relative contribution of the ‘tell me if I get it wrong rule’ to children’s reduced suggestibility.

‘I Wasn’t There, I Don’t Know What Happened’. In everyday discourse, children typically converse with knowledgeable adults. Of concern is that, as adults, children may adopt this same perception of investigative interviewers. This is problematic given that children are more likely to guess when they perceive an interviewer to be knowledgeable, as opposed to naive (Waterman et al., 2004). A statement highlighting the interviewer’s naivety (e.g., ‘I wasn’t there, I don’t know what happened’) is therefore essential. The statement also emphasises to the child the importance of providing sufficient detail along with reducing the power imbalance that exists within an investigative interview. In the CI (Fisher & Geiselman, 1992) this is referred to as the transfer of control from the interviewer to the witness. It is hoped that by transferring control to the child it will reduce the likelihood that the child will acquiesce to the interviewer’s suggestions (Mulder & Vrij, 1996). A number of studies have examined this ground rule (e.g., Beuscher & Roebbers, 2005; Córdón et al., 2005; Krackow & Lynn, 2010). Beuscher and Roebbers (2005) found the instruction to have no impact upon children’s performance. It did not improve recall accuracy, nor did it increase the frequency of ‘I don’t know’ responses to unanswerable questions. Although other studies have elicited more positive findings (e.g., Córdón et al., 2005; Krackow & Lynn, 2010), very few studies have examined the rule in isolation. Thus, making it very difficult to ascertain its beneficial effects. One study that can demonstrate the rule’s utility is that of Mulder and Vrij (1996). They crossed the naivety rule with the ‘I don’t know’ rule. The naivety rule was found to increase children’s resistance to suggestion, when given the naivety rule children provided less incorrect responses (26% vs. 40%) to misleading questions. The rule was found to be beneficial for younger (4- and 5-year-olds) and older children (8- to 10-year-olds). However,

it is important to note that the combination of the two rules (i.e., interviewer naivety and ‘I don't know’) was found to be more effective than either in isolation.

‘I Don't Know’ / ‘Don't Guess’. The ‘I don't know’ rule has been the most extensively researched of all the ground rules. It is also argued that it is the most salient of the rules. The rationale for this claim being that even if children are unable to indicate why a question poses a challenge (i.e., the interviewer has made a mistake or the child does not understand) the ‘I don't know’ rule reduces their propensity to guess (Brubacher et al., 2015). Brubacher et al. (2015) argue that an ‘I don't know’ response could serve a number of different purposes including ‘I don't remember’, ‘I don't understand’, or ‘you got it wrong’. Thus, an ‘I don't know’ response does not always accurately reflect a child's knowledge, hence the necessity for additional ground rules. It would however suggest that an ‘I don't know’ response may be more common than other requests for clarification. This claim is bolstered by research which suggests that the ‘I don't know’ rule is relatively easy to comprehend, in comparison to the other rules (i.e., ‘I don't understand’; Dickinson et al., 2015).

Despite children having the ability to comprehend the rule, ‘I don't know’ responses have been found to be relatively rare in investigative interviews (Earhart et al., 2014). As with all of the ground rules, the findings have been mixed regarding the efficacy of this rule. Some studies have found the rule to have little impact upon the frequency of ‘I don't know’ responses (Earhart et al., 2014). Earhart et al. (2014) reviewed transcripts of 76 investigative interviews with children aged 4 to 13 years old. They found that the introduction of the rule did not elicit more ‘I don't know’ responses. However, the rule was not modelled or practised at the outset of any of the investigative interviews. Yet a failure to practise the ground rules is not the only explanation that could account for the lack of ‘I don't know’ responses in Earhart et al.'s (2014) study. It may have also been confounded by the interviewers' behaviour.

Interviewers rejected 'I don't know' responses on approximately 30% of occasions and then often preceded by asking more risky questions about the same topic. This undermines the permissibility of an 'I don't know' response and likely made children believe that such a response was undesirable and unacceptable. However, earlier studies have yielded equally disconcerting results. Geddie et al. (2001) and Ellis et al. (2003) found no beneficial effects of the 'I don't know' rule on children's accuracy and resistance to suggestion. The limitations of each of these studies are discussed above.

Other studies of the 'I don't know' rule have reported more positive results (Cordón et al., 2005; Gee et al., 1999; Mulder & Vrij, 1996; Waterman & Blades, 2011). Mulder and Vrij (1996) found that, compared to children who received no interview instructions, children presented with the 'I don't know' rule provided fewer incorrect responses. However, when children were presented solely with this rule, children also provided less meaningful answers. This effect was mediated by the inclusion of an additional rule, associated with the interviewer's naivety. The additional rule increased the number of meaningful responses from 43% to 68% (Mulder & Vrij, 1996). A number of other studies have demonstrated the superiority of multiple ground rules (Cordón et al., 2005; Waterman & Blades., 2011). Waterman and Blades (2011) found that for younger children the inclusion of an additional rule associated with interviewer naivety, alongside the 'I don't know' rule, significantly increased the number of appropriate responses to unanswerable questions. This brought the performance of the younger children (5- to 6-year-olds) in line with that of the older children (7- to 8-year-olds). A further study, which demonstrates the utility of multiple rules was conducted by Cordón et al. (2005). They found that when children received conversational rules ('I don't know', 'I can't help', 'I may trick you') errors were reduced by 14%, compared to a control group (i.e., received placebo instructions).

Additional Instructions. A recent study by Quas et al. (2018) looked at the efficacy of two additional interview instructions: promising to tell the truth and the putative confession (i.e., telling children that the alleged offender “told me everything that happened and wants you to tell the truth”). ABE guidance (MoJ, 2011) recommends exploring truth and lies with children, during the rapport phase of the interview. However, the guidance explicitly states that “**no** attempt should be made to get the witness to swear an oath” (MoJ, 2011, p. 72). The guidance also makes no reference to the putative confession. This may be because of the ethical implications of giving such an instruction when the alleged offender has not made a confession. Nevertheless, the findings of Quas et al.’s (2018) study are worthy of mention, given the different strategies that are potentially permissible internationally. Their study involved children, aged 4 to 9 years old (maltreated and non-maltreated), being interviewed regarding an incident in which they were playing with a stranger and a number of toys got broken. The children were instructed by the stranger to not disclose the transgression. The researchers found that promising to tell the truth increased disclosures (63% who received the rule vs. 31% in the control condition) to open-ended questions amongst older children but not younger children. The authors attributed this to younger children having less appreciation of the importance of a promise and how failing to comply with a promise can undermine trust. The putative confession, on the other hand was effective in increasing disclosures (63% disclosed) to open-ended questions irrespective of age. However, neither instruction was found to increase children’s resistance to suggestion when asked yes/no questions that explicitly mentioned wrongdoing. Overall, Quas et al.’s (2018) findings suggest that both promising to tell the truth and the putative confession may increase disclosures from children. However, the putative confession in particular has significant risks. It could potentially seriously jeopardise the rapport between the interviewer and child, if the child were to discover that they had been misled and the alleged offender had not confessed. Thus, these

two additional instructions will not be considered in further discussion of the ground rules literature.

The Utility of Ground Rules. Given the lack of agreement across best practice guidelines (Brubacher, et al., 2015) it is unsurprising that the aforementioned studies vary greatly in their methodologies. The studies have varied in the number of ground rules presented and the extent to which the rules were modelled and practised. Some studies have examined the efficacy of a single rule (e.g., Saywitz et al., 1999), whilst others have examined the efficacy of multiple rules (e.g., Cordón et al., 2005; Mulder & Vrij, 1996; Krackow & Lynn, 2010; Waterman & Blades, 2011). As mentioned above, the latter approach makes it very difficult to determine the relative contribution of each rule. Although this is interesting from a theoretical perspective, practice guidelines tend to advocate the introduction of multiple ground rules (e.g., ABE, MoJ, 2011; NICHD Protocol, Lamb et al., 2007). Thus, it could be argued that research adopting this approach is more salient within an applied context. Research which has simultaneously employed multiple rules has generally yielded positive findings (e.g., Cordón et al., 2005; Krackow & Lynn, 2010; Mulder & Vrij, 1996; Waterman & Blades, 2011). One benefit of simultaneously introducing multiple rules is that it can safeguard against any one rule being overused (Mulder & Vrij, 1996; Gee et al., 1999). This is very important from an applied perspective. For example, an over-reliance on the 'I don't know' rule could potentially result in a child providing very little forensically relevant information, which in turn could jeopardise the investigation. Therefore, interviewers need to give careful consideration as to the number of rules to include. The more ground rules, the lengthier this phase of the interview and the less potential time available for discussing substantive issues.

Time spent discussing the ground rules is also dependent upon their delivery. Some studies have sought to teach children a rule by way of a simple statement (e.g., Beuscher &

Roebbers, 2005; Ellis et al., 2003; Waterman & Blades, 2011), whilst others have utilised practise opportunities (e.g., Geddie et al., 2001; Krackow & Lynn, 2010; Peters & Nunez, 1999; Saywitz et al., 1999). Although there have been some positive findings when the rules have simply been stated (Waterman & Blades, 2011), the research suggests that practice is generally a more effective approach when teaching children the rules (Danby et al., 2015). However, the extent of practice required, in order to learn a rule, has been found to vary subject to the rule's complexity (Danby et al., 2015). This can be further compounded by the age of the child (Dickinson et al., 2015). The aforementioned studies have examined the efficacy of ground rules with children aged 4 to 13 years old. Positive findings have emerged for all ages (e.g., 4 to 5 years, Krackow & Lynn, 2011; 5 to 8 years, Peters & Nunez, 1999; 4 to 5 years / 8 to 10 years, Mulder & Vrij, 1996; 9 to 13 years, Gee et al., 1999). This is somewhat surprising, given Brubacher et al's (2015) concerns surrounding the acquisition of ground rules by younger children. However, there is evidence to suggest that their concerns hold some weight. Studies have shown that younger children have more difficulty acquiring and employing some of the more cognitively complex rules (i.e., 'I don't understand'; Danby et al., 2015; Dickinson et al., 2015), and require more intensive training protocols than older children. Without intensive training the introduction of these rules would be futile (Danby et al., 2015). A balance therefore has to be struck between time efficiency and narrative accuracy. Interviewers need to decide whether the extra time spent practising ground rules outweighs the risks of fatigue or delay. Overall, there is evidence to suggest that all of the ground rules have some beneficial effects.

Criticisms of Ground Rules. Research on the efficacy of ground rules has produced mixed results. With practise, beneficial effects of ground rules have been observed in children as young as 4 years old (Krackow & Lynn, 2010). Nevertheless, some argue that ground rules instructions should be omitted from best practice guidelines and interview protocols. They

have offered a number of reasons for their omission. Firstly, it is argued that ground rules are too abstract and practise questions are developmentally inappropriate (Geddie et al., 2001). Dickinson et al. (2015) found marked differences in children's ability to comprehend the different ground rules, with the 'I don't understand' rule being the most challenging. However, it is important to note that the practise questions asked in this study varied considerably in their structure and linguistic complexity. Some rules utilised specific-closed questions (e.g., 'don't guess' - "what is my dog's name?") whilst others employed yes/no questions (e.g., 'tell me when you don't understand' - "is my shirt gridelin?"). Research has found that children may simply respond in the affirmative to yes/no questions (see section 2.2.2), thus this could have potentially confounded the results. It is also argued that children's inability to answer ground rules questions could damage their credibility and perceived competence (Evans & Lyon, 2012). Furthermore, it is argued that children only have limited attentional resources and those resources would be better used eliciting investigative relevant information (Anderson et al., 2009). Contrary to this argument, Dickinson et al. (2015) found that it generally took less than 4 minutes to introduce and practise ground rules. However, Dickinson et al.'s (2015) study did not examine whether children then went on to employ the rules in an interview scenario. As cautioned by Brubacher et al. (2015) children may not possess the executive skills to hold the rule in mind whilst inhibiting more dominant responses. They therefore may be unable to utilise the rule within an interview context. Although of concern this can potentially be overcome with the addition of non-verbal scaffolds such as communication aids (e.g., rule cards which may act, as an aide memoire, in a similar way to that of 'think sheets'; Matheson & Hutchinson, n.d.).

2.2.4 Communication Aids

Children may be reluctant to disclose abuse, have difficulties with recall, or lack the vocabulary to put their experiences into words (Morgan et al., 2013). Hence, they may require

additional support or scaffolding in order to provide a complete and accurate narrative account. Communication aids can be used to facilitate and support children's communication. Examples include body diagrams, drawings, rule cards, visual timetables, emotions scales, calming objects, dolls, props, and figures. Communication aids have multiple functions (Mattison, 2015; Plotnikoff & Woolfson, 2015). However, for the purpose of this research the focus will be on the communication aids that are associated with gathering and clarifying evidence namely dolls, drawings, and body diagrams. Each of which should only be used following careful and extensive planning (Marchant, 2013) and never in conjunction with leading questions (MoJ, 2011).

Anatomical Dolls. Anatomical dolls were created in the 1970s. The dolls rapidly gained popularity amongst forensic interviewers (Poole & Bruck, 2012). Although the use of dolls has been relatively widespread (Hlavka et al., 2010), since the 1980's, the practice has been seen as controversial (Salmon et al., 2012). During the 1980's there was a common misconception that anatomical dolls could be used as a diagnostic test of sexual abuse (Hlavka et al., 2010). It was thought that children who had been sexually abused would play and interact with the dolls differently to non-abused children, and thus through clinical interpretation of the child's behaviour one could determine the likelihood that abuse had taken place (Everson & Boat, 1994). However, in a review of interview guidelines, conducted by Everson and Boat (1994), none of the guidelines or protocols advocated the use of anatomical dolls as a diagnostic test of abuse. In fact, several of the guidelines cautioned against the over-interpretation of children's behaviour. This is because the over-interpretation of children's behaviour can lead to the interviewer posing overly suggestive questions to the child and can potentially steer an investigation in a direction which would have otherwise been unwarranted (Everson & Boat, 1994).

However, the dolls are seen as possessing a multitude of other functions and have been described by interviewers as valuable for the purposes of clarification, consistency, distancing, and communication (Hlavka et al, 2010). Research studies have yielded mixed findings, some of which raise major concerns relating to the use of dolls as a communication aid within a high-stakes legal arena. An early study by Saywitz et al. (1991) found that dolls paired with yes/no and specific-closed questions can increase true disclosures of anal and vaginal touching whilst eliciting only small numbers of false reports (i.e., 5.56% rate of false reports of anal touching; 2.86% rate of false reports of vaginal touching). Within their study commission errors (i.e., falsely reporting vaginal / anal touching) were far less frequent than errors of omission (i.e., failing to report true incidents of vaginal / anal touching). Contrary to these findings, Bruck et al. (2000) found that when asked to show on the doll children were as likely to make errors of commission as errors of omission. Many of the children in their study demonstrated sexualised behaviours with the dolls including inaccurately using a spoon or other prop to show anal or vaginal touching (i.e., 28% of 3-year-old children; 27% of 4-year-old children). Furthermore, Santtila et al. (2004) found that the introduction of dolls can hinder children's communication. In their study, dolls were associated with shorter and less detailed responses from children and poorer practices by the interviewers. When dolls were introduced, interviewers tended to ask longer questions (containing more words) along with more suggestive utterances.

In light of these findings, some researchers have suggested that dolls are not an effective tool for scaffolding children's communication (Santtila et al., 2004) and should not be used in investigative interviews with children under the age of 5 (Bruck et al., 2000). The assertion that dolls should not be used with very young children is due to research showing that dolls can increase suggestibility (e.g., Bruck et al., 2000) and young children having a lack of representational insight. Representational insight refers to the ability to comprehend

that the doll is simultaneously an object and a symbol representing a particular person. It is a prerequisite for conveying information about oneself using a doll (Poole & Bruck, 2012).

Despite the belief that children develop this insight at approximately 3 years old, studies have shown that older children can have difficulty with more complex representational tasks (Poole et al., 2011). Lytle et al. (2015) examined children's ability to map the location of body touches on dolls. They observed significant deficits in the ability of 3-year-olds to use symbols to demonstrate bodily touch. Although performance did improve with age, 5-year-olds still made considerable errors with less than half of the 5-year-olds achieving perfect scores. In addition to representational insight, Poole et al. (2011) identified two other cognitive skills that children require in order to use dolls to accurately report past events. Namely children must be able to map past events onto the dolls and have a level of attentional focus whereby they do not digress into fantasy play. Following challenges to doll-assisted interviews in court and in a bid to overcome the cognitive requirements associated with the use of dolls, interviewers began using body diagrams (Poole & Bruck, 2012).

Body Diagrams. As with dolls, body diagrams take the onus off 'telling' and place it on 'showing'. Body diagrams are characteristically black and white line-drawings of male and female, children and adults. Diagrams vary in that they can be either clothed or unclothed; and some graphically depict male and female body parts, whilst others are gender neutral (Poole & Bruck, 2012). Body diagrams are believed, by some, to be superior to dolls. It is argued that they are less likely to encourage exploratory sexualised play and are recognised earlier in development as a symbolic representation of self (Poole & Bruck, 2012). Contrary to this belief, Lytle et al. (2015) found that 3- and 4-year-old children performed better when asked to locate body touches on a 3D doll as opposed to a 2D body diagram.

Overall evidence for the application of body diagrams is inconsistent. Aldridge et al. (2004) found that body diagrams can increase the amount of forensically relevant information

provided by children. They found that with the introduction of body diagrams, 4- to 7-year-olds provided on average 95 additional details. A recent study by Dickinson and Poole (2017) equally found some support for the use of body diagrams. Body diagrams were found to be effective in eliciting detailed reports of touching, but only amongst children 5 years and older who had not previously disclosed. When probing for additional disclosures, however, body diagrams were associated with higher error rates. Other studies have also seen accuracy compromised with the use of body diagrams. Bruck et al. (2016) found that body diagrams elicited more disclosures of sexual touching. However, this was found with younger children to come at the expense of accuracy (i.e., 50% of 3-year-olds, 33% of 4-year-olds, and 22% of the 5-year-olds incorrectly reported genital touching). A study by Willcock et al. (2006) again found a high percentage of commission errors amongst children interviewed with a body diagram. Only 47.8% of children's responses were found to be accurate with 11.3% of children inaccurately reporting genital touching. These studies are not isolated cases, other studies have reported inaccuracies in children's accounts following the use of body diagrams (e.g., Brown et al., 2012; Otgaar et al., 2012; Poole & Dickinson, 2011). Any factor that could compromise accuracy within a legal setting is highly problematic. This raises the fundamental question of whether these props should be introduced within an investigative interview.

However, the above findings need to be interpreted with caution. The first difficulty lies with the field studies (e.g., Aldridge et al., 2004). As collaborative evidence is seldom available in cases of CSA (Willcock et al., 2006), it is impossible to determine the accuracy of children's disclosures. Although experimental studies (e.g., Brown et al., 2012; Salmon et al., 2012; Willcock et al., 2006) are able to measure accuracy it is unclear as to whether the innocuous touches involved in experimental studies are salient enough to generalise the findings to real-world allegations of sexual touching and abuse. The high omission errors observed in the studies by Brown et al. (2012) and Willcock et al. (2006) suggest that the

touches involved in these studies may not have been attended to by the children. In the context of more exciting aspects of the events, the touches may not have proved salient enough to be encoded into memory. As such some studies (e.g., Bruck et al., 2000; Saywitz et al., 1991) have examined children's recall of medical examinations involving genital and anal touching. However, using events that incorporate salient touch does not overcome the second major limitation of the previous studies - the increase in recognition questions and poor interview practices that accompanied the introduction of the props (e.g., Aldridge et al., 2004; Santtila et al., 2004). Research has shown that children are more likely to provide erroneous responses to these questions forms (Lamb et al., 2007). Thus, this makes it very difficult to ascertain whether accuracy was compromised in these studies as a result of the props (e.g., dolls or body diagrams) or as a result of poor questioning practices. Furthermore, many of the studies failed to assess whether children had achieved representational insight (Hlavka et al., 2010). This is important as many of the studies involved young children, aged 7 years old and under (e.g., Brown et al., 2012; Salmon et al., 2012; Saywitz et al., 1991; Willcock et al., 2006). Without representational insight children are unable to use dolls or body diagrams to accurately convey information about themselves (Poole & Bruck, 2012). As this was not assessed, it is impossible to establish whether props are inherently suggestive or whether this practice was simply developmentally inappropriate for the samples used in the aforementioned studies.

Drawing. According to the ABE guidelines (MoJ, 2011) children find it easier to understand the symbolic nature of their own drawings, compared to either dolls or body diagrams. As with dolls and body diagrams, there is an extensive body of research exploring the impact of their drawing on the richness and accuracy of children's narrative accounts (Katz & Hershkowitz, 2010). Evidence has emerged which suggests that drawing facilitates children's verbal communication. There is an abundance of research showing that drawing

increases the amount of information elicited from children (e.g., Barlow et al., 2011; Katz & Hamama, 2013; Katz & Hershkowitz, 2010; Patterson & Hayne, 2011). Some studies have found that children who draw and tell about their experiences report approximately twice as much information compared to those who are only asked to tell (Butler et al., 1995; Gross & Hayne, 1998; Macleod et al., 2013; Woolford et al., 2015). Furthermore, this increase in information does not appear to compromise the accuracy of children's accounts (Barlow et al., 2011; Butler et al., 1995; Gross & Hayne, 1998; Patterson & Hayne, 2011). In fact, Gentle et al. (2014) suggest that drawing may act as a protective tool. They found that children who were asked to draw and tell, as opposed to just tell, were more resilient to suggestion. However, it is important to note that Gentle et al. (2014) found drawing to have no impact on children's informativeness during free recall. Nevertheless, there appears to be a general consensus amongst researchers that drawing is effective in facilitating children's recall and communication (Woolford et al., 2015). There is less of a consensus, however, regarding the mechanisms underpinning the effectiveness of drawing as a communication tool.

Researchers have proposed a number of explanations that could account for the facilitative effects of drawing on children's recall and communication. It is suggested that both cognitive and emotional processes could be involved (Katz & Hershkowitz, 2010). One of the most frequently cited cognitive explanations is that drawing can act as an auto-generated retrieval cue (Barlow et al., 2011; Butler et al., 1995; Gross & Hayne, 1998; Katz & Hershkowitz, 2010). Gross and Hayne (1998) found preliminary evidence to support this memory retrieval hypothesis. They found that as children's drawing ability increased so did the amount of information the children provided when interviewed. It was asserted that the better the representational quality of the drawing the more effective it would be in its role as a contextual retrieval cue. However, other researchers have proposed that the facilitative effects of drawing may be the result of interviewers adapting their verbal behaviour (Macleod et al.,

2013; Patterson & Hayne, 2011). Some studies have found that when children draw interviewers use more minimal responses such as ‘uh huh’ (Woolford et al., 2015). Minimal responses have been found to positively correlate with the amount of information children report (Macleod et al., 2013). Yet this explanation alone cannot account for the positive effects of drawing, due to evidence that drawing can also elicit detailed and accurate accounts when used in conjunction with specific-closed questions (Barlow et al., 2011; Butler et al., 1995). A further explanation for the facilitative effects of drawing is that interviews that involve drawing tend to be longer and therefore offer extended retrieval opportunities (Butler et al., 1995; Macleod et al., 2013). A study by Salmon et al. (2012) offers some support for this explanation. Salmon et al. (2012) found that, following an opportunity to draw, children reported approximately four additional details when re-interviewed. Yet the positive effect of drawing was shown to be no greater than other cognitive activities (e.g., puzzles). In sum, it is very difficult to ascertain which, if any of these explanations account for the positive effects of drawing as a communication aid. This lies in the fact that controlling the duration of the interview and the number of minimal responses would prove very difficult as it would serve to make the adult-child interaction artificial and uncomfortable, reducing rapport (Macleod et al., 2013). Nevertheless, failing to control for these variables makes it impossible to ascertain whether drawing per se is responsible for children’s enhanced recall or whether it is the result of an interaction between drawing, extended opportunities for recall, and improved interviewer practice.

The Utility of Communication Aids. Wolfman et al. (2018) examined interviewers’ use of communication aids in 98 investigative interviews with children alleging sexual abuse. Sixty-two percent of the interviews reviewed included at least one communication aid, with sketch-plans the most commonly used tool. Sketch-plans involve children drawing the location or spatial layout of an event. It is not surprising that drawing emerged as the most

commonly used tool in Wolfman et al.'s (2018) study as previous research suggests that it is a safer and more effective communication tool than either dolls or body diagrams. However, it is important to note that in contrast to drawing many of the studies involving dolls and body diagrams were concerned with eliciting disclosures of touch (e.g., Brown et al., 2012; Bruck et al., 2000; Willcock et al., 2006). The concept of touch can be difficult for young children to comprehend. It is often perceived by young children as something you do with your fingers. This can subsequently lead to incomplete accounts (Marchant, 2013) and may have exacerbated the observed differences. Although, the above studies offer very little evidence to suggest that dolls or body diagrams are capable of eliciting accurate and reliable information from children, none of the above studies used communication aids following a pre-interview assessment or in conjunction with an RI. A survey of RIs found that 95% used communication aids in their work, with drawings and body diagrams perceived as the most effective at facilitating communication (Owen, 2016). The popularity of drawing was further emphasised in a recent study by Mattison and Dando (2020). Twenty-two RIs ($N = 35$) reported using communication aids in their practice. Of those, 20 frequently used drawing as a method of facilitating communication (i.e., responded 'often', 'almost always', or 'always' on the survey). RIs have also reported using pipe cleaner figures, timelines, photographs, and prompt cards, to name a few (Plotnikoff & Woolfson, 2015).

2.2.5 What Happens in Practice

Despite an international consensus regarding what constitutes best practice, research suggests that most investigative interviewers fail to adhere to interviewing guidelines (Powell et al., 2010) and possess little knowledge of the psychological mechanisms underpinning best practice (Dodier, Tomas, et al., 2019). Currently, there appears to be a significant gap between knowledge and practice. For example, best practice guidelines advocate a phased approach. The approach involves four phases: rapport, free narrative, questioning, and

closure. Although, guidance is given as to what each phase should comprise of (MoJ, 2011), evidence has emerged to suggest that interviewers fail to heed these recommendations, thus omitting crucial elements of the interview (Westcott & Kynan, 2006; Wolfman et al., 2016). Best practice guidelines (MoJ, 2011) also emphasise the importance of using open-ended questions. Yet studies have shown that interviewers rely predominantly upon specific-closed, forced-choice, and yes/no questions when eliciting narrative accounts from children (Johnson et al., 2015; Korkman et al., 2006; Wolfman et al., 2016).

Interviewers have offered explanations to account for their deviations from best practice including needing to elicit specific information from children (Guadagno et al., 2013). Further explanations include interviewers having a lack of understanding regarding the distinction between different question types (Wright & Powell, 2006) and the unusual nature of maintaining an open-ended discourse (Yarbrough et al., 2013). Maintaining an open-ended discourse, within an investigative interview, is challenging as it is such a stark contrast from how people from English-speaking countries typically converse with one another (Powell, 2000). Typically, conversations consist of question-and-answer exchanges wherein the questions often seek a desirable response; open-ended questions are rarely used within everyday interactions (Yarbrough et al., 2013). The aforementioned explanations are discussed in more detail below, along with potential strategies with which to address deviations from best practice (see Powell et al., 2010).

Knowledge of Psychological Mechanisms Underpinning Best Practice. Eyewitness reports can be instrumental in solving criminal cases (Dodier, Tomas, et al., 2019). These reports are also one of the main causes of wrongful convictions (Innocence Project, 2015). In order to avoid such negative outcomes, professionals working within the CJS are advised to keep up to date with the latest advances in memory research. Amongst other things, police officers should have an awareness of the detrimental effects of leading questions; the impact

of co-witnesses; the differences between children's and adults' memory abilities; and the effect of stress / trauma on memory (Dodier, Tomas, et al., 2019). However, research has shown that police officers often have a limited knowledge of memory, suggesting that they are not familiar with these factors that affect eyewitness reports (e.g., Chaplin & Shaw, 2016; Dodier, Tomas, et al., 2019). A recent study, by Dodier, Tomas, et al. (2019), compared memory-related knowledge and erroneous beliefs of police officers (25% of the officers specialised in child victim and suspect cases) and lay persons. Both groups scored poorly in terms of knowledge, with police officers demonstrating more erroneous beliefs. This was investigated further by comparing officers who had undergone training in investigative interviewing and their untrained counterparts. Similar low scores were observed for both groups. However, officers who reported conducting five or more interviews per month were found to hold more erroneous beliefs than officers who reported conducting no interviews at all. The lack of knowledge and acceptance of false beliefs held by this active group of interviewers may account for the low quality of investigative interviews reported in previous studies.

Adherence to Recommended Interview Structure. Research has found that interviews rarely comprise of all of the requisite parts (e.g., rapport, free narrative, questioning, and closure; Hill & Davies, 2013; Westcott & Kynan, 2006). A study by Hill and Davies (2013) examined whether the revision and development of the guidance had succeeded in improving the practice of investigative interviewers. Their study compared MOGP (Home Office, 1992) and ABE (MoJ, 2011) interviews. They found that neither the Memorandum or ABE interviews consistently included all four interview phases (i.e., 25% of interviews), with the closure and rapport phases most frequently omitted. The interviews were also comparable in their ability to elicit a free narrative account. Yet there was evidence of poor practices during this phase, namely raising the allegation directly with the child and

citing the allegation of another child. Of further concern was the proportion of open-ended questions across both Memorandum (Home Office, 1992) and ABE (MoJ, 2011) interviews. Only 7.4% of interviewer questions were open-ended in the Memorandum interviews (Home Office, 1992) and 8.1% in the ABE (MoJ, 2011). Specific-closed questions also dominated both interviews, consisting of 40.1% in the Memorandum (Home Office, 1992) and 41.9% in the ABE (MoJ, 2011) interviews. There appeared to be difficulties across both the Memorandum (Home Office, 1992) and ABE (MoJ, 2011) interviews in maintaining an open-ended discourse.

In 2014 a joint inspection was conducted by Her Majesty's Inspectorate of Constabulary (HMIC) and Her Majesty's Crown Prosecution Service Inspectorate (HMCPSP) into whether police forces were adhering to the ABE guidance (Criminal Justice Joint Inspection, 2014). Overall, it was found that compliance with the guidance was poor. The report echoed many of the deficiencies reported in previous research. Many of the interviews failed to include all of the ground rules; the free narrative and closure phases were often too brief; and there was an over-reliance on specific-closed, forced-choice, and leading questions. Furthermore, interviewers had a propensity to ask developmentally inappropriate questions regarding complex concepts such as dates, times, lengths, and heights. Interestingly, during the investigation it emerged that the practice of some forces was far superior to that of other forces. The investigation found greater consistency and coherence in practice within forces with specialist and centralised models. Despite some forces demonstrating better practice the conclusion of the investigation was that the guidance is not achieving what it set out to do - the failure being in the implementation.

Adherence to Recommended Question Forms. International best practice guidelines state that investigative interviews should consist predominantly of open-ended questions (e.g., MoJ, 2011). Yet research has found that this recommendation is rarely implemented by

practitioners. Deviations from best practice guidelines have been documented in the UK (Aldridge & Cameron, 1999; Davies et al., 2000; Sternberg, Lamb, Davies, & Westcott, 2001) Sweden (Cederborg et al., 2000), Norway (Johnson et al., 2015), Finland (Korkman et al., 2006), United States (Lamb et al., 2000; Warren et al., 1999), Canada (Luther et al., 2015), and New Zealand (Wolfman et al., 2016). A recent study conducted in Norway found that a mere 2% of interviewers' questions were open-ended (Johnson et al., 2015). Although, other studies have not yielded quite as disconcerting results, open-ended questions have frequently been found to constitute less than 25% of the total interviewer utterances (e.g., Cederborg et al., 2000; Luther et al., 2015; Wolfman et al., 2016). Research has found that interviewers tend to have an over reliance on specific-closed, forced-choice, and yes/no questions (Luther et al., 2015; Wolfman et al., 2016). Of further concern is that these questions are often introduced very early on in the substantive phase of the interview (Cederborg et al., 2000; Wolfman et al., 2016). A Finnish study found that in 66% of cases interviewers actually used forced-choice, yes/no, or leading questions in order to introduce the topic (Korkman et al., 2006). Research has also found that multiple questions are frequently used in child investigative interviews (Katz & Hershkowitz, 2012). This directly contravenes best practice guidance (MoJ, 2011). Despite an abundance of research and guidance deeming open-ended questions to be the most effective in eliciting detailed and accurate accounts from children their use appears infrequent at best.

Reasons for Failing to Comply with Recommendations. A number of reasons have been identified that could account for why investigative interviewers fail to adhere to best practice guidelines. The first explanation is that the investigative interview context differs significantly from other social contexts in which question-answer exchanges take place (Yarbrough et al., 2013). Best practice guidance (ABE; MoJ, 2011) advocates the use of open-ended questions. However, open-ended questions are uncommon in everyday discourse.

Instead, everyday discourse can promote bad habits such as phrasing questions in order to achieve a desired response (Yarbrough et al., 2013). These bad habits can potentially permeate the investigative interview, particularly given the high cognitive demands within this context (i.e., listening to and remembering what the witness has said whilst also appraising this information and generating further questions). These cognitive demands can lead to the interviewer experiencing increased cognitive load and can reduce their ability to accurately recall what a witness has said which may impact upon their subsequent questions and adherence to the best practice guidance (Hanway et al., 2020). Another explanation for interviewers failing to comply with the guidelines is their lack of understanding regarding what constitutes an open-ended question. Interviewers have been shown to have difficulty distinguishing between open-ended and specific-closed questions (Wright & Powell, 2006). A recent study, by Yi and Lamb (2018), looked at the ability of police officers to accurately identify different question types. One hundred and twenty Korean police officers were required to classify 50 questions into their respective categories. Overall, the accuracy rate was only 65.6%. It has also been found that interviewers do not recognise the potential of open-ended questions and fail to comprehend their importance. Many interviewers perceive specific-closed questions as more apt for eliciting specific details from children (Guadagno et al., 2013; Powell et al., 2010; Wright & Powell, 2006). A study by Guadagno et al. (2013) identified five themes whereby interviewers felt it necessary to deviate from best practice guidelines and use specific-closed and yes/no questions (i.e., identifying the alleged perpetrator, determining the offender's intent, ascertaining the time and location of the incident, establishing whether penetration occurred, and determining the child's terminology for describing genitalia). Furthermore, it has been suggested that the utility of open-ended questions is not adequately enforced by superiors. This, it has been proposed, is particularly problematic within the police service where there exists a mentality of learning 'on the job'

(Powell et al., 2010). Trainees may experience pressure from peers who rely on old methods, they claim ‘work’. These peers or ‘old hands’ not only fail to recognise the necessity of improving their own practice but also impede the learning and development of their colleagues (St-Yves et al., 2014).

When considering the quality of ABE interviews and adherence to best practice guidelines (MoJ, 2011) one must also consider the unprecedented pressures that austerity has placed upon police forces in England and Wales. Recent inspections paint a very bleak picture of the current climate. In a recent National Child Protection Post-Inspection Review of the Metropolitan Police Service officers reported being “under significant pressure, with factors such as the capacity, capability and current vacancy levels affecting their ability to provide a consistently good service” (HMICFRS, 2019, p.13). In one missing person unit, visited during the review, there were only three staff on duty with 77 missing people, 57 of whom were children (HMICFRS, 2019). Unfortunately, this is not an isolated incident. Nationally, one in five police investigator posts is either vacant or filled with ‘untrained’ officers (HMICFRS, 2018). In fact, in response to funding cuts some forces are moving away from specialist child protection units to merge public protection units or ‘omni-competent’ policing. The ‘omni-competent’ model involves equipping detectives to deal with child protection alongside other, more general, investigative work (Plotnikoff & Woolfson, 2019). It is feared that this move away from specialist services could risk diluting standards and compromising the quality of child interviews (Plotnikoff & Woolfson, 2019).

Similar concerns regarding a lack of resources and high staff turnover have been reported in other countries (e.g., Sweden, Ernberg, 2018). Such pressures leave staff with limited scope to attend additional training, practise skills, and seek feedback from colleagues (Powell et al., 2010). Finally, it has been suggested that the structure of training programs is not always effective in promoting best practices. Some programs have been found to be

unsuccessful (Aldridge & Cameron, 1999; Warren et al., 1999). Studies by both Aldridge and Cameron (1999) and Warren et al. (1999) found training to be ineffective in instigating behavioural change. Warren et al. (1999) found that despite improving the knowledge base of interviewers, the training did little to impact upon their questioning style. Following the training, interviewers still had a propensity to rely predominantly upon yes/no questions (although use decreased from 74% to 66%). Their use of open-ended questions did not significantly increase, nor did the amount of accurate information elicited from the children. Powell et al. (2005) attribute the ineffectiveness of some training programs to four main factors. Namely, the programs being intensive and not promoting continuous development, a lack of regular feedback, a lack of good practice examples, and limited organisational incentives (Powell et al., 2005). Motivation is also a key factor in the success of training programs. Although some investigators may enrol out of a desire to learn and improve their interviewing skills, others enrol out of obligation or in the hope of promotion (St-Yvette et al., 2014). These latter reasons create a negative platform upon which to build and could potentially limit the utility of training.

How Failure to Comply Can Be Overcome. Some training programs have addressed a number of the issues raised by Powell et al. (2005). Research has shown that training programs that incorporate regular feedback and are long-term can enhance the practice of investigative interviewers (Cederborg et al., 2013; Cyr et al., 2012; Price & Roberts, 2011). Cederborg et al. (2013) examined the efficacy of a 6-month training program implemented in Sweden. The program used a combined model that incorporated the NICHD Protocol (Lamb et al., 2007) and the PEACE model (Milne & Bull, 1999). Following the completion of the program, interviewers were found to use three times as many open-ended questions and two-thirds fewer forced-choice / yes/no questions than they had prior. These improvements were maintained for at least 2 months after the completion of the course (Lindholm et al., 2016, as

cited in Lamb, 2016). A similar study was conducted by Price and Roberts (2011). The study examined the efficacy of an 8-month training programme in Canada. The training programme was again based on the principles of the NICHD Protocol (Lamb et al., 2007). With the completion of the programme interviewers were found to use significantly more open-ended and less forced-choice, yes/no, and leading questions. Interviewers, post-training, were also more likely to employ a practice narrative and introduce the child to ground rules.

Despite these positive findings neither of the aforementioned studies conducted a long-term follow-up assessment of interview quality. It is therefore impossible to determine whether improvements in practice were maintained over sustained periods. Research has shown that without regular supervision and feedback interviewers can revert to poor and undesirable interview practices (Lamb et al., 2002). Lamb et al. (2002) found that the cessation of supervision led to interviewers using less open-ended and more forced-choice and leading questions when interviewing children. In addition, after supervision had ended, they found that interviewers tended to introduce the forced-choice and leading questions earlier in the interview. Despite the obvious benefits of regular feedback and supervision it is argued that in practice it would be too costly to implement and would be beyond the capabilities of large-scale police services. Police forces often have tight budgets and a high staff turnover (Powell et al., 2005). It may also prove difficult to identify sufficiently qualified supervisors. Enlisting more experienced police officers into this role is unlikely to be a successful approach if those officers rely on old methods and remain resistant to change (St-Yves et al., 2014). This raises the question of how police forces ensure that their officers are maintaining desirable interview practices.

Research suggests that effective professional practice requires knowledge to be organised around underlying principles, as opposed to surface facts (McMahon, 2006). This implies that to achieve mastery interviewers cannot rely solely on their knowledge of the

guidelines. Being able to regurgitate the guidelines is not enough, they need to possess an understanding of the theoretical underpinnings (Ernberg, 2018). Best practice guidelines are largely underpinned by the principles of developmental and cognitive psychology. It is therefore unsurprising that some training programmes incorporate elements of child development (Lamb, 2016). However, there is little mention in the literature of the utility of this element or whether this knowledge is maintained post-training. There is also, to the author's knowledge, no previous research examining the influence of a pre-interview assessment on interviewers' practice. In order to conduct a pre-interview assessment of a child, interviewers must possess some knowledge of child development. As this knowledge is deemed necessary to achieve mastery, encouraging interviewers to conduct child assessments could theoretically improve their practice. It would serve to enhance their awareness of children's capabilities and offer an additional opportunity to hone their interviewing skills.

2.2.6 Registered Intermediaries and Pre-Interview Communication Assessments

Fessinger and McAuliff (2020) recently conducted a survey of American child forensic interviewers ($N = 781$). The survey asked the interviewers to identify the types of information they wanted prior to interviewing a child. Fifty-nine percent of the interviewers wanted information about the child's developmental abilities. This information could be gleaned from a pre-interview communication assessment. ABE (MoJ, 2011) advocates conducting an assessment of a child, if it is in the child's best interests, before embarking upon an investigative interview. The primary purpose of an assessment is to determine how best to conduct the interview (Smith & Milne, 2017). Although the nature of an assessment is highly dependent upon a witness' vulnerability (e.g., age, disability, or both), Smith and Milne (2017) provide some examples of factors that may be explored:

- The nature of the vulnerability;
- The witness' first language or preferred means of communication;

- How the witness copes with a change of routine;
- Whether the interviewer is able to establish adequate rapport with the witness;
- The ability of the witness to maintain concentration;
- How the witness responds to trauma and how to potentially manage this;
- The witness' understanding of concepts such as prepositions, temporal concepts, frequency, and sequential concepts;
- Practising the ground rules;
- The witness' ability to draw;
- The witness' ability to recount a neutral event.

ABE (MoJ, 2011) recommends exploring similar factors but does not provide detailed guidance as to how the interviewer should seek to assess these factors. It is therefore unsurprising that police officers have reported being unclear as to what a formal communication assessment should involve (McCullough, 2017). Smith and Milne (2017) suggest engaging the child in an informal conversation or practice narrative. Alternatively, if they have received specialist interview training, they may wish to engage the child in a more extensive assessment task. For example, the Achieving Best Evidence Language Screen (ABELS; Jackson, 2016, as cited in McCullough, 2017) or 'Unpacking the Box' (Triangle, 2015).

ABELS. ABELS (Jackson, 2016, as cited in McCullough, 2017) was developed by a specialist Speech and Language Therapist and RI. ABELS is a picture-based assessment tool, designed for use with children under the age of 11 and teenagers with special educational needs. The tool was developed for use by police officers. It provides a structured framework for addressing the pre-assessment requirements set out in the ABE guidance (MoJ, 2011). The tool uses a traffic light-based scoring system in order to identify any communication difficulties that may impact upon the investigative interview (e.g., the ability to comprehend

ground rules, produce a clear and coherent narrative account, understanding of positional words, ability to draw) and recognise when RI assistance is required.

A recent study by McCullough (2017) looked at practitioners' (i.e., police officers from Norfolk Constabulary) perceptions of the ABELS (Jackson, 2016, as cited in McCullough, 2017). The feedback from the officers was generally very positive. Overall, it was felt that the tool improved their practice. It was thought to have enhanced rapport particularly with the most reticent of children, allowed defensible decisions to be made regarding employing the assistance of an RI, and enabled them to better understand and adapt to children's difficulties resulting in more extensive interview planning. Concerns were raised regarding the supplementary section of the tool which tests children's understanding of complex language and idioms. It was felt that ending the pre-interview assessment with this task could compromise rapport due to the difficulty of the task and the potential for a high proportion of incorrect responses. There were also concerns that the full assessment may be too lengthy and could impact negatively upon the investigative interview. Given the concerns previously voiced by researchers regarding lengthy pre-interview activities (e.g., Davies et al., 2000), this is an area that warrants additional research. The ABELS website (www.abels.org.uk/services/child-abels/) indicates that further evaluations of the tool will be underway in the near future. It is essential that the future research looks at whether the decisions made on the basis of ABELS (Jackson, 2016, as cited in McCullough, 2017) are correct and whether the tool does, quantifiably, improve the quality of child interviews. Neither of these factors were addressed in McCullough's study (2017) and therefore the efficacy of the framework is currently undetermined.

'Unpacking the Box'. 'Unpacking the Box' is an assessment tool developed by Triangle (2015). Triangle is a UK based company that specialises in working with children. 'Unpacking the Box' (Triangle, 2015) provides a structured method of assessing children's

receptive communication; expressive communication; attention, anxiety, and behaviour. More specifically the tool can assess a child's understanding and use of sequencing vocabulary (e.g., before, after), prepositions (e.g., in, on, under), and comparatives (e.g., same, different). It can also test a child's auditory working memory and ToM. The tool consists of a silver box containing small objects for the child to work with and an accompanying guidance manual. Despite having been used in applied settings for a number of years (by police officers, RIs, and NRIs), to date, the only piece of research involving 'Unpacking the Box' (Triangle, 2015) looked at its ability to reliably measure children's working memory capacity (Iranzo, 2016). The study found that children's scores on 'Unpacking the Box' (Triangle, 2015) correlated with their scores on three other well-established working memory tasks: digit span, corsi blocks, and listening span (Iranzo, 2016). The children's level of enjoyment was also found to be higher for 'Unpacking the Box' (Triangle, 2015) compared to the other three tasks (Iranzo, 2016). This suggests that it provides a reliable and engaging method for measuring working memory in young children, aged 6 to 8 years old.

The lack of research regarding this tool is of concern. There appears to be a growing emphasis, in the legal arena, upon evidence-based practice. Practices lacking a strong theoretical framework are coming under increased scrutiny. One example is the '20 principles' underpinning vulnerable witness advocacy training in England and Wales. Some researchers have called for the training to be overhauled and the '20 principles' reformed on the grounds of insufficient empirical support (Cooper et al., 2018). Thus, further research is essential to ensure 'Unpacking the Box' (Triangle, 2015) is a reliable tool.

Police officers should be using a communication assessment to plan and inform the interview (Smith & Milne, 2017). For example, during the assessment the police officers should consider the needs of the child and whether the child would benefit from additional support. If it arises, during the assessment, that the child may be unable to recognise a

problematic question or indicate that a question is problematic an assessment by an RI should be considered.

RIs and the Witness Intermediary Scheme (WIS). Currently, intermediaries are being utilised in Northern Ireland, South Africa, Australia, Japan, England, and Wales. The YJCEA (1999) introduced the intermediary ‘special measure’ in England and Wales. The WIS was created to implement the ‘special measure’. The scheme was first introduced as a pilot project in 2004. However, it is now available in all 43 police forces and CPS areas across England and Wales (MoJ, 2015). To be eligible for RI assistance a witness must be: under the age of 18, suffer from a mental disorder, possess a significant cognitive or social impairment, or have a physical disability that has the potential to diminish the quality of their evidence (YJCEA, 1999). All RIs are specialists in communication who have been recruited, trained, and accredited by the MoJ (Plotnikoff & Woolfson, 2007). All RIs have to attend an initial face-to-face training course delivered across a two-week period (for a more comprehensive description of the current training see Collins and Krähenbühl, 2020). Since its implementation, demand for RIs has rapidly increased (Plotnikoff & Woolfson, 2015) with a 483% increase between 2010 and 2020 (MoJ, 2020b). In 2019/20 the WIS received 6,907 requests for RI assistance which constituted the highest number of requests in a 12-month period since the scheme’s national roll out in 2008. Of these requests, 4677 were made for child witnesses (MoJ, 2020b).

The Role of an RI. The role of an RI is impartial (O’Mahoney et al., 2011) and involves facilitating two-way communication between witnesses and criminal justice practitioners to ensure that communication is as coherent, complete, and accurate as possible (MoJ, 2020a). There are many ways in which RIs can facilitate effective communication. For example, an RI will often brief police officers and court officials on a witness’ needs and limitations, and recommend strategies to maximise the witness’ ability to provide accurate

evidence. The RI may also communicate questions to the witness that are put to them during an interview or cross-examination and request overly complex questions to be rephrased. They also help the witness understand and become familiar with the legal process (Powell et al., 2015). The RI can be present during the ABE interview (MoJ, 2011), suspect identification procedure, pre-court familiarisation visit, memory refreshment of the investigative interview DVD, during the witness' evidence in court, and when the witness is informed of the trial outcome (Plotnikoff & Woolfson, 2013).

The intermediary role is now far wider than initially envisaged. It is important to be aware of the distinction between RIs and NRIs. RIs work through the WIS. As such their practice is regulated by the MoJ (Cooper, 2014). However, when RIs choose to operate outside the scheme's remit (i.e., work with vulnerable witnesses in the family court or defendants) they are referred to as Non-Registered for the purpose of that instruction. Those, not trained by the MoJ, may also operate within this role. The appointment of a NRI is at the judge's discretion as there is currently no statutory framework in place. However, section 104 of the Coroners and Justice Act (2009) is attempting to amend this situation. Section 104 extends the intermediary 'special measure' to vulnerable defendants. Once it is implemented it will be inserted into the YJCEA (1999). Until such time, it is likely that the appointment of NRIs for defendants will continue to cause considerable debate.

Debates have centred on the necessity of having an NRI in attendance for the duration of a trial (some agencies will have this as a requisite of their assistance) and the costs associated with this (Geddes, 2016). Due to rising costs incurred by the CJS, the Criminal Procedure Rules Committee redrafted the Criminal Practice Directions (2015). The Directions state that the appointment of an NRI for a defendant's evidence will "be rare, but for the entire trial extremely rare" (pg. 18). Even when section 104 is implemented there are no guarantees that this position will change. The live link is the only 'special measure' currently

available to defendants by statute. Yet, it is rarely invoked (for a discussion of the barriers to implementation see Fairclough, 2017) – it is unclear whether similar barriers would limit the accessibility of RI support (Fairclough, 2018). This thesis will not delve further into these contentious issues, beyond alerting the reader to the RI scheme that is now active in Northern Ireland. Information about the scheme can be found on the Department of Justice website (www.justice-ni.gov.uk/publications/registered-intermediary-schemes). The scheme accommodates witnesses, victims, defendants, and suspects. Although the scheme has been praised for recognising the utility of RIs at the earliest stages of the justice process (i.e., within police custody), there has been criticism of the circumscribed role of the RI at court (Taggart, 2018). RI assistance has been limited to the defendant's evidence. A recent study by Taggart (2021) compared the experiences of RIs from Northern Ireland and from England and Wales. Significant differences emerged both in relation to the RI's standing in the CJS and in how the RIs experience the role and its demands. For example, the RIs from Northern Ireland perceived their impartiality as absolute whilst the RIs from England and Wales saw their impartiality as more malleable which sometimes led them to enter complex ethical territory (this will be discussed in greater detail throughout the thesis). Due to divergences in the role as well as in how the RIs conceptualise their role, further discussion will be limited to the role of the RI, as it is defined in England and Wales, prior to and during the child interview.

The Role of an RI Prior To and During the Investigative Interview. The Registered Intermediary Procedural Guidance Manual (MoJ, 2020a) outlines the involvement of the RI in a criminal case. It states that when a police officer identifies that a witness would benefit from the assistance of an RI, the officer should submit a Request-for-Service form (RfS) to the National Crime Agency Witness Intermediary Team. The case will then be matched according to the location, skills, knowledge, and experience of the RI, and the needs of the witness. Once the case has been allocated, the RI will contact the police officer,

relatives of the witness, and other professionals to gain a better insight into the witness' needs which will then inform their assessment of the witness. The purpose of the communication assessment is to ascertain whether the witness has the ability to communicate their evidence, whether the assistance of an RI will improve the quality of the witness' evidence, to provide guidance to the police and Advocates on the most effective way of questioning the witness, and to recommend any 'special measures' or adjustments that will enable the witness to give their best evidence. There is no set procedure or guidance on how to conduct the assessment. Although having some flexibility is beneficial as it enables the assessment to be tailored to the witness' communication needs and the RI's specialism, having no framework in place means that crucial areas of communication may be missed by the RI and/or inappropriate topics explored.

Following the assessment, the RI will provide the police officer with either written or oral recommendations as to how best to communicate with the witness during the interview. Recommendations may relate to language used during the interview, the length and complexity of sentences, the types of questions posed to the witness, the room set-up, the frequency and duration of breaks, how to check the witness' understanding, the use of communication aids, and how best to intervene if necessary. During the interview, the RI is not a second interviewing officer. Their role is purely to facilitate communication and to only intervene if required. For example, to indicate the need for a break, to check the witness' understanding, or to rephrase a question.

As noted above, it is the responsibility of the officer in the case to identify whether a child requires the support of an RI. This is somewhat concerning given the infrequency with which officers conduct pre-interview assessments. Potentially, children with communicative difficulties are not being identified and are therefore not being afforded the support they require in order to give their best evidence. This assertion is in line with the findings of the

Criminal Justice Joint Inspection (2014). The inspection found that RIs are being under-utilised. Of the interviews reviewed in the inspection, none involved an RI. This is somewhat worrying given that some of the interviews involved very young children - under the age of 6 years. When case files were reviewed, it emerged that RIs were not being used because of a tendency by some interviewers to underestimate the communication needs of vulnerable witnesses and / or overestimate their own ability. Other studies have similarly reported a disparity in practitioners' confidence in their abilities and their actual skill levels (Goetzold, 2015). A recent unpublished study found that many interviewers do not request the assistance of an RI as the interviewers feel that "they did not need one as they [knew] how to talk with children because it is something everyone does" (Cots & Garrett, 2018, as cited in Aldridge-Waddon, 2019). Another frequently cited reason for not eliciting the services of an RI relates to the implications associated with delaying the interview. Delays were reported in a recent project conducted in County Durham and Darlington, a region in the North East of England (Leake & Jeffels, 2017). The project sought to provide a baseline assessment of CSA practices and procedures. Although RI and police joint working was considered to be effective, there was again evidence that RIs were being under-utilised. Financial constraints and concerns regarding availability were two of the reasons given as to why referrals were not being made.

RI availability was again highlighted as a major concern in a review by the Victims' Commissioner (Newlove, 2018). The review looked at how the WIS operates and how this impacts upon the provision of RIs for vulnerable victims and witnesses. Input was sought from RIs, the CPS, police, and other relevant agencies (e.g., National Crime Agency Witness Intermediary Team and the College of Policing). The review concluded that:

RIs are invaluable in providing communication assistance for vulnerable victims and witnesses, giving them a voice in the criminal justice system (CJS) and in turn, providing them with equality of access to justice. (p. 7)

However, it was found that not all vulnerable victims and witnesses who were entitled to, and would benefit from, RI support received it. This was attributed to there being too few RIs to meet the current demand (Newlove, 2018). In 2017, the MoJ estimated 470 RIs would be required to meet demand. Yet, as of March 2018 there were only 183 RIs on the Register (Plotnikoff & Woolfson, 2019). This has inevitably led to long delays in matching cases - the average waiting time was estimated to be approximately 4 weeks (Newlove, 2018). These long delays resulted in some police officers going ahead with the interview without an RI and, as reported by Leake and Jeffels (2017), deterred others from requesting their assistance in the first place. Although the MoJ is currently in the process of recruiting more RIs, evidence from the review suggests that this will do little to meet demand and reduce waiting time (Newlove, 2018). The review (Newlove, 2018) also unearthed inconsistencies in how vulnerability is assessed by the Police and the CPS. As a consequence, there was found to be considerable disparity across, and within, police forces as to whether an RI was requested. Some forces were reported as using the ABELS (Jackson, 2016, as cited in McCullough, 2017) to determine whether victims require additional support. Both the Police and CPS were also shown to have a lack of awareness regarding the RIs' role and how to work with them effectively. Thus, a recommendation of the review was that it should be mandatory to include in all training on 'special measures' information pertaining to the RIs' role and guidance on collaborative working to ensure that a victim or witness can achieve their best evidence.

RI Research. Despite the WIS having now been active for over a decade, the work of RIs has not been subject to rigorous research (Collins et al., 2017). The research that has been conducted has focused predominantly on perceptions of the role. In 2007, Plotnikoff and

Woolfson conducted an evaluation of the pilot project. Feedback from witnesses and their carers was wholly positive. Carers felt that the RIs successfully facilitated communication, whilst also providing witnesses with the support required to cope with the stress of the legal process. Generally, criminal justice personnel shared the carers' enthusiasm for the scheme. It was felt that it increased access to justice for vulnerable witnesses. However, some challenges did arise. These included difficulties identifying eligible witnesses, misunderstandings of the role, a lack of planning (i.e., conflict between the desire to conduct the interview in a timely manner and conduct a full communication assessment), and a lack of appropriate intervention during questioning (Plotnikoff & Woolfson, 2007).

A more recent evaluation study (Plotnikoff & Woolfson, 2019) found that some of these challenges have yet to be resolved. There appears to still be some confusion regarding the RI's role. More specifically, how it overlaps / complements other roles within the CJS (e.g., independent sexual violence advisor). Until this is resolved, vulnerable witnesses may not receive as comprehensive support and assistance as they require due to the unfounded belief that one professional can fulfil multiple roles. However, of greater concern is that 21% of judges, in the study, reported having dispensed with the RI's presence at cross-examination following the receipt of their report (Plotnikoff & Woolfson, 2019). Some believed the RIs overstated the need for their presence and provided generic recommendations that added little if anything to counsels' understanding (Plotnikoff & Woolfson, 2019). Some judges appeared to overestimate their ability to monitor the child from afar (courtroom to live link room) and overlook the many other ways in which the RIs' presence can contribute to best evidence (e.g., improves child's confidence). A recent study by Collins and Krähenbühl (2020) emphasised the importance of a judge's support for ensuring that an RI's recommendations are adhered to. The study involved semi-structured interviews with 17 practicing RIs. Although some reported having positive experiences with judges they had worked with,

others reported working with judges who were very dismissive of their assessment findings and recommendations. Unfortunately, some of the RIs also reported a lack of cooperation from barristers. This was felt to have a negative impact on the RIs' ability to fulfil their role, ultimately compromising the quality of the child's evidence.

A number of surveys (e.g., Cooper, 2009; Cooper, 2011; Cooper, 2012; Cooper, 2014) have also been completed by the RIs themselves. Challenges identified by the RIs included a lack of awareness of the role, late referrals (i.e., post interview), and legal professionals failing to fully discuss their recommendations (more prominent within a court environment) and then subsequently not adhering / enforcing them. Similar challenges were reported in a recent study, by Agneswaran (2018), in which 12 RIs were interviewed about their experiences of working with adults in the CJS. The RIs again spoke of a lack of awareness but also of feelings of isolation and the need for additional support. Despite these challenges the RIs did report feeling as if they were becoming more accepted by the CJS (Agneswaran, 2018). In both Agneswaran (2018) and Cooper's (2009; 2011; 2012; 2014) research the work of RIs had been commended by other professionals including police officers, barristers, and judges. Police officers have reported learning from and adopting different interview strategies as a result of working with RIs (Plotnikoff & Woolfson, 2015). Positive experiences of collaborative working have been echoed in Northern Ireland, with the initial review of the project deeming RIs as an "integral part of the justice process" (Department of Justice, 2015, p. 4). Further research related to the work of intermediaries has been conducted in South Africa (Matthias & Zaal, 2011). Although, internationally the roles differ slightly, the onus on facilitating communication remains. Despite implementation difficulties in South Africa, intermediaries have proved highly effective and are thought to have enhanced the fairness of legal proceedings (Matthias & Zaal, 2011).

There is also an emerging body of research which has looked at how the presence of an RI impacts upon perceptions of a child witness and the legal process. Ridley et al. (2015) presented a mock interview transcript either with or without RI interventions to barristers, police officers, and mock jurors. The presence of the RI led to more positive ratings pertaining to interview quality. However, ratings of the witness' honesty, credibility, confidence, and the completeness of their account were unaffected. This study was criticised on the grounds that it used transcripts as opposed to video evidence as would be the standard procedure with ABE interviews (Collins et al., 2017). A more recent study by Collins et al. (2017) presented mock jurors with a video-recorded cross-examination of either a 4- or 13-year-old child (the child was answering questions about the content of a cartoon), with or without an RI. During the cross-examination the RI intervened on five occasions and the questions were rephrased. When the RI was absent no interventions or rephrasing of questions occurred. When the RI was present the jurors rated the quality of the cross-examination and the child's behaviour higher than when there was no RI. Smethurst and Collins (2019) reported similar findings when mock jurors were asked to rate the quality of a defendant's evidence either with or without an NRI. In the NRI condition the NRI intervened when the questioning was inappropriate and the barrister was thus instructed by the judge to rephrase the question. In the no NRI condition the defendant's evidence proceeded without any interruptions. When the defendant was accompanied by an NRI the defendant was perceived as less anxious and more confident, cooperative, and coherent compared to when the defendant was unsupported.

However, these findings have not been consistent across studies. Krähenbühl (2019) conducted a study in which adult mock jurors watched a mock cross-examination of a child witness (this was based on an original cross-examination of a 6-year-old child giving evidence regarding an allegation of physical abuse). The study found that RI presence and / or the

inclusion of interventions had no impact upon mock jurors' ratings in relation to the quality of cross-examination (i.e., child credibility, child understanding, and legal professional's behaviour) but did impact upon their ratings in relation to trial progression. The mock jurors' ratings of trial progression (i.e., expectations of a guilty verdict) were highest when an RI was present but did not intervene and when an RI was absent but interventions were made by the judge. Krähenbühl (2019) thus concluded that jurors may require further instruction during proceedings to fully understand the implications of the RI's presence and / or their interventions. The conflicting findings, of the aforementioned studies, could be attributed to methodological differences (e.g., the type of evidence presented).

To date only one study has explored the impact of RIs on children's communication during investigative interviews (Henry, Crane, et al., 2017). Henry, Crane, et al. (2017) examined the impact of three interventions (RIs, sketch reinstatement of context, and visual labels) designed to improve interview performance. The study examined how each intervention impacted upon the communication of typically developing (TD) children and children with autism spectrum disorder (ASD). TD children were found to perform significantly better in the RI interview compared to the baseline best practice interview. TD children in the RI interview reported 18.96 more items of correct information than children in the best practice interview. This increase in information was not found to jeopardise the children's accuracy. Further research has also found that children, who receive RI assistance, are more accurate when identifying a perpetrator from an identification parade (Wilcock et al., 2018).

Contrary to the researchers' expectations the RI intervention, in Henry, Crane, et al. (2017), did not improve the interview performance of children with ASD. Henry et al. (2017) offer a number of explanations to account for the lack of effect. They assert that the RI provision may offer different benefits for children with ASD compared to TD children.

Benefits may include informing the police about ASD, facilitating and developing rapport, familiarising the child with the investigative process, and adapting the interview environment accordingly. These aspects of the role were not addressed in Henry, Crane, et al.'s (2017) study. This, however, is not the only limitation of the research. The research has undergone considerable scrutiny and the findings vehemently contested by some researchers. Dando et al. (2018) argue that the study failed to test the validity of RIs, Sketch-RC, or verbal labels for supporting children with ASD in real-world forensic interviews. Four main reasons were provided to support this claim. First, the authors argue that the experimental paradigm lacked ecological validity. The mock crime events were very minor (i.e., movement of phone or keys) and are therefore unlikely to reflect a salient and traumatic incident such as CSA. The events were also presented to the children via different mediums (i.e., live and video) leading to distinctly different encoding experiences. Second, the authors question the rationale for including an initial interview prior to the ABE. They argue that this does not reflect practice and would render the interventions less effective as they are designed to support recall at first retrieval. Third, it is queried as to whether the interviewers used in the research had the appropriate training for conducting child interviews. It is also unclear as to the experience and expertise of the RIs with regards to working with children with ASD. Finally, the authors question the decision to transform some of the data to meet assumptions of normality. They argue that as a result of this the variance associated with ASD may have been hidden which may have reduced the likelihood of finding effects that would be of relevance to practitioners. Dando et al. (2018) conclude that rejecting RIs, Sketch-RC, or verbal labels on the basis of Henry, Crane, et al.'s (2017) findings would be unjustified and potentially damaging if used in legal proceedings to undermine the testimony of vulnerable victims and witnesses. Given Dando et al.'s (2018) significant criticisms, more research is warranted into the efficacy of

each of these techniques - research that either better reflects practice or involves real-world interviews.

Despite the obvious constraints of Henry, Crane, et al.'s (2017) study one more aspect warrants further discussion. Children in the RI condition were assessed and then accompanied in the interview by one of two experienced, practising RIs. Although both RIs were deemed as 'experienced', a significant difference was found between the number of correct details recalled by the TD children dependent upon the RI involved. One had more of a beneficial effect on recall than the other, with children reporting on average 62.21 pieces of correct information compared to 44.63 respectively. No differences were found in relation to incorrect or confabulated details. Nevertheless, the disparity between the two RIs in relation to correct information is of concern. The findings suggest that practice is not standardised and that there is possibly considerable variation in practice. The RIs interviewed in Agneswaran's (2018) research acknowledged this variability and were concerned with how this impacts upon police and legal practitioners' attitudes towards them. This variability definitely warrants further exploration upon a larger scale. It is important, in any profession, to find out 'what works'.

Nevertheless, recent studies have produced findings that should increase confidence in RIs' abilities. A recent study by Henry et al. (2021) examined whether the presence of an RI has an impact upon children's resilience to misleading questions during cross-examination. Henry et al.'s (2021) study involved 6 to 11-year-old children witnessing a staged event (i.e., a mock crime). One week later the children were interviewed about their experiences. Following an 8 to 13 month delay the children were cross-examined by one of several experienced barristers and their accounts challenged using a defence statement from a mock perpetrator. In the RI condition, the RIs re-assessed the children (in order to gather up-to-date and accurate information regarding their communication needs) and a ground rules hearing

was conducted between the barristers and RIs prior to the cross-examination. The RIs then accompanied the children during the cross-examination, simplifying the instructions provided by the judge, and intervening if the barristers' questions did not adhere to the recommendations discussed during the ground rules hearing. Overall, 94% of the children complied with at least one of the barristers' false suggestions. However, the children that were assisted by an RI, during cross-examination, were less likely to comply with false information given by the barristers. This could, in part, be due to the barristers' questions in the RI condition generally aligning more closely with those advocated in best practice guidance. Henry et al.'s (2021) findings suggest that the presence of an RI does improve the accuracy of children's testimony thus demonstrating the efficacy of the role.

However, there are a number of limitations of Henry et al.'s (2021) study. The first limitation relates to the ecological validity of the study. Due to ethical reasons, it was not possible for the study to replicate the unfamiliarity, anxiety, and trauma associated with a real court case. As such the event the children witnessed was relatively innocuous and the barristers described as both approachable and experienced (all of the barristers were chosen due to having previous experience in cross-examining children). A further difficulty when interpreting the findings of the current study is that the study fails to state when the ground rules hearing took place (i.e., before all of the cross-examinations took place or just prior to those involving the RIs). If the ground rules hearing took place before all of the cross-examinations, then the barristers may have been applying the recommendations to the children in all of the conditions. This would suggest that any significant differences that did emerge were due directly to the RIs presence and perhaps a greater sense of scrutiny / social pressure on the barristers' behalf. Conversely, if the ground rules hearing took place only prior to the RI assisted cross-examinations then the ground rules hearing may have also contributed to the barristers' improved practice.

Positive findings were also reported in a study by Hanna and Henderson (2018) which compared lawyers (from New Zealand) and RIs' (from England / Wales) perceptions of what constitutes developmentally appropriate language. The participants each assessed an anonymised transcript of an 11-year-old child's courtroom questioning. Although the lawyers and RIs were largely in agreement as to what categories of language might confuse children (e.g., word choice, unclear reference, multiple questions, negation, and lengthy / complicated questions), they varied in their perceptions of what constituted admission into each category. Compared to the lawyers, the RIs identified more potentially problematic vocabulary, subtle multiple questions, unclear references, questions involving negation, and the use of the passive voice. The RIs were also more successful than the lawyers at rephrasing questions as the lawyers had a tendency to incorporate new complexities. These findings would suggest that RIs are both aware and receptive to the needs of young witnesses.

This is supported by the work of Owen (2016) who found that generally intermediaries (from UK and Australia) are aware of the needs of vulnerable witnesses and are able to highlight the challenges this population faces when entering into the CJS. RIs have also been shown to regularly use resources such as The Advocates Gateway (i.e., a free online resource for those working with vulnerable witnesses and defendants in the CJS) to inform their practice. Although this is really positive, a thorough knowledge of best practice interview guidance is probably equally, if not more, important to their role. Krähenbühl (2011) sought to examine whether intermediaries have knowledge of ABE (MoJ, 2011). She examined the intermediaries' perceptions of appropriate communication in the context of an interview and cross-examination. The intermediaries were shown mock interview transcripts and were asked to indicate which questions they believed to be inappropriate, state what made the questions inappropriate, and provide a more appropriate alternative. Although the intermediaries demonstrated an awareness of developmental and emotional factors, a lack of knowledge

emerged pertaining to best practice guidelines. For example, many of the alternatives provided by the intermediaries were not open-ended. This is not in line with the ABE guidance (MoJ, 2011), which advocates the use of predominantly open-ended questions. The intermediaries also criticised three-part forced-choice questions for having too many options, despite the guidelines stipulating that forced-choice questions should contain three parts, one of which should be informing the child of the permissibility of an ‘I don't know’ response. These findings are somewhat disconcerting given that an intermediary’s primary rule is to facilitate communication within the CJS. It would thus be assumed that a knowledge of the guidance governing this practice would be essential. Given Krähenbühl’s (2011) findings it therefore comes as no surprise that the role has been met with some scepticism.

Concerns Associated with the RI Role. There are a number of concerns surrounding intermediary schemes. One concern relates to the ability of RIs to maintain impartiality. Critics fear that there is a risk that an RI could distort the evidence or interfere with the trial (Agnew, 2006). However, there is no empirical evidence to support this claim. In fact, RIs have been reported to provide very good support whilst maintaining neutrality and impartially (Leake & Jeffels, 2017). Another concern is that of training. It has been argued that the training RIs receive is insufficient and that a short training course may not be substantial in establishing and ensuring expertise (Powell et al., 2015). It is likely that this argument has been further fuelled by the Victims’ Commissioner’s report (Newlove, 2018) which concluded that there is a “lack of overall management and governance of the WIS” (p.7). The most pertinent issues associated with this being a reduction in the provision of continuing professional development for RIs, a lack of funded mentoring, no funded clinical supervision, and insufficient quality assurance procedures.

In terms of quality assurance procedures, the RI Quality Assurance Board may want to consider more closely monitoring the work of RIs. This could potentially be achieved by each

RI submitting, on a yearly basis, a copy of an ABE interview for feedback, along with a reflection log (i.e., a record which demonstrates that the RI has spent time critically appraising their own practice). Research, involving police interviewers, has shown that regular feedback (which is what this initiative would offer) can improve practice (see section 2.2.5). It could also help to identify and resolve any widespread shortcomings, such as a lack of knowledge pertaining to the best practice guidelines, and ensure that subsequent training is planned in accordance with this. Furthermore, it is suggested that intermediaries possess insufficient knowledge of the best practice guidelines. A detailed knowledge of the guidelines is required in order to understand the underlying principles used to elicit accurate and complete testimony from a child (Krähenbühl, 2011).

Although other countries have considered the implementation of an intermediary scheme, these and other concerns have created some trepidation. Powell et al. (2015) interviewed criminal justice practitioners in Australia. The practitioners voiced concerns regarding the implementation of such a scheme. They feared that the implementation of an intermediary scheme would be premature given the gap that currently exists in Australia between best practice interview guidelines and the questions used to elicit information during interview and trial. Without addressing these underlying problems, it was felt that the intermediary scheme would serve merely as a 'Band-Aid' reform and would have little benefit. Practitioners also raised concerns specifically related to the intermediary role. These centred around three issues: difficulties establishing and maintaining professional competency, possible conflicts that may arise due to the dual-purpose of the role (i.e., supportive vs. communicative), and detrimental effects associated with another party being introduced into the system which may result in delays and exacerbate witness anxiety. Despite such concerns a pilot scheme did take place in Australia (Owen, 2016). An evaluation of the pilot found widespread support for the use of witness intermediaries. The witness

intermediaries were thought to reduce children's stress at court and improve the quality of their evidence. The witness intermediaries were also seen as playing an educational role, increasing the knowledge of police officers and lawyers in regards to developmentally appropriate questioning (Cashmore & Shackel, 2018). Given this positive response it is not surprising that the scheme has continued to operate following the review of the pilot.

Overall, the body of work regarding RIs is very limited, with the majority of studies having simply looked at perceptions of the role, as opposed to the direct impact of the RI provision upon practice. The current thesis will attempt to address this gap in the literature. The thesis comprises of four studies designed to examine different facets of the RI role including the RIs knowledge of children's memory and current best practice guidance, the efficacy of pre-interview communication assessments, individual difference factors related to children's event recall, and the impact of RIs on practice in real-world investigative interviews. The studies utilise a variety of methodological approaches (i.e., online questionnaire, experimental laboratory research, and the examination of real-world interviews) in order to provide a more holistic evaluation of RI practice.

Chapter Three: Registered Intermediaries' Beliefs about Children's Memory and Investigative Interview Practice (Study 1)

Eyewitness reports are a central component of many criminal investigations (Kebbell & Milne, 1998) and can be instrumental in solving criminal cases (Dodier, Tomas, et al., 2019). However, these reports, due to the fragility of memory, are also one of the main causes of wrongful convictions (Innocence Project, 2015). As such it is recommended that law professionals stay up to date with the latest advancements in memory research and are aware of the factors that can influence memory in a criminal and judicial context (Dodier, Tomas, et al., 2019). Yet, in a recent study, by Dodier, Tomas, et al. (2019), French police officers were found to have a poor understanding of these factors and held a number of erroneous beliefs regarding witness memory. However, police officers are not alone in their lack of knowledge. Previous research has found that many professionals, working within CJSs internationally, have a limited understanding of and hold incorrect beliefs regarding witness memory (e.g., Dodier, Melinder et al., 2019; Erens et al., 2020; Granhag et al., 2005; Jiang & Luo, 2016). For example, Granhag et al. (2005) examined Swedish legal professionals' (i.e., police officers, prosecutors, and judges) beliefs regarding eyewitness testimony. The participants were presented with 13 items related to eyewitness testimony and for each item they had to indicate which of the four alternatives best reflected their opinion on the topic (e.g., "children's testimonies tend to be *less / more / as* reliable as adults" or "don't know"). For some items, the legal professionals' beliefs were in line with empirical research (e.g., weapon focus, completeness of children's accounts). However, for other items some of the professionals' expressed beliefs which were in contrast with the current research findings (e.g., simultaneous vs. sequential lineups⁵, forgetting curve). Judges, for example, were very skeptical about the reliability of children's testimony. This view, however, was far less prevalent amongst the police officers and prosecutors. Not only were the different

professionals found to hold different beliefs, but professionals within the same occupational group had differing opinions. This raises serious questions regarding objectivity and equity within the CJS.

Similar findings were reported by Jiang and Luo (2016) who examined legal professionals' (i.e., judges, prosecutors, police officers, and defense attorneys) knowledge of eyewitness testimony in China. The legal professionals indicated on a 4-point scale from 1 (strongly disagree) to 4 (strongly agree) their beliefs about 12 statements related to eyewitness testimony. The proportions of correct responses were 57% for both the prosecutors and police officers, 58% for the judges, and 61% for the defense attorneys. The four groups of professionals differed in their beliefs regarding four items including the impact of stress (it is not surprising that the professionals' beliefs differed in respect to this item given the mixed research findings on this topic, see section 3.1.3 for further discussion) and the impact of exposure time on eyewitness recall. There was also very little consensus within each professional group – only four items were considered to have elicited a rough unanimous response (i.e., exceeding 75% agreement). Of those items, one was related to child suggestibility. Seventy-five percent of judges and 78% of defense attorneys held the correct belief that children are more likely than adults to be influenced by interviewer suggestion, peers, and other social influences (prosecutors and police officers did not meet the threshold

⁵The legal professionals held the belief that the risk of inaccurate identifications is lower for simultaneous as opposed to sequential line-ups. This was deemed by Granhag et al. (2005) to be incorrect based on the research evidence available at the time of the study. However, more recent research has yielded conflicting findings (e.g., Collof et al., 2022; Seale-Carlisle et al., 2019). For example, Seale-Carlisle et al. (2019) demonstrated the superiority of simultaneous line-ups. They concluded that simultaneous line-ups lead to more correct and fewer incorrect identifications.

for a unanimous response with 67% and 64% correct respectively). The percentage of correct responses pertaining to child witness accuracy was comparatively low, ranging from 45 – 58%.

A more recent study by Erens et al. (2020) examined the beliefs of professionals that work at Safe Home about memory and forensic interviewing practices (Safe Home is a Dutch hotline where professionals, such as teachers and sports coaches, as well as citizens can report their concerns about domestic violence or child abuse). The professionals from Safe Home comprised of social workers, behavioral scientists, and medical doctors. All the professionals completed an online questionnaire which required them to agree or disagree with nine statements related to eyewitness memory and one statement related to questioning techniques used within child investigative interviews. For the 10 statements, correct responses ranged from 16-94%. Thus, some beliefs were in keeping, whilst others ran counter, to current scientific evidence. For example, most professionals were shown to hold accurate beliefs regarding the development of false memories and the influence of suggestion. Yet far less responded correctly to the statements regarding repressed memories and appropriate questioning practices during child forensic interviews. That being said, a statistically significant difference between the behavioral scientists and social workers was only found for 1 of the 10 statements, suggesting that beliefs were relatively consistent amongst these two groups of professionals.

A further study by Dodier, Melinder, et al. (2019) examined a different group of professionals who work within the legal system – expert witnesses. Dodier, Melinder, et al. (2019) used a 12-item questionnaire to compare French and Norwegian expert witnesses' knowledge about memory. The French and Norwegian expert witnesses both had a limited knowledge of memory and the factors that could distort it. On average, the Norwegian expert witnesses answered 7 of the 12 questions correctly whilst the French expert witnesses only

answered 5. The Norwegian experts outperformed the French experts on seven items including the item pertaining to children's recall with correct responses of 57.4% and 27.5% respectively.

Further studies that have examined the memory-related knowledge of legal professionals have also incorporated a sample of people, who do not work within the CJS, for comparison. Kask (2011) compared the knowledge of Estonian legal professionals (i.e., judges, prosecutors, preliminary investigators, and juvenile police officers) and lay persons. All were asked to indicate whether they agreed, disagreed, or were unsure about 32 statements related to eyewitness testimony. For many of the statements, the beliefs of the legal professionals were in line with current research on the topic. However, there were six statements whereby the opinions of the legal professionals differed from those of the research base (e.g., minor details, confidence-accuracy⁶, suspect lineup format). Kask (2011) also found very few differences in the beliefs of the different legal professionals. Differences emerged in relation to five statements, predominantly related to suspect identification procedures. When the beliefs of the legal professionals were compared to lay persons, the percentage of erroneous responses was equal across the two groups. Yet the lay persons were found to hold more beliefs that were in line with the research base than the legal professionals, as the legal professionals were more likely to air on the side of caution opting for a 'don't know' response.

⁶The legal professionals agreed with the statement 'at trial, an eyewitness's confidence is a good predictor of his or her accuracy in identifying the suspect'. This was deemed by Kask (2011) to be incorrect based on the research evidence available at the time of the study. However, more recent research has found that identifications made with high confidence are more likely to be accurate than those made at lower levels of confidence (Seale-Carlisle et al., 2019; Wixted & Wells, 2019). That being said, concerns have been raised regarding putting too much weight on the confidence-accuracy relationship when evaluating the veracity of claims in real-world cases (Berkowitz et al., 2022). The relationship is dependent upon a number of factors including age, lineup characteristics, and the witness' facial recognition skills (Berkowitz et al., 2022).

Houston et al. (2013) also examined knowledge of eyewitness memory in Scottish judges and jury-eligible members of the public. The judges and members of the public shared comparable levels of knowledge. Sixty-seven percent of the judges' and 61% of the public's responses, to a multiple-choice questionnaire, were consistent with the research base. For the individual items on the questionnaire, correct responses (i.e., based on previous research) ranged from 40% - 97% for the judges and 31%-92% for the members of the public (with correct responses for the item relating to child suggestibility at 77.3% and 63.5% respectively). An earlier study conducted in the US by Wise and Safer (2010) compared the memory-related knowledge of judges, undergraduate, and advanced law students. All the participants completed a 12-item knowledge scale. The law students (66%) answered significantly more items correctly than either the undergraduate students (58%) or the judges (55%). The judges and undergraduate students did not differ from one another in their knowledge of eyewitness memory.

Magnussen et al. (2010) elicited slightly more promising findings. Magnussen et al. (2010) compared the responses of Norwegian judges and jurors⁷ to 12 statements related to eyewitness testimony. For the 12 items, correct responses ranged from 15-89% for the jurors and 31-98% for the judges. Although the judges scored higher than the jurors for 10 of the 12 items on the questionnaire, the judges' knowledge of eyewitness memory was still shown to be lacking in certain areas (e.g., forgetting curve). Magnussen et al. (2010) also compared the responses of the jurors to the general public. Comparisons were only made in respect to 7 items. No statistically significant difference was found between the knowledge of the two

⁷ For serious cases in Norway the Court of Appeal uses 10-person lay juries ('lagmannsretten'). Jurors are appointed for 4 years and often serve on multiple trials. The rationale for comparing jurors and the general public was to look at the effect of trial experience upon knowledge.

groups. On average 56% of the general public and 53% of jurors gave a correct response to each item.

Overall, the findings of previous studies suggest that the knowledge of legal professionals regarding eyewitness memory and the factors that can affect it is not sufficient, and is on some occasions no better than that of lay people (e.g., Houston et al., 2013; Wise & Safer, 2010). Given how important eyewitness reports are in many criminal investigations, particularly those involving child abuse (Wilcock et al., 2006), this conclusion is very concerning. A lack of knowledge on behalf of legal professionals could lead to poor decisions being made throughout the criminal justice process, potentially undermining the quality of and weight given to a witness' evidence (Dodier, Tomas, et al., 2019). Thus, it is essential that future research examines the memory-related knowledge of all those working within the CJS, so that training and initiatives to address any gaps in knowledge can be put in place. To the author's knowledge, no previous research has examined the knowledge of RIs despite their active role within legal proceedings, in England and Wales, for over a decade. This research is important as concerns have been raised regarding RIs' understanding of the factors (specifically communication aids) which can impact upon children's recall (Collins & Krähenbühl, 2020).

The Current Study

This exploratory study examined whether RIs' beliefs regarding children's memory and investigative interview practice differed from those of lay people. The purpose of the study was to identify any gaps in the RIs' knowledge which could be used to inform the content of future training and CPD opportunities. This was achieved by both RIs and lay persons completing an online questionnaire. The questionnaire comprised of 20 items related to eyewitness memory and best practice interview techniques. The items were all based upon the findings of previous research and current best practice guidance (i.e., ABE; MoJ, 2011). A

total knowledge score (i.e., number of correct responses) and an erroneous belief score (i.e., number of incorrect responses) were calculated for each participant and compared across groups. It was hypothesised that:

- RIs would score higher on the total knowledge score (total number of responses that align with current research findings) and lower on the erroneous belief score (total number of responses that conflict with current research findings) than lay people. As such, RIs would also be more likely to answer each of the individual items correctly and provide less ‘uncertain’ responses. This is due to an important part of the RIs role being to facilitate communication in investigative interviews and thus RIs should have an understanding of best practice and children’s memory capabilities.

3.1 Method

3.1.1 Design

The study was an experimental independent measures design. The independent variable was participant with two levels: RI and lay person. The dependent variables were the total knowledge score (i.e., number of correct responses for each participant), erroneous belief score (i.e., number of incorrect responses for each participant), number of ‘uncertain’ responses, and response to each individual item (i.e., correct, incorrect, or ‘don’t know’).

3.1.2 Participants

The questionnaire was distributed to RIs via the RI National Regional Coordinator. The Coordinator sent an email with a link to the questionnaire to all Regional Leads who then disseminated this to all the RIs in their respective areas. A link to the questionnaire was also shared on Registered Intermediaries Online (an online forum for RIs) and with Triangle (a company that employs RIs). Thirty-two RIs completed the questionnaire (approximately 30% of active RIs in England and Wales, registered to accept cases involving witnesses under the age of 18; Plotnikoff & Woolfson, 2019). All the RIs were currently listed on the MoJ

Register in England and Wales as accepting appointments for witnesses under the age of 18. The mean age of the RIs was 53.69 years with an age range of 30 to 78 years. Of the RIs, 96.9% ($n=31$) were female and one participant chose not to disclose their gender; 59.4% ($n=19$) had a background in speech and language therapy, 31.3% ($n=10$) in teaching, and 9.4% ($n=3$) in 'other' (i.e., police officer, youth worker, early years). The RIs reported having worked within the role for between 1 to 14 years with a mean length of time in the role of 4.68 years. The mean number of referrals the RIs had accepted for witnesses under the age of 18 was 134.41 (range = 12 to 500). The RIs had supported witnesses under the age of 18 at a mean of 123.94 interviews (range = 12 to 500) and 33.03 trials (range = 0 to 309).

A link with the questionnaire was distributed to lay people via social media networks (i.e., Facebook). Seventy-one people completed the questionnaire. However, 10 of these were excluded from the final sample by virtue of their occupation as they could not be deemed as lay people (e.g., police officer, lawyer, psychologist). Of the 61 lay persons included in the final sample, 75.4% ($n=46$) were female and 24.6% ($n=15$) were male. The lay person sample ranged in age from 23 to 69 with a mean age of 48.38 years.

3.1.3 Materials

There were two versions of the questionnaire: RI and layperson. Both versions included socio-demographic questions (i.e., age, gender). The layperson questionnaire also included a question about occupation in order to determine whether the person may possess any prior knowledge / expertise about the topic, and the RI questionnaire included questions related to the RI's experience (i.e., number of years in the role and number of referrals). The 20 multiple-choice items associated with children's memory and investigative interview practice were identical across the two versions of the questionnaire. Nine of the items were taken from Dodier, Tomas, et al. (2019) who based their questionnaire on Magnussen and Melinder (2012), Melinder and Magnussen (2015), and Dodier and Payoux (2017). Although

the item relating to the impact of stress was taken from the aforementioned studies, the response considered to be most in line with current empirical evidence was amended from ‘agree’ to ‘disagree’. Previous studies (e.g., Magnussen & Melinder, 2012) have cited a meta-analysis by Deffenbacher et al. (2004) as the evidential basis for this item. The meta-analysis concluded that high levels of stress have a negative impact upon the accuracy of eyewitness memory. However, the meta-analysis found that heightened stress impacted upon the recall of adult witnesses but not children which is the focus of the current study. Furthermore, a number of field studies, have yielded conflicting findings to Deffenbacher et al. (2004) in respect to adults, with memory for stressful events (e.g., bank robbery, assault, shooting) generally found to be accurate (e.g., Christiansen & Hubinette, 1993; Odnot et al., 2009; Woolnough & MacLeod, 2001; Yuille and Cutshall, 1986). It is, however, acknowledged that the impact of stress on eyewitness memory is a controversial topic (for a review and discussion of the disparate findings see Marr et al., 2021). One item was adapted from Magnussen and Melinder (2012). The remaining 10 items were added due to their relevance to the RI role and current best practice interview guidance. The 20 statements and response options are presented in Table 1. The answer that is considered to be correct based on current empirical evidence is indicated by an asterisk. The empirical evidence supporting each item is also provided in Appendix A.

3.1.4 Procedure

The questionnaire was run using JISC Online Surveys – a tool used for developing and distributing online questionnaires. First the participants were required to read the participant information. The participants were then asked to check the boxes to confirm that they met the inclusion criteria for the study and gave their consent to take part. The inclusion criteria for lay people were that they must be over the age of 18 and live in England or Wales. The inclusion criteria for the RIs were that they needed to be listed on the MoJ Register in

England and Wales and accept appointments for witnesses under the age of 18. The participants were then asked a number of demographic and employment / experience questions. Following this the participants were presented with the 20 items related to children's memory and investigative interview practice (see Table 1). For each item, the participants were asked to select the response which best reflected their belief regarding the statement. Once the questionnaire was complete the participants were presented with a debrief form.

Table 3.1*Questionnaire Items*

Name of Item	Statement
Multiple Witnesses	Children who witness the same event will recall it differently. Agree* Disagree Uncertain
Effect of Post-Event Information	Eyewitness testimony about an event often reflects not only what a witness actually saw but also information obtained later on from other witnesses, the police, the media, etc. Agree* Disagree Uncertain
Minor Details	A witness's ability to recall minor details about a crime is a good indicator of the accuracy of the witness's identification of the perpetrator of the crime. Agree Disagree* Uncertain
Confidence-Accuracy**	An eyewitness's confidence is a good predictor of his or her accuracy in identifying the defendant as the perpetrator of the crime. Agree Disagree* Uncertain
Impact of Stress	Very high stress at the time of observation has a negative effect on the accuracy of testimony. Agree Disagree* Uncertain
Attitudes and expectations	An eyewitness's perception and memory of an event may be affected by his or her attitude and expectations. Agree* Disagree Uncertain
Weapon Focus	The presence of a weapon can impair an eyewitness's ability to identify the perpetrator's face accurately. Generally true* Generally false Uncertain
Forgetting Curve	The rate of memory loss for an event is greatest right after the event and then levels off over time. Generally true* Generally false Uncertain
Children's Recall	When small children talk about events they have experienced, do you think they remember better, as well as, or worse than adults? Better As well as Worse* Uncertain
Dramatic Events	Sometimes people become witnesses to dramatic events. Do you think the memory for such events is worse, as good as, or better compared to the memory for everyday events? Better* As good as Worse Uncertain
Immediate Acceptance of Suggested Information	An affirmation to a suggestive question asked by a professional does not necessarily mean that the witness remembers the suggested information. Agree* Disagree Uncertain

Name of Item	Statement
False Memories	Under certain circumstances older children are more likely than younger children to recall false information. Agree* Disagree Uncertain
Response Bias	Young children can possess response biases (e.g., indiscriminately respond 'yes') to certain types of questions. Agree* Disagree Uncertain
Anatomical Dolls	Anatomical dolls can lead to inaccuracies or distortions in a child's account. Agree* Disagree Uncertain
Communication Aids	Drawing is generally a safer method than either dolls or body diagrams for eliciting accurate information from children in an investigative interview. Agree* Disagree Uncertain
Question Types	According to current best practice guidance (i.e., ABE), open-ended questions should be used predominantly during an investigative interview. These questions generally elicit the most accurate accounts. Agree* Disagree Uncertain
Open-Ended Questions	'Can you tell me what happened?' is an example of an open-ended question. Agree Disagree* Uncertain
Recall vs. Recognition Memory	Open-ended questions utilise recall memory processes whilst forced-choice questions (e.g., Did you have tea or coffee?) utilise recognition memory processes. Agree* Disagree Uncertain
Repeating Questions	Repeating who / what / when / where / why questions verbatim is recommended practice as it can elicit more information from reticent children. Agree Disagree* Uncertain
Children's Narratives	Children's false narratives can be as detailed and coherent as true narratives. Agree* Disagree Uncertain

*Correct response based on empirical evidence / previous research

** This item should be viewed with caution. 'Disagree' was chosen as the correct response for this item as the confidence-accuracy relationship has been found not to hold for young children (Berkowitz et al., 2022). Previous research has shown that children (under 11 years old) tend to be overconfident in their inaccurate judgements (Brackman et al., 2019; Brewer & Day, 2005)

3.2 Results

Knowledge Score

An independent-samples t-test was run to determine if there were differences in knowledge score between RIs and laypersons. There were no outliers in the data as assessed by inspection of a boxplot. Knowledge scores for the two groups (RI and layperson) were normally distributed as assessed by a Shapiro-Wilk test ($p > .05$) and there was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .083$). The knowledge score was higher for laypersons ($M = 11.51$, $SD = 2.36$) compared to RIs ($M = 10.19$, $SD = 3.03$), a statistically significant difference was found, $M = -1.32$, 95% $CI [-2.45, -0.19]$, $t(91) = -2.32$, $p = .023$, and represented an effect of $d = .236$.

Erroneous Belief Score

An independent-samples t-test was run to determine if there were differences in erroneous belief score between RIs and laypersons. There were no outliers in the data as assessed by inspection of a boxplot. Erroneous belief scores for the RI group were normally distributed as assessed by a Shapiro-Wilk test ($p > .05$), but the scores for the layperson group violated this assumption ($p = .020$). Running an equivalent non-parametric test had no appreciable effect on the findings thus the decision was made to report the results of the t-test. There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .256$). There was no statistically significant difference between the erroneous belief scores of the laypersons ($M = 4.43$, $SD = 1.98$) and the RIs ($M = 3.75$, $SD = 1.69$), $M = -0.68$, 95% $CI [-1.49, -0.14]$, $t(91) = -1.65$, $p = .103$.

Expression of Uncertainty

An independent-samples t-test was run to determine if there were differences in expressions of uncertainty between RIs and laypersons. There were no outliers in the data as assessed by inspection of a boxplot. Expressions of uncertainty for the RI group were

normally distributed as assessed by a Shapiro-Wilk test ($p > .05$), but the scores for the layperson group violated this assumption ($p = .010$). Running an equivalent non-parametric test had no appreciable effect on the findings thus the decision was made to report the results of the t-test. There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .288$). Expressions of uncertainty were more common for the RIs ($M = 5.97$, $SD = 2.91$) compared to the laypersons ($M = 4.02$, $SD = 2.56$), a statistically significant difference was found, $M = 1.95$, 95% $CI [0.79, 3.12]$, $t(91) = 3.33$, $p = .001$, and represented an effect of $d = .330$.

Questionnaire Items

Chi Square tests were conducted to examine beliefs for each of the 20 items of the questionnaire individually. Where 20% of cells had an expected count less than 5 exact significance tests were run.

Multiple Witnesses

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to multiple witnesses, $\chi^2(2, N = 93) = 1.11$, $p = .608$, Cramer's $V = .109$ (see Table 3.2).

Effect of Post-Event Information

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to the effect of post-event information, $\chi^2(2, N = 91) = 3.51$, $p = .139$, Cramer's $V = .196$ (see Table 3.2).

Minor Details

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to minor details, $\chi^2(2, N = 93) = 10.82$, $p = .004$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .341$. A greater proportion of the laypersons held an erroneous belief regarding this item compared to the RIs

(41.0% vs. 12.5%). The RIs were more likely than the laypersons to express uncertainty (46.9% vs. 19.7%) (see Table 3.2).

Confidence-Accuracy

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to confidence-accuracy, $\chi^2(2, N = 93) = 5.19, p = .075$, Cramer's $V = .236$ (see Table 3.2).

Impact of Stress

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to the impact of stress, $\chi^2(2, N = 93) = 1.43, p = .490$, Cramer's $V = .124$ (see Table 3.2).

Attitudes and Expectations

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to attitudes and expectations, $\chi^2(2, N = 93) = 0.55, p = .809$, Cramer's $V = .077$ (see Table 3.2).

Weapon Focus

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to weapon focus, $\chi^2(2, N = 93) = 31.95, p < .001$. The association was of large strength (Cohen, 1988), Cramer's $V = .586$. A greater proportion of the laypersons held a belief in line with the current research regarding this item compared to the RIs (70.5% vs. 18.8%). The RIs were more likely than the laypersons to express uncertainty (78.1% vs. 18.0%) (see Table 3.2).

Forgetting Curve

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to the forgetting curve, $\chi^2(2, N = 92) = 1.08, p = .583$, Cramer's $V = .108$ (see Table 3.2).

Children's Recall

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to children's recall, $\chi^2(3, N = 93) = 9.04, p = .029$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .312$. A greater proportion of the laypersons held a belief in line with the current research regarding this item compared to the RIs (16.4 % vs. 0.00%). The RIs were more likely than the laypersons to hold the belief that children remember as well as adults (71.9% vs. 44.3%) (see Table 3.2).

Dramatic Events

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to dramatic events, $\chi^2(3, N = 93) = 22.13, p < .001$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .488$. A greater proportion of the laypersons held a belief in line with the current research regarding this item compared to the RIs (67.2% vs. 31.3%). The RIs were more likely than the laypersons to express uncertainty (50.0% vs. 8.2%) (see Table 3.2).

Immediate Acceptance of Suggested Information

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to immediate acceptance of suggested information, $\chi^2(2, N = 93) = 3.98, p = .160$, Cramer's $V = .207$ (see Table 3.2).

Response Bias

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to response bias, $\chi^2(2, N = 92) = 0.77, p = .846$, Cramer's $V = .091$ (see Table 3.2).

False Memories

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to false memories, $\chi^2(2, N = 93) = 2.43, p = .274$, Cramer's $V = .162$ (see Table 3.2).

Anatomical Dolls

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to anatomical dolls, $\chi^2(2, N = 93) = 18.89, p < .001$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .451$. A greater proportion of the RIs held a belief in line with the current research regarding this item compared to the laypersons (62.5% vs. 18.90%). The laypersons were more likely than the RIs to express uncertainty (67.2% vs. 28.1%) (see Table 3.2).

Communication Aids

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to communication aids, $\chi^2(2, N = 93) = 5.41, p = .067$, Cramer's $V = .241$ (see Table 3.2).

Question Types

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to question types, $\chi^2(2, N = 92) = 9.75, p = .006$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .325$. A greater proportion of the laypersons held a belief in line with the current research regarding this item compared to the RIs (88.5 % vs. 67.7%). The RIs were more likely than the laypersons to hold an erroneous belief (19.4% vs. 1.6%) (see Table 3.2).

Open-ended Questions

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to open-ended questions, $\chi^2(2, N = 93) = 2.84, p = .242$, Cramer's $V = .175$ (see Table 3.2).

Recall vs. Recognition Memory

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to recall vs. recognition memory, $\chi^2(2, N = 93) = 7.57, p = .017$. The association was of small strength (Cohen, 1988), Cramer's $V = .285$. A greater proportion of the laypersons held a belief in line with the current research regarding this item compared to the RIs (80.3% vs. 71.0%). The RIs were more likely than the laypersons to express uncertainty (43.8% vs. 18.0%) (see Table 3.2).

Repeating Questions

There was a statistically significant association between participant type (RI vs. Layperson) and belief for the item related to repeating questions, $\chi^2(2, N = 93) = 13.38, p = .001$. The association was of moderate strength (Cohen, 1988), Cramer's $V = .379$. A greater proportion of the RIs held a belief in line with the current research regarding this item compared to the laypersons (53.1% vs. 21.3%). The laypersons were more likely than the RIs to hold an erroneous belief (50.8% vs. 15.6%) (see Table 3.2).

Children's Narratives

There was a non-significant association between participant type (RI vs. Layperson) and belief for the item related to children's narratives, $\chi^2(2, N = 92) = 3.90, p = .142$, Cramer's $V = .206$ (see Table 3.2).

Table 3.2*Registered Intermediary and Layperson Beliefs Regarding the 20 Item Questionnaire*

Questionnaire Item	Belief	Number / Percentage Held Belief	
		RI	Layperson
Multiple Witnesses	Agree*	28 (87.5%)	54 (88.5%)
	Disagree	1 (3.1%)	4 (6.6%)
	Uncertain	3 (9.4%)	3 (4.9%)
Effect of Post-Event Information	Agree*	23 (71.9%)	51 (86.4%)
	Disagree	1 (3.1%)	2 (3.4%)
	Uncertain	8 (25.0%)	6 (10.2%)
Minor Details	Agree	4 (12.5%)	25 (41.0%)
	Disagree*	13 (40.6%)	24 (39.3%)
	Uncertain	15 (46.9%)	12 (19.7%)
Confidence-Accuracy	Agree	3 (9.4%)	13 (21.3%)
	Disagree*	20 (62.5%)	41 (67.2%)
	Uncertain	9 (28.1%)	7 (11.5%)
Impact of Stress ^a	Agree	17 (53.1%)	40 (65.6%)
	Disagree*	7 (21.9%)	9 (14.8%)
	Uncertain	8 (25.0%)	12 (19.7%)
Attitudes and expectations	Agree*	26 (81.3%)	53 (86.9%)
	Disagree	2 (6.3%)	3 (4.9%)
	Uncertain	4 (12.5%)	5 (8.2%)
Weapon Focus	Generally True*	6 (18.8%)	43 (70.5%)
	Generally False	1 (3.1%)	7 (11.5%)
	Uncertain	25 (78.1%)	11 (18.0%)
Forgetting Curve	Generally True*	12 (38.7%)	27 (44.3%)
	Generally False	6 (19.4%)	7 (11.5%)
	Uncertain	13 (41.9%)	27 (44.3%)
Children's Recall	Better	4 (12.5%)	12 (19.7%)
	As Well As	23 (71.9%)	27 (44.3%)
	Worse*	0 (0.0%)	10 (16.4%)
	Uncertain	5 (15.6%)	12 (19.7%)
Dramatic Events	Better*	10 (31.3%)	41 (67.2%)
	As Good As	4 (12.5%)	7 (11.5%)
	Worse	2 (6.3%)	8 (13.1%)
	Uncertain	16 (50.0%)	5 (8.2%)
Immediate Acceptance of Suggested Information	Agree*	27 (84.4%)	46 (75.4%)
	Disagree	0 (0.0%)	7 (11.5%)
	Uncertain	5 (15.6%)	8 (13.1%)
Response Bias	Agree*	27 (87.1%)	56 (91.8%)
	Disagree	1 (3.2%)	2 (3.3%)
	Uncertain	3 (9.7%)	3 (4.9%)

Questionnaire Item	Belief	Number / Percentage Held Belief	
		RI	Layperson
False Memories	Agree*	12 (37.5%)	32 (52.5%)
	Disagree	2 (6.3%)	5 (8.2%)
	Uncertain	18 (56.3%)	24 (39.3%)
Anatomical Dolls	Agree*	20 (62.5%)	11 (18.0%)
	Disagree	3 (9.4%)	9 (14.8%)
	Uncertain	9 (28.1%)	41 (67.2%)
Communication Aids	Agree*	12 (37.5%)	32 (52.5%)
	Disagree	9 (28.1%)	6 (9.8%)
	Uncertain	11 (34.4%)	23 (37.7%)
Question Types	Agree*	21 (67.7%)	54 (88.5%)
	Disagree	6 (19.4%)	1 (1.6%)
	Uncertain	4 (12.9%)	6 (9.8%)
Open-Ended Questions	Agree	21 (65.6%)	48 (78.7%)
	Disagree*	11 (34.4%)	12 (19.7%)
	Uncertain	0 (0.0%)	1 (1.6%)
Recall vs. Recognition Memory	Agree*	17 (53.1%)	49 (80.3%)
	Disagree	1 (3.1%)	1 (1.6%)
	Uncertain	14 (43.8%)	11 (18.0%)
Repeating Questions	Agree	5 (15.6%)	31 (50.8%)
	Disagree*	17 (53.1%)	13 (21.3%)
	Uncertain	10 (31.3%)	17 (27.9%)
Children's Narratives	Agree*	17 (53.1%)	44 (73.3%)
	Disagree	4 (12.5%)	5 (8.3%)
	Uncertain	11 (34.4%)	11 (18.3%)

Note. The acronym RI refers to Registered Intermediary. Not all participants responded to every question.

*Correct response based on empirical evidence / previous research

^a Due to the controversial nature of this item, the analyses related to knowledge scores, erroneous belief scores, and expressions of uncertainty were conducted with and without this item. The items inclusion was found to have no appreciable effect on the findings.

3.3 Discussion

This exploratory study is the first to compare RIs' and lay peoples' beliefs regarding children's memory and investigative interview practices. The hypotheses were that RIs would score higher on the total knowledge score and lower on the erroneous belief score than lay people. These hypotheses were not supported. No difference was found in the erroneous belief scores of the RIs and lay people. However, a difference was found in the knowledge scores, with the lay people scoring significantly higher than the RIs. Explanations for this counterintuitive finding will be discussed below, along with areas of memory and investigative interview practice in which the RIs appear to both possess and lack knowledge.

All of the RIs in the current study reported having experience supporting child witnesses in investigative interviews. One may presume that this experience would result in a greater knowledge of children's memory and investigative interview practice. However, in the current study the RIs achieved lower knowledge scores than the lay persons and thus appear to possess less beliefs that are in line with current research findings. The difference in knowledge scores appears to be a by-product of the RIs' uncertainty, with the RIs found to opt for the 'uncertain' response alternative significantly more than the laypersons. When the results of the 20 items were analysed individually, there were four items whereby the RIs were more likely than the lay people to state that they were 'uncertain'. For three of these items (e.g., weapon focus, dramatic events, recall vs. recognition) the laypersons were more likely to provide a correct response – a pattern of responding that could account for the significant difference in knowledge score. Criminal justice practitioners airing on the side of caution has been observed in previous research (Kask, 2011). Kask (2011) attributed this to criminal justice practitioners (e.g., judges, prosecutors) having to make weighty and highly consequential professional judgements. Part of an RIs role is to make recommendations to these other criminal justice practitioners and as such their involvement could potentially have

a profound influence on the progression of an investigative interview. With that in mind it is perhaps not surprising that the RIs, in the current study, responded with similar caution. This caution may also stem from the fact that there have been concerns regarding the training RIs receive and their level of expertise (Powell et al., 2015). If the RIs in the current study were aware of such concerns they may have been more likely, if not confident in their response, to state that they were 'uncertain'. This may have been an attempt to safeguard against any criticisms towards their profession.

Irrespective of the motives behind this pattern of responding, the RIs' mean knowledge score (10.19) is a cause for concern. It indicates that there are gaps in the RIs' knowledge and that further training on the topic of children's memory and investigative interview practice is required. That being said, there was a considerable amount of variability in the knowledge scores of the RIs, with scores ranging from 4 to 15. This finding could potentially explain why in Henry, Crane, et al.'s (2017) study there was a significant difference in the amount of correct information the children provided as a product of the RI. One would suspect that RIs who possess a better knowledge of memory and best practice guidance, would be more successful in scaffolding children's communication in an investigative interview context. The variability in knowledge scores could also relate to when the RIs completed their initial training. The training has recently been revised and covers question types in more detail than the previous 5-day course new RIs received (e.g., for an overview of the revised training see Collins & Krähenbühl, 2020). The experience of the RIs in the current study ranged from 1 to 14 years so although some of the RIs will have undergone the revised training, with the increased coverage of question types, many will not have and thus may have performed more poorly on these items.

Overall, the findings of the current study are in line with the previous research. Previous research has found that criminal justice practitioners generally have a limited

knowledge and hold erroneous beliefs regarding eyewitness testimony (e.g., Jiang & Luo, 2016, Wise & Safer, 2010), with considerable variability in beliefs even amongst professionals within the same occupational group (e.g., Granhag et al., 2005). As with the professionals in Erens et al.'s (2020) study, the RIs, in the current study, held some beliefs that are in keeping with and others that are in conflict with current scientific evidence. For the 20 items, included in the current study, correct responses from the RIs ranged from 0.0% to 87.5%. Many of the RIs held accurate beliefs in relation to the recall of multiple witnesses (87.5%), the influence of attitudes and expectations on event memory (81.3%), the immediate acceptance of suggested information (84.4%), and response biases in young children (87.1%). In contrast, many of the RIs held erroneous beliefs in relation to children's recall (84.4%) and what qualifies as an open-ended question (65.6%).

In terms of children's recall 12.5% of the RIs' incorrectly stated that young children remember events they have experienced 'better' than adults, 71.9% stated that young children remember the events 'as well as' adults, and 15.6% RIs stated that they were 'uncertain' regarding this item. None of the RIs, in the current study, held the belief that young children remember experienced events 'worse' than adults. Although previous research has found that children's free-recall accounts tend to be less detailed than those of adults (e.g., Allwood et al., 2008; Jack et al., 2014), their accounts are often just as accurate (i.e., Jack et al., 2014; Knutsson & Allwood, 2014). In the current study, the item pertaining to children's recall did not specify whether it was referring to accuracy (i.e., quality) or detail (i.e., quantity), which could potentially account for the RIs appearing to perform poorly on this item. In addition, the item did not state whether the children were recalling the event with or without support. Given the RIs respective role, the RIs may have been responding to this item as if the child was recalling the event, with support, under optimal recall conditions. According to Vygotsky's zone of proximal development (1978), with appropriate scaffolding and support,

children can achieve tasks and levels of competence (i.e., providing a detailed narrative account) which would have otherwise been outside of their reach. An interesting avenue for future research would be to examine whether with RI support children are able to provide accounts which are more comparable to adults in respect to both accuracy and detail.

The second item with which many RIs held an erroneous belief related to defining an open-ended question. Of the RIs, 65.6% stated that ‘Can you tell me what happened?’ is an example of an open-ended question. As discussed in section 2.2.2 of this thesis this is not an open-ended question but an indirect speech act. Indirect speech acts ask if the child knows, whilst indirectly asking what they know (Evans et al., 2014). Although the difference between open-ended questions (e.g., ‘Tell me what happened’) and indirect speech acts (e.g., ‘Can you tell me what happened?’) can be very subtle, RIs with their expertise in communication should be aware of such nuances. These nuances can have a profound impact upon the response elicited from a child. Whilst an open-ended question may elicit a long, descriptive response; a similar indirect speech act may elicit a simple ‘yes’, with the child having failed to recognise the indirect question. This finding suggests that RIs may require more training in respect to different question types.

There was another finding, in the current study, that would support the assertion that RIs require additional training on question types. A significant association was found between participant type (RI vs. Layperson) and belief for the item pertaining to question types (i.e., ‘According to current best practice guidance (i.e., ABE), open-ended questions should be used predominantly during an investigative interview. These questions generally elicit the most accurate accounts’). The RIs were more likely than the laypersons to hold an erroneous belief (19.4% vs. 1.6%). Given that the role of RIs is to make recommendations regarding communication (including question types) in investigative interviews one would expect that the RIs would have a good understanding of the current best practice guidance. That being

said, previous research has shown that RIs have a lack of knowledge in this area (Krähenbühl, 2011). One of the findings of Krähenbühl's (2011) study was that many of the alternatives provided by RIs are not open-ended. This, along with the findings of the current study, may suggest that RIs fail to recognise the utility and potential for open-ended questions (Wright & Powell, 2006).

Potentially such a belief may have arisen due to the RIs having a lot of experience working with very young children or children with more profound communication needs. Research has found that young children (3- to 4-year-olds) may respond more informatively to specific-closed, as opposed to open-ended questions, due to specific-closed questions requiring less retrieval effort (Hershkowitz et al., 2012). Furthermore, ABE (MoJ, 2011) acknowledges that for some vulnerable witnesses the use of forced-choice questions may be the only viable option. Hence, there is evidence to suggest that asking predominantly open-ended questions is not always the most effective approach. However, this does not justify holding the erroneous belief which was expressed by almost 20% of the RIs in the current study. The current study asked about something that is explicitly stated in ABE (MoJ, 2011), which arguably leaves little room for interpretation. Open-ended questions are considered in ABE (MoJ, 2011) to be the gold-standard as these questions elicit the most accurate information. As such, the guidance advises interviewers to begin the interview with an open-ended invitation and to use further open-ended questions wherever possible to elicit additional information (MoJ, 2011).

Limitations and Areas for Future Research

One limitation of the current study is the modest sample size. That being said, the number of RIs that participated in the current study is comparable to previous research (e.g., Mattison & Dando, 2020). This is despite the current study having more stringent inclusion criteria (i.e., the RIs had to be registered to accept appointments for witnesses under the age of

18) which would reduce the limited pool of potential participants even further. A second limitation of the current study is that the item pertaining to children's recall did not specify whether it was referring to accuracy or detail (it was important that the questionnaire was accessible to lay persons). Future research should incorporate items relating to both aspects of recall, whilst also specifying the conditions in which recall is taking place (e.g., with support). A further limitation is that there was no opportunity, in the current study, for the RIs to elaborate upon their answers. It would have been very interesting, particularly in respect to children's recall and question types, to find out what fostered the RIs' beliefs. It is recommended that future research explore this further, using a more qualitative approach, that can delve deeper and potentially challenge these beliefs.

Conclusions

The current study is the first to examine RIs beliefs regarding children's memory and investigative interview practices. The findings were comparable to previous research (e.g., (e.g., Jiang & Luo, 2016, Wise & Safer, 2010), which has shown that criminal justice practitioners have a limited knowledge of these topics. Perhaps of greatest concern was that many of the RIs, in the current study, held erroneous beliefs regarding question types – an area that one would presume is a prerequisite to performing the RI role. It is therefore recommended that RIs receive additional training on children's memory and current best practice interview guidance (e.g., ABE, MoJ, 2011). This is not the first study to suggest that RIs require additional training (e.g., Collins & Krähenbühl, 2020). However, the findings of Collins and Krähenbühl (2020) would suggest that generally RIs are very open to the prospect of additional training and improving their practice. As such, it is recommended that additional training is implemented as soon as possible due to the profound impact this training could have upon the practice of RIs and the quality of child investigative interviews.

Chapter Four: The Ability of Pre-Interview Assessments to Inform Interviewers (and Registered Intermediaries) about Children's Communication (Study 2)

Pre-interview communication assessments can serve a dual purpose. The first is to help a child understand what is expected of them during the interview process and the second is to gain an understanding of the child's communication abilities. Having an awareness of a child's communication abilities can help a police officer (potentially in collaboration with an RI) plan an interview that is more in line with the child's needs. Although pre-interview assessments are ubiquitous to the RI role, forming the basis of the RIs' recommendations, assessments may also need to be conducted by police officers. The ABE guidance (MoJ, 2011) states that a pre-interview assessment "should be considered for all child witnesses" (p.27). Yet, the assistance of an RI has until recently been relatively rare (Criminal Justice Joint Inspection, 2014). Although the number of referrals for RI assistance has almost doubled in the last six years (MoJ, 2020b), some cases (0.6%) still go unmatched. It is also likely, due to time pressures, that for some cases RI referrals are still not being made at all (see section 2.2.6 of this thesis). Thus, placing the responsibility on the investigative interviewer to conduct the assessment.

ABE (MoJ, 2011) provides examples of factors that investigative interviewers may want to explore during a pre-interview assessment (see section 2.2.6 of this thesis). However, there is no mention in the guidance as to *how* to actually assess these factors. This has left police officers feeling unsure about what a formal pre-interview assessment should involve (McCullough, 2017) and may be responsible for their infrequent use. In a Criminal Justice Joint Inspection, conducted in 2014, only 13 of 69 interviews reviewed showed evidence of a pre-interview assessment having taken place. Of these only 10 cases had recorded the specific needs of the child (i.e., social, cognitive, linguistic, physical, and sexual). If a lack of clarity regarding what an assessment should involve is the main barrier to implementation, this could

be overcome through introducing and training officers in the use of specific assessment tools including ABELS (Jackson, 2016, as cited in McCullough, 2017) and ‘Unpacking the Box’ (Triangle, 2015). Despite these tools currently being used in practice, both lack empirical support. However, the ABELS website (www.abels.org.uk/services/child-abels/) indicates that further evaluations of the tool are imminent. Hence, the decision was made to focus, in this thesis, on the efficacy of the ‘Unpacking the Box’ tool (Triangle, 2015).

‘Unpacking the Box’ (Triangle, 2015) provides a structured method of assessing children’s receptive communication; expressive communication; attention, anxiety, and behaviour (see section 2.2.6). Although ‘Unpacking the Box’ (Triangle, 2015) has been used in applied settings for a number of years (mainly by RIs and NRIs), to date, only one piece of research has examined the tool’s efficacy and this was only in respect to one aspect of communication (Iranzo, 2016). Iranzo (2016) found ‘Unpacking the Box’ to be a reliable and engaging method for measuring working memory in young children. However, the lack of research regarding the other facets of communication the tool purports to assess is of concern. There is a growing emphasis, within the legal arena, upon evidence-based practice, with practices lacking a strong theoretical framework coming under increasing scrutiny. One example, mentioned in chapter two, is the ‘20 principles’ underpinning vulnerable witness advocacy training. There are calls for the training to be overhauled and the ‘20 principles’ reformed due to insufficient empirical support (Cooper et al., 2018). Research is therefore essential to determine whether ‘Unpacking the Box’ (Triangle, 2015) is a reliable tool.

Aims

It is important to have an insight into a child’s abilities, both social and cognitive, in order to plan and conduct an effective investigative interview (Smith & Milne, 2017). How successful the interviewer is in doing this can ultimately determine the quality of a child’s evidence. The present small-scale exploratory study examined whether a communication

assessment, using 'Unpacking the Box' (Triangle, 2015), can provide a reliable indication of a child's abilities. This was achieved by children, aged 4 to 9 years old, participating in a to-be-remembered event. The children were allocated to one of three experimental conditions: assessment, no assessment, and colouring activity. One week later, the children were interviewed about their experiences. Prior to the interviews, predictions were made regarding the children's interview performance. Predictions were made regarding the children's use of ground rules, attention span, responsiveness, resistance to suggestion, and drawing ability. For the children in the assessment condition predictions were based on the findings of the 'Unpacking the Box' (Triangle, 2015) assessment. As for the children in the no assessment and colouring conditions the predictions were based solely on professional judgement. The analysis examined whether the predictions based on the findings of a communication assessment were more accurate than those based entirely on professional judgement.

When there is no pre-interview communication assessment, interviewers are relying solely upon their professional judgement. Professional judgement relies on the interviewer having knowledge of age-based norms / developmental milestones. However, not all children develop at the same rate (e.g., language competence; Adams et al., 1999). Thus, the milestones merely act as a guide. A pre-interview communication assessment, on the other hand, assesses an *individual* child's abilities irrespective of age-based norms. Theoretically, this should give a more accurate representation of an individual child's communication abilities. Hence, it was hypothesised that predictions based upon the findings of a communication assessment would be more accurate than those based solely upon professional judgement – the researcher's judgement as a trained investigative interviewer and intermediary (see Appendix B for reflection on how this role has impacted upon the PhD process). More specifically, it was hypothesised that in the assessment condition, compared to

the colouring and no assessment conditions, the one interviewer (i.e., the researcher) would be more accurate in predicting the children's:

- use of ground rules,
- attention,
- responsiveness (conceptualised as the amount of information provided to open-ended questions),
- resistance to suggestion / compliance,
- and ability to draw and use that drawing to correctly identify body parts.

4.1 Method

4.1.1 Design

The study was an experimental independent measures design. The independent variable was the type of assessment, with three levels: pre-interview assessment, colouring activity, and no assessment. The children were allocated to one of the three experimental conditions. The dependent variables were the accuracy of predictions relating to the children's use of the ground rules, attention span, responsiveness, resistance to suggestion, drawing ability, and ability to use the drawing to identify body parts. The study consisted of five phases: cognitive testing, staged event (Mrs Science), assessment / colouring activity, predictions, and interview.

4.1.2 Participants

Sixty-five children from a primary school in the North East of England participated in the study. Following the cognitive testing, 12 children were excluded from the study due to insufficient attentional / cognitive abilities to complete the tests ($n = 9$), not meeting the inclusion criteria ($n = 1$), being absent from school ($n = 1$), and requesting to withdraw from the study ($n = 1$). A further two children were excluded from the study as their interviews contained fewer questions (the preliminary analysis identified these interviews as outliers),

leaving a total sample of 51 children. This comprised of 24 males and 27 females with a mean age of 83.14 months (6.93 years) and age range of 49 to 118 months (4.08 to 9.83 years).

Forty-six children were White, four were Asian, and one child was Mixed Race. All of the children spoke English as their first language, none of the children had been identified as having any additional needs, and none were currently involved or had previously been involved in an investigative interview or criminal proceedings.

The children were allocated to the three assessment conditions. Sixteen children (males = 8, females = 8, mean age = 84.19 months, age range = 59 to 118 months) were allocated to the pre-interview assessment condition, 15 children (males = 7, females = 8, mean age = 79.80 months, age range = 49 to 116 months) were allocated to the colouring activity condition, and 20 children (males = 9, females = 11, mean age = 84.80, age range = 57 to 116 months) were allocated to the no assessment condition. There were no significant differences between the three experimental conditions in relation to demographic (e.g., age) and cognitive factors (e.g., receptive language abilities; see section 4.2).

Due to the high attrition rate in the current study the decision was made to conduct a sensitivity power analysis using G* Power software. With an alpha level of $p < .05$ and power of .8 an effect size of .43 was calculated. Thus, there was adequate power to detect a large effect size (.5, Cohen, 1992). Due to the study only having enough power to detect a large effect the findings of the study need to be interpreted with caution and considered exploratory.

4.1.3 Materials

Cognitive Tests. To increase confidence in the findings (i.e., that any differences between the assessment conditions could not be attributed to differences in cognition) a number of cognitive tests were administered. Statistical analyses were conducted to ensure that there were no statistically significant differences between the three groups of children (i.e., assessment, colouring activity, and no assessment condition) for each cognitive

assessment. The cognitive tests were chosen to represent the skills that children utilise most frequently within an investigative interview context. Given the relatively wide age range included in this study, it was challenging to choose cognitive tests which were appropriate for use with the whole sample. As such, a considerable amount of time was spent discussing the potential options with an experienced developmental researcher. Although some of the cognitive tests fall outside the age ranges used in the current study, this is negated by the fact that the researcher was using the raw scores from these tests to control for differences between child groups. The researcher was not concerned with the standardised scores.

Receptive Language. The British Picture Vocabulary Scale Third Edition (BPVS3; Dunn et al., 2009) is a test of receptive language. BPVS3 is suitable for use with children between the ages of 3 and 16 years old. The test involves the child selecting the correct picture from the four options provided. The BPVS3 comprises of 14 sets. Once, the child provides eight incorrect responses in a single set the test is terminated and scored.

Expressive Language. The Renfrew Action Picture Test (RAPT; Renfrew, 1997) Measures children's expressive language in respect of information given and grammatical structure. RAPT is designed for use with children aged 3 to 8.5 years old. The test comprises of a series of 10 cards each with a picture and accompanying question (e.g., card 1 depicts a girl holding a teddy bear. The question 'what is the girl doing?'). The child is shown each card in turn and asked the question on the reverse of the card. The exact words spoken by the child are recorded and their response scored. Additional prompts may be required to encourage the child to provide more information. However, these are used sparingly.

Visuospatial ability. Ravens Coloured Progressive Matrices (CPM; Raven, 2008) measures visuospatial ability and is designed for use with children from 5 to 11 years old. The test requires the child to select the missing piece of a pattern from six possible options. The CPM consists of 36 items. The child receives 1 point for every correct response.

Inhibition. The Day / Night Task (Gerstadt et al., 1994) measures inhibition and is suitable for use with children between the ages of 3 and 6 years old. The task requires the child to first identify pictures of the sun / moon and what time of day they appear (day / night time). The child is then told about 'Wally the Whale' who mixes things up (i.e., Wally says 'night-time' when he sees the sun and vice versa). The child is instructed to do the same and is shown shuffled pictures of the sun / moon (14 trials). The child is scored 2 for answering correctly the first time and 1 for answering incorrectly then correcting themselves.

The Stroop Task (Stroop, 1935) measures inhibition and is suitable for use with children 6 years and above. Prior to the task the child's ability to identify four colours and the corresponding words is ascertained. During the task the child is presented with a list of 12 incongruent words. Incongruent words are written with a different ink colour than their meaning (e.g., the word "red" printed in blue ink). The children are then asked to name the ink colour. The children are scored 2 for answering correctly the first time and 1 for answering incorrectly then correcting themselves.

Attention. Standard Dimensional Change Card Sort (SDCCS; Carlson, 2005; Frye et al., 1995) measures executive functioning and is designed for use with children from 3 to 7 years old. This is achieved through the use of red / blue cards featuring cars and stars. Two boxes with slots on the lids are placed in front of the child, one with a blue car attached and one with a red star. First, the child plays the 'colour game'. This involves the child placing the cards in the corresponding box regardless of shape. If the child gets less than five of the trials correct the test is terminated and the child given a score of 0. If the child answers five of the six trials correctly, the child plays the 'shape game'. This involves the child placing the cards in the correct box regardless of colour. If the child gets less than five of the trials correct the test is terminated and the child given a score of 1. If the child answers five of the six trials correctly, the child sorts the cards depending on whether the cards have a border. If the card

has a border the child has to play the 'colour game' and if the card does not have a border the child plays the 'shape game' – there are 12 trials. If the child gets less than nine trials correct, they are given a score of 2, if they get more than nine correct, they are given a score of 3.

Staged Event. The staged event was adapted from Dickinson and Poole's (2017) Mr Science Germ Detective paradigm. The event was scripted. It was also video recorded to ensure that any minor deviations from the script could be accounted for when assessing the accuracy of the children's accounts. The event began by a research assistant describing to the child (each child took part in the staged event individually) the potential contaminating effects of touching. The research assistant also explained that, to avoid spreading germs, Mrs Science had been told not to touch children's skin and that the child should remind Mrs Science of this rule if she forgets. Mrs Science then introduced herself to the child and helped him / her put on a lab coat and safety glasses. Next the child took part in the first germ detective activity. The activity was about the importance of covering your face when you sneeze. It involved the child spraying water with a bottle. The child sprayed the water on two occasions (with and without Mrs Sciences' hand in front of the sprayer) and measured how far it travelled. Once the activity was complete Mrs Science attempted to brush water off the child's face. The second activity was about how germs can easily transfer from one object to another. The child was asked to put on some gloves and dip their fingers in some petroleum jelly. Mrs Science then sprinkled glitter on the gloves and asked the child to touch a cup. Mrs Science also touched the cup. She then asked the child to close their eyes whilst she touched something else. The child had to find the glitter (i.e., where she had touched) using a magnifying glass. The final activity involved showing the child how to wash their hands correctly. Mrs Science put some germ-glow on her hands she then asked the child to shine a black light on her hands so the child could see the pretend germs. Mrs Science then washed her hands once according to the child's instructions and once using a correct hand washing technique (during the latter

the child sang “happy birthday”). After each hand washing the child looked at Mrs Sciences’ hands under the black light. Finally, Mrs Science thanked the child and attempted to shake their hand. If the child did not alert Mrs Science of the rule break, Mrs Science reminded the child of the rule.

Colouring Materials. The children were presented with five different books / activity packs (e.g., Spiderman, Unicorns, Enchanting Nature) from which they could choose a picture to colour. The children were also provided with a pack of 32 colouring pencils.

‘Unpacking the Box’. ‘Unpacking the Box’ is an assessment tool developed by Triangle (2015). The tool can be used to quickly assess children’s receptive communication (i.e., ability to understand language); expressive communication (i.e., ability to use language); attention, anxiety, and behaviour. More specifically it can test a child’s understanding and use of sequencing vocabulary (e.g., before, after), prepositions (e.g., in, on, under), comparatives (e.g., same, different), auditory working memory capacity, and ToM. It consists of a silver box which contains small objects such as keys, thimble, paperclips, for the child to work with. It also comes with a guidance manual. The guidance manual was used as the basis for the assessment plan below (see figure 4.1). A number of additional materials were required for the assessment. These included pencils and paper for the children to draw, outlines of gingerbread people (these were to be used if a child was unable to draw a picture of themselves where body parts were recognisable), a 90 second timer (to limit the amount of time children spent drawing their pictures), and a smarties box containing paperclips (can be used to test ToM).

Figure 4.1*Assessment Plan Based Upon the 'Unpacking the Box' Framework*

Activity	Complete / Notes
<p>Introductions</p> <p>Hi (child's name). My name is Alex and I have got some jobs that I need your help with. The jobs will help me find out how you think and how you use your words. Is that ok?</p> <p>Thank you. If you want a break while we are doing the jobs that is ok. Please tell me or press this big red stop button here (shows child stop card).</p> <p>We have got three jobs to do. (Shows child job cards). Talking, drawing, and silver box. We are going to do the talking job first.</p>	
<p>Practice narrative (including introduction and practise of the ground rules, as well as the below question formats)</p>	
<p>Introduction of the rules:</p> <p>(Child's name) I have got something really important to show you (Shows child rule cards: 'I don't know', 'If I get it wrong, tell me', 'I don't understand'). These are talking rules. (Encourages child to read rules and quickly explains).</p>	
<p>Open-Ended Questions</p> <p>Examples:</p> <p>So now we are ready to start our talking jobs. I heard that you (activity child has taken part in).</p> <p>Tell me everything about ... (Open-ended invitation)</p> <p>Then what happened? (Open-ended breadth)</p> <p>What happened next?</p> <p>Tell me more about the part where... (Open-ended depth)</p> <p>You said you went ... Tell me about that</p>	
<p>Specific Closed Questions (ensure pairing - follow with open-ended questions)</p> <p>Examples:</p> <p>When?</p> <p>Who?</p> <p>Where?</p> <p>What?</p> <p>Why?</p>	

What made?

How?

How many?

How far?

How long?

How did you / others feel?

Yes/No Questions

Forced-Choice Questions

Two-option

Three-option

Tagged Questions

Negative tag (...is it?)

Positive tag (...isn't it?)

Statements

Rule Practise (deliberately violated ground rules to evaluate child's response)

I don't know (Asked the child a question they could not possibly know the answer to)

Tell me if I get it wrong (Summarise incorrectly)

I don't understand (What happened when you got there, went around the other way and there you were on the next bit before Tuesday?)

Drawing

I would like you to draw a picture of yourself. You have got to be fast. You only have until the bubbles get to the other side (of the timer). There are lots of nice colours for you to use (shows child box of crayons).

Now I would like you to draw a picture of your teacher.

Deliberate Naming Error (deliberately confused drawings of child and teacher)

Is this you? (Praise for correct response / recap rules if an error)

Body Parts (ask child to identify body parts on drawing / gingerbread figure)

Show me... (neutral parts e.g., not where the child was touched during the staged event / practise don't understand rule)

What is this...?

Silver Box

Please take everything out of the box for me.

Names

What's this called? (Thimble / Star / Paper Clip / Hair clip / Triangle / Square)

Counting

How many paperclips are there?

Are there more keys or paperclips?

Same / Different

Pick two things that are a different colour.

Pick two things that are the same size.

Pick two things that are a different shape.

ICW (words the child needs to understand in order to follow the instruction) and Concepts

Where is the key? (1 ICWs)

Point to the thimble. (2 ICWs)

Give me the big key. (3 ICWs)

Point to the black bag. (3 ICWs)

Put the thimble in the black bag. (4 ICWs)

Give me the big key and hair clip. (4 ICWs)

Put the triangle under the big box. (4 ICWs)

Put the thimble between the two boxes. (4 ICWs)

Put two paper clips on the square. (4 ICWs)

Put the thimble and hair clip in the silver bag. (5 ICWs)

Put the black bag next to the small box. (5 ICWs)

Give me the thimble after pointing to the hair clip. (5 ICWs)

Point to the black bag and give me the triangle and thimble. (6 ICWs)

Put the thimble on the square after giving me the hair clip. (6 ICWs)

Put the big key and small star in the black bag. (7 ICWs)

Give me the triangle, put the thimble on the square and point to the hair clip. (7 ICWs)

Theory of Mind

I have one more box (shows child box of crayons or Lego). What do you think is in here?

If I showed your teacher the box what do you think she would say was in it?

Conclusions

Thank you for all your help. I will come back tomorrow so that we can do some more talking.

Note. ICW refers to Information Carrying Words

Rule Cards. The decision was made to present the ground rules visually, as well as verbally, to ensure consistency with the ‘Unpacking the Box’ framework (Triangle, 2015). Four rule cards (approximately 2 x 2 inches) were introduced to the child at assessment and again at interview. Three cards had a picture of a cartoon child accompanied by a rule written in bold underneath. The rules written on the cards were: ‘If I get it wrong, tell me’ (with a child holding both hands up in a stop position), ‘I don’t understand’, and ‘I don’t know’ (both with a child looking confused). The fourth and final card was a red stop sign which the children were asked to press if they needed a break (see Appendix B for examples). These ground rules were chosen to reflect sections 3.12 and 3.14 of the ABE guidance (MoJ, 2011).

Truth and Lies Video Clip. The child’s understanding of truth and lies was assessed, at the outset of each interview (in line with current best practice guidance, ABE, MoJ, 2011), using a video clip developed by Triangle (www.triangle.org.uk/resource-categories/downloads). The video clip is approximately 15 seconds long. It shows a girl and a boy sitting in a room, each has one sweet. When the girl leaves the room, the boy eats her sweet. When the girl returns, she asks the boy ‘Did you eat my sweet?’ and he replies ‘no’. The child is then asked a number of questions designed to assess their understanding of truth and lies.

Investigative Interview Script. An interview script was developed so that children’s responses could be compared across conditions. Having a script also sought to safeguard against any interviewer bias that may have unconsciously arisen due to the researcher not being blind to the condition (having previously conducted a pre-interview communication assessment or a colouring activity with some of the children). Previous research (e.g., Ginot &

Verkampt, 2007) cited the use of a written protocol for this same reason. Creating a written protocol is thought to minimise unintentional variation in the experiment by standardising the process as much as possible (Price et al., 2015)⁸.

The interview script consisted of a maximum of 52 questions (see Appendix C). The questions asked about the staged event. The questions varied in terms of type (e.g., open-ended, specific-closed) and complexity (e.g., length of questions, vocabulary). In line with ABE guidance (MoJ, 2011) after assessing the children's understanding of truth / lies and introducing / recapping the ground rules, a free narrative was elicited using open-ended questions (e.g., 'Tell me what happened?', 'Then what happened?'). This was followed by a combination of open-ended, specific-closed (e.g., 'How many times did Mrs Science wash her hands?'), forced-choice questions (e.g., 'Did you play one game, more than one game, or don't you know?'), and multiple questions (e.g., 'Did you find the glitter? Where did you find it?'). The interview also included misleading questions (e.g., 'Mrs Science had a light that was red, did she not?') to assess children's use of ground rules and resistance to suggestion. Which questions each child was asked was determined and directed by their previous response/s at the interview. This was to prevent some children from potentially being led throughout the entirety of the interview. During the interview, the children were asked to draw a picture of themselves to show / clarify the location of any touches.

Recording Equipment. An Olympus Digital Voice Recorder WS-852 was used to record the cognitive tests. Two Canon Legria HF R506 Camcorders were used to record the staged event. This allowed the accuracy of the children's accounts to be verified. A Canon Legria HF R506 was also used to record the interviews. Video recording was essential in order to code both children's verbal and non-verbal behaviours.

⁸ The researchers' behaviour was not inter-rated. Some previous studies also do not appear to have inter-rated the interviewers' behaviour (e.g., Dando et al., 2009; Ginet & Verkampt, 2007).

4.1.4 Procedure

Ethical approval was granted from Teesside University ethics committee, and consent gathered from the school and parents. The study procedure involved five phases:

Phase 1 – Cognitive Measures. Prior to the staged event, all of the children were subject to a number of cognitive tests (e.g., language, attention). In order to accommodate school timetables and maintain the children's engagement the cognitive tests were split over multiple sessions. The children's scores on each test were used to assess whether there were any statistically significant differences in cognition across the three assessment conditions. The tests were conducted by three research assistants (i.e., not the researcher conducting the interviews to minimise the time the primary researcher spent with the children before the interviews). After the assistants were recruited, they all had to undergo a DBS check and be fully trained in the cognitive test procedures.

Phase 2 – Staged Event. The staged event was conducted by two research assistants (the primary researcher was not involved in this phase). Very careful consideration was given to the nature of the to-be-remembered event in this study. Previous research has been criticised for incorporating innocuous touches that are not emotionally salient or memorable, limiting the extent to which the findings can be applied to real-world investigations of CSA (Lyon, 2012). The event chosen, for study two, was adapted from the empirically tested Mr Science Germ Detective paradigm. The original paradigm was developed by Dickinson and Poole (2017) and is about germ transmission and contagion prevention. The paradigm uses children's disgust for germs and contaminated objects to create memorable and inappropriate touches. At the beginning of the event, a research assistant described to the child the potential contaminating effects of touching. The assistant explained that, to prevent spreading germs, Mrs Science has been instructed not to touch children's skin. The child was told to remind her of this rule if she forgets. Following this explanation, the child took part in three germ

detective activities. Mrs Science attempted to touch the child on two occasions – once on their cheek and once on their hand. The staged event, generally, lasted between 10 and 15 minutes. In order to accommodate all of the children the entire experiment took place over multiple weeks. However, the experiment was carefully planned to ensure that the delay between the three test phases (i.e., staged-event, assessment or colouring activity, and interview) was the same for each child.

Phase 3 - Communication Assessment. Six days after the staged event, 16 children had a communication assessment. All of the assessments followed the ‘Unpacking the Box’ framework (see figure 4.1 for assessment plan) and were conducted by the primary researcher. The communication assessments ranged in length from 14 minutes 34 seconds to 26 minutes 49 seconds, with a mean length of 18 minutes 23 seconds. Fifteen children took part in a colouring activity. This was a collaborative activity with the researcher. The child chose a picture and the researcher and child coloured in together. It was felt that avoiding all conversation during the colouring activity may actively damage rapport. Hence, unstructured conversation (i.e., the content was not planned and no attempt was made to assess the child’s communication abilities or to teach them rules / behaviours relevant to the investigative interview) did take place during the colouring activity. Topics included the picture the children were colouring in and what they had been doing at school. The colouring activity ranged in length from 12 minutes 26 seconds to 32 minutes, with a mean length of 17 minutes 49 seconds. The remaining 20 children did not take part in either the colouring activity or the communication assessment. The only contact these children had with the primary researcher, prior to the interviews, was when they or other children were being collected from their classrooms (it is estimated that this would not have exceeded five minutes).

Phase 4 – Predictions. Predictions were made regarding whether the child would use the ground rules, what their attention span would be, how responsive they would be, their

resistance to suggestion, ability to draw, and their ability to use that drawing to identify body parts. For the children that had a communication assessment, this informed those predictions. The predictions, for the other children, were based upon the researcher's professional judgement and the child developmental literature.

Phase 5 - Interview. A week after the staged event, all of the children took part in an interview about the germ detective activities. All interviews were conducted by the primary researcher (the researcher who conducted the assessments and colouring activities). The interviews followed the general structure of an ABE (i.e., rapport, free-narrative, questioning, and closure; MoJ, 2011). However, in order to fully test the hypotheses, the interviews purposefully included some examples of poor questioning. The mean number of questions asked in the interview was 40, with a range from 27 to 49 (although the interview was scripted whether a child was asked a particular question was dependent upon the child's previous response this was to prevent non-leading questions from potentially becoming leading). The interviews were relatively short. The mean length of the interviews was 12 minutes 7 seconds, with a range from 8 minutes 20 seconds to 18 minutes 30 seconds.

4.1.5 Predictions and Coding

The researcher made predictions based upon the communication assessment or age-related expectations (informed largely by the developmental literature, see Appendix E).

Predictions related to the following:

Ground Rules. Prior to the interview, the researcher allocated each child a score (i.e., 1 = unlikely, 2 = sometimes, 3 = likely) pertaining to their likelihood to use the ground rules. This score was informed by the developmental literature (for those children in the colouring or no assessment conditions) or based upon the frequency with which the children utilised the rules during the pre-interview assessment, specifically during the practice narrative. For example, if the child did not use the ground rules during the assessment, despite multiple

opportunities to do so, the child was allocated a score of 1. The number of times the child appropriately used the ground rules during the interview was then calculated. Use of ground rules was scored as follows:

- 1 – child does not employ ground rules during the interview.
- 2 – child employs ground rules on one occasion during the interview.
- 3 – child employs ground rules on two or more occasions during the interview.

If this score corresponded with the pre-interview score, the interviewer's prediction was classified as correct, if not it was classified as incorrect.

Attention. Prior to the interview, the researcher allocated each child a category (0-4minutes, 5-9 minutes, 10-14minutes, 15-19minutes, 20minutes +) pertaining to their likely attention span. The category a child was allocated was informed either by the developmental literature (for those children in the colouring or no assessment conditions) or based upon the researcher's observations of the child during the pre-interview assessment (i.e., the point at which the child appeared to lose focus was noted and the corresponding category chosen). At what point the child began to get distracted, and lost focus during the interview was recorded. If this score corresponded with the allotted pre-interview category, the interviewer's prediction was classified as correct, if not it was classified as incorrect.

Responsiveness. Prior to the interview, the researcher allocated each child a score (i.e., 1 = low, 2 = medium, 3 = high) pertaining to their likely responsiveness to open-ended questions. Responsiveness was scored, during the interviews as follows:

- 1 – child responds with a single word or phrase.
- 2 – child responds with a full sentence.
- 3 – child gives an extensive narrative (multiple sentences).

If this score corresponded with the pre-interview score, the interviewer's prediction was classified as correct, if not it was classified as incorrect.

Resistance to Suggestion. Prior to the interview, the researcher allocated each child a score (i.e., 1 = low, 2 = medium, 3 = high) pertaining to their likelihood to acquiesce to suggestion. The number of times the child acquiesced during the interview was recorded.

Resistance to suggestion was scored as follows:

- 1 – child does not acquiesce to suggestion.
- 2 – child acquiesces to a misleading suggestion once during the interview.
- 3 – child acquiesces to a misleading suggestion on two or more occasions during the interview.

If this score corresponded with the pre-interview score, the interviewer's prediction was classified as correct, if not it was classified as incorrect.

Drawing. Prior to the interview, the researcher allocated each child's drawing a score pertaining to quality (i.e., 1 = unusable [cannot distinguish body parts], 2 = usable [can distinguish body parts]), and the child's ability to use the picture / body diagram to identify body parts (1 = able; 2 = unable). The child's drawings, during the interview, were judged using the same criteria. If this score corresponded with the pre-interview score, the interviewer's prediction was classified as correct, if not it was classified as incorrect.

4.1.6 Inter-Rater Reliability

Two raters (i.e., the primary researcher and one other) independently coded 10% of the transcripts. The second rater was blind to condition. There was an inter-rater reliability (percent agreement) of .8 for drawing ability and an inter-rater reliability of 1.0 for all other measures (i.e., ground rules, attention, responsiveness, resistance to suggestion). Any differences between raters were resolved by discussion.

4.2 Results

Preliminary analyses were conducted to determine whether there were any significant differences between the three assessment conditions (i.e., assessment, colouring, and no assessment) in terms of demographic, cognitive, and procedural variables.

Demographic variables

Age. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of age. There were no outliers as assessed by boxplots, and the data were normally distributed for each group as assessed by the Shapiro-Wilk test ($p > .05$). There was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .830$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.32, p = .732$.

Gender. A chi-square test of independence was conducted between gender and the assessment condition. All expected cell frequencies were greater than five. There was no statistically significant association: $\chi^2(2, N = 51) = 0.09, p = .956$.

School Year. The analysis showed that 13 cells had an expected count less than 5, so an exact significance test was selected for Pearson's chi square. There was no statistically significant association between school year and the assessment condition: $\chi^2(8, N = 51) = 12.63, p = .125$.

Ethnicity. The analysis showed that 6 cells had an expected count less than 5, so an exact significance test was selected for Pearson's chi square. There was no statistically significant association between ethnicity and the assessment condition: $\chi^2(4, N = 51) = 3.16, p = .619$.

Cognitive variables

BPVS III. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of BPVS III score. There were no outliers as assessed by boxplots,

and the data were normally distributed for each group as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .292$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.26, p = .769$.

RAPT. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of RAPT scores. Boxplots identified one outlier in the colouring condition. Removing the outlier created additional outliers and did not have an appreciable effect on the analysis. Therefore, the decision was made to include the outlier in the analysis. The data were normally distributed for each group as assessed by Shapiro-Wilk test ($p > .05$), and there was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .246$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.56, p = .574$.

Ravens. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of Ravens score. There were no outliers as assessed by boxplots, and the data were normally distributed for each group as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .688$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.02, p = .981$.

SDCCS. A Kruskal-Wallis test was conducted to determine if there was a difference in SDCCS between the three assessment conditions. There was no statistically significant difference between the groups: $\chi^2(2, N = 51) = 0.75, p = .689$.

Stroop Test. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of scores on the Stroop test. Boxplots identified two outliers and there were deviations from normality in all three conditions. However, the decision was made to include the outliers in the analysis as their removal did not have an appreciable effect on

the analysis, nor did running an equivalent non-parametric test. There was homogeneity of variances as assessed by Levene's test of homogeneity of variances ($p = .185$). The difference between the three assessment conditions was not statistically significant: $F(2, 34) = 1.23, p = .304$.

Day and Night Task. A one-way Welch ANOVA was conducted to examine whether the assessment conditions varied in terms of scores on the Day / Night task. There were no outliers and the data were normally distributed for each group. Homogeneity of variances was violated as assessed by Levene's test of homogeneity of variances ($p = .003$). The difference between the three assessment conditions was not statistically significant: Welch's $F(2, 5.03) = 2.77, p = .155$.

Assessment Variables

Duration. An independent-samples t-test was run to determine if there were differences in duration between the assessment and colouring activity. Boxplots identified two outliers and there were deviations from normality in both conditions. However, the decision was made to include the outliers in the analysis as their removal did not have an appreciable effect on the analysis, nor did running an equivalent non-parametric test. There was homogeneity of variances as assessed by Levene's test of homogeneity of variances ($p = .339$). The difference between the two assessment conditions was not statistically significant: $t(29) = 0.40, p = .692$.

Interview Variables

Number of Questions. A one-way ANOVA was conducted to determine if the assessment conditions varied in terms of the number of questions asked during the interviews. Boxplots identified two outliers in the colouring condition. The outliers were removed from further analysis. Once removed, the data were normally distributed for each group as assessed by Shapiro-Wilk test ($p > .05$), and there was homogeneity of variances as assessed by

Levene's test of homogeneity of variances ($p = .134$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 2.45, p = .097$.

Duration. A second ANOVA was conducted to examine whether the assessment conditions varied in terms of the length of the interviews. Boxplots identified two outliers and there were deviations from normality in the colouring condition. However, the decision was made to include the outliers in the analysis as their removal did not have an appreciable effect on the analysis, nor did running an equivalent non-parametric test. There was homogeneity of variances as assessed by Levene's test of homogeneity of variances ($p = .070$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.92, p = .407$.

Misleading Questions. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of the number of misleading questions. Boxplots identified three outliers and there were deviations from normality in all three conditions as assessed by the Shapiro-Wilk test. Neither transforming the data nor conducting an equivalent non-parametric test had an appreciable effect on the analysis. Therefore, the decision was made to report the results with the original data. There was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .289$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 1.99, p = .148$.

Open-ended Questions. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of the number of open-ended questions. There were no outliers as assessed by boxplots. However, there were deviations from normality in the assessment and no assessment conditions as assessed by the Shapiro-Wilk test. Neither transforming the data nor conducting an equivalent non-parametric test had an appreciable effect on the analysis. Therefore, the decision was made to report the results with the original data. There was homogeneity of variance as assessed by Levene's test of homogeneity of

variances ($p = .323$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.99, p = .379$.

Specific-Closed Questions. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of the number of specific-closed questions. Boxplots identified one outlier and the colouring condition was not normally distributed as assessed by the Shapiro-Wilk test. Transforming the data had no appreciable effect on the analysis. However, an equivalent non-parametric test gave a significant result ($p = .029$). In terms of the ANOVA there was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .450$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 2.63, p = .082$.

Forced-Choice Questions. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of the number of forced-choice questions. Boxplots identified one outlier and the no assessment condition was not normally distributed as assessed by the Shapiro-Wilk test. Neither transforming the data nor conducting an equivalent non-parametric test had an appreciable effect on the analysis. Therefore, the decision was made to report the results with the original data. There was homogeneity of variance as assessed by Levene's test of homogeneity of variances ($p = .449$). The difference between the three assessment conditions was not statistically significant: $F(2, 48) = 0.31, p = .738$.

Multiple Questions. A one-way ANOVA was conducted to examine whether the assessment conditions varied in terms of the number of multiple questions. Boxplots identified multiple outliers and deviations from normality in two conditions as assessed by the Shapiro-Wilk test. Transforming the data did not have an appreciable effect on the analysis, nor did running an equivalent non-parametric test. There was homogeneity of variances as assessed by Levene's test of homogeneity of variances ($p = .344$). The difference between the three assessment conditions was statistically significant⁹: $F(2, 48) = 6.57, p = .003$. The

number of multiple questions increased from the colouring condition ($M = 1.07$, $SD = .80$) to the assessment condition ($M = 1.38$, $SD = .96$) to the no assessment condition ($M = 2.10$, $SD = .85$). Tukey post hoc analysis revealed that the mean increase from the colouring to the no assessment condition (1.03, 95% CI [0.31, 1.75], $p = .003$) was statistically significant and the mean increase from the assessment to the no assessment condition (0.73, 95% CI [0.02, 1.43], $p = .043$) was statistically significant. No other group differences were statistically significant.

Predictions

Chi Square tests were conducted to examine whether communication assessments provide a more accurate representation of children's abilities (i.e., ability to use ground rules, attentiveness, responsiveness, resistance to suggestion, ability to draw, and use that drawing to identify body parts), operationalised as correct or incorrect, than professional judgement alone.

Ground Rules. There was a statistically significant association between the accuracy of predictions relating to the children's use of ground rules and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 51) = 9.58$, $p = .008$ (see table 4.1). The association was of moderate strength (Cohen, 1988), Cramer's $V = .433$. The assessment condition had the highest percentage correct (87.5%) and the colouring condition the lowest (33.3%).

Attention. The analysis showed that 3 cells had an expected count less than 5, so an exact significance test was chosen for Pearson's chi square. There was no significant association between the number of correct predictions pertaining to the children's attention

⁹It is difficult to determine as to what caused this significant difference. The questions children were asked were dependent upon their previous response. For example, the child would not have been asked the question 'Did you find the glitter? Where did you find it?' if they had not previously mentioned glitter being involved in the activity.

and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 51) = 4.94, p = .103$ (see table 4.1).

Responsiveness. There was a statistically significant association between the accuracy of predictions relating to the children's responsiveness and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 51) = 9.08, p = .011$ (see table 4.1). The association was of moderate strength (Cohen, 1988), Cramer's $V = .422$. The assessment condition had the highest percentage correct (87.5%) and the no assessment condition the lowest (40.0%).

Resistance to Suggestion. There was no significant association between the accuracy of predictions relating to the children's acquiescence and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 51) = 0.09, p = .955$ (see table 4.1).

Drawing Ability. The analysis showed that 3 cells had an expected count less than 5, so an exact significance test was chosen for Pearson's chi square. There was a significant association between the accuracy of predictions relating to the children's ability to draw a person whereby body parts could be sufficiently identified and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 49) = 7.58, p = .024$ (see table 4.1). The association was of moderate strength (Cohen, 1988), Cramer's $V = .393$. The assessment condition had the highest percentage correct (100.0%) and the colouring condition the lowest (60.0%).

The analysis showed that 3 cells had an expected count less than 5, so an exact significance test was chosen for Pearson's chi square. There was no significant association between the accuracy of predictions relating to the children's ability to use drawings to correctly identify body parts and the type of assessment method (Assessment vs. Colouring vs. No Assessment): $\chi^2(2, N = 50) = 1.02, p = .752$ (see table 4.1).

Table 4.1*Number of Correct and Incorrect Predictions Pertaining to Interview Performance*

Communicative Factor	Condition	Correct	Incorrect
Ground rules	Assessment	14 (87.5%)	2 (12.5%)
	Colouring	5 (33.3%)	10 (66.7%)
	No Assessment	11 (55.0%)	9 (45.0%)
Attention	Assessment	16 (100.0%)	0 (0.0%)
	Colouring	15 (100.0%)	0 (0.0%)
	No Assessment	17 (85.0%)	3 (15.0%)
Responsiveness	Assessment	14 (87.5%)	2 (12.5%)
	Colouring	7 (46.7%)	8 (53.3%)
	No Assessment	8 (40.0%)	12 (60.0%)
Suggestibility	Assessment	6 (37.5%)	10 (62.5%)
	Colouring	6 (40.0%)	9 (60.0%)
	No Assessment	7 (35.0%)	13 (65.0%)
Drawing ability ^a	Assessment	16 (100.0%)	0 (0.0%)
	Colouring	9 (60.0%)	6 (40.0%)
	No Assessment	13 (72.2%)	5 (27.8%)
Ability to use drawing to identify body parts	Assessment	16 (100.0%)	0 (0.0%)
	Colouring	14 (93.3%)	1 (6.7%)
	No Assessment	18 (94.7%)	1 (5.3%)

^a Two children did not draw pictures.

Results Summary

A significant association between the assessment condition and accuracy of predictions was found for the children's use of ground rules. Of the predictions in the assessment condition, 87.5% were correct compared to 33.3% in the colouring condition and 55.0% in the no assessment condition. Significant associations were also found for responsiveness and drawing ability with 87.5% (assessment), 46.7% (colouring), 40.0% (no assessment) and 100.0% (assessment), 60.0% (colouring), 72.2% (no assessment) of predictions correct respectively. No significant associations were found between the assessment condition and accuracy of predictions for attentiveness, resistance to suggestion, and ability to use the drawing to identify body parts. However, the predictions for both attentiveness and ability to use the drawing to identify body parts were almost at ceiling at 100.0% (assessment), 100.0% (colouring), 85.0% (no assessment) and 100% (assessment), 93.3% (colouring), 94.7% (no assessment) respectively.

4.3 Discussion

This is an exploratory study looking at whether a pre-interview communication assessment using the 'Unpacking the Box' framework (Triangle, 2015) achieves its primary purpose, namely to provide a reliable indication of a child's communication abilities. As children develop at different rates (Anderson et al., 2009), it was hypothesised that predictions informed by a communication assessment would be more accurate than those based solely upon professional judgement. The hypothesis was partially supported. Each of the predictions will now be considered in turn.

Use of Ground Rules

Predictions regarding the children's use of ground rules were found to be more accurate when based upon the findings of communication assessments as opposed to professional judgement alone. Knowing whether a child will use the ground rules could be

important in planning an investigative interview. Having an awareness of whether the child will use the 'I don't know' rule will to some extent dictate what questions can be asked (e.g., forced-choice questions - see section 2.2.2 for an overview). Although best practice guidance recommends using these questions very sparingly (MoJ, 2011), forced-choice questions may sometimes be necessary to elicit further information. For example, if an interviewer needs to establish whether it was a single or repeated incident, they may ask 'Did it happen once, more than once, or don't you know?' However, with a child that is prone to guessing and is unable or unwilling to say that they do not know the answer the interviewer needs to consider the benefits of obtaining this additional information against the risks of eliciting potentially inaccurate evidence, that may serve to undermine the child's credibility.

Misunderstandings can also lead to children providing inaccurate evidence and are likely perpetuated in children who have difficulty using the 'I don't understand' rule. As such, extra care should be taken with these children in the framing of questions (i.e., length, complexity, vocabulary). Interviewers should, for example, try to use the child's own words (ABE; MoJ, 2011). This is particularly important when paraphrasing what the child has said. Previous research has found that interviewers occasionally paraphrase children's statements incorrectly (Evans et al., 2010) and confuse details across occurrences (Pichler, 2018). For a child who is reluctant to correct the interviewer or use the 'you got it wrong' rule, such errors can be very problematic potentially creating inconsistencies in the child's accounts. Although it is beyond the scope of this thesis, having some knowledge about whether a child will use the ground rules is also important within a courtroom setting where the child will be cross-examined and their account challenged. If a child is unable to comprehend or use the ground rules, criminal justice practitioners (i.e., police officers, RIs) need to be hypervigilant to the child's non-verbal behaviour as this may allude to the fact that there are difficulties.

Attention

The study found no statistically significant difference in predictions regarding the children's attentiveness based upon the findings of the communication assessments or professional judgement alone. There is however an important caveat when considering this finding. The interviews, in the current study, were very short averaging just over 12 minutes. The interviews were limited to a maximum of 20 minutes in order to accommodate school timetables. A general guideline is that children are able to attend for between 3 and 5 minutes per year of age. A 4-year-old, for example, should be able to attend for approximately 12 to 20 minutes (Schmitt, 1999, as cited in Anderson et al., 2009). It is likely, given this guideline, that the interviews in the current study were too short to present a challenge to the majority of children's attentional abilities. Thus predictions, relating to attention span, were not fully tested. However, a number of children in the study did ask to leave / stop the interview early. These requests emphasise the difficulties associated with generic guidelines and the necessity for pre-interview communication assessments. A very long interview can lead to a child becoming fatigued. Once this happens the child is likely to give the questions less thought and consideration and may even begin to respond to the questions randomly (Anderson et al., 2009). This could potentially jeopardise both the credibility and accuracy of their account. Thus, it is of the utmost importance that interviewers plan an interview in line with a child's attentional abilities and recognise as well as be receptive to signs of fatigue as they appear.

Responsiveness

Predictions regarding the children's responsiveness (conceptualised as the amount of Investigation Relevant Information [IRI] provided to open-ended questions) were found to be more accurate when based upon the findings of communication assessments as opposed to professional judgement alone. This is a very important finding given that longer responses to open-ended questions have been found to result in more convictions (Myklebust & Bjorkland,

2009). Responses to open-ended questions are also considered to be the most accurate as they rely upon free recall memory processes (Orbach & Lamb, 2000). As such, having the means to reliably predict children's responsiveness during an assessment should help interviewers to establish both how likely the child is to provide an accurate account at interview and how likely that account is to result in a conviction.

Establishing a child's responsiveness during an assessment can also help an interviewer plan their questioning accordingly (see section 2.2.2 for an overview of question types) in order to scaffold a child's account. Some children, particularly those that are very young or do not tend to speak a lot, may have difficulty responding to very broad open-ended invitations such as 'Tell me everything that happened?' For these children open-ended depth questions such as 'You said you went to the park. Tell me more about that' or open-ended breadth questions such as 'You said you went on the slide. Then what happened?' may be more appropriate. These questions include pre-disclosed details that can scaffold children's recall by re-focusing their attention (Orbach & Lamb, 2000). Responsiveness can easily be explored with a practice narrative (see section 2.2.1 for more information) during a pre-interview communication assessment. From this a plan can be made for the investigative interview which may, for example, include different strategies to introduce the topic (i.e., alleged incident).

However, responsiveness is not determined entirely by a child's cognitive ability. It is also reliant upon social factors, namely does that child want to engage and disclose the requested information. This situation did not arise in the current study. The information the children were disclosing was neither traumatic nor emotive. Previous research has found that children are reluctant to disclose information of this nature e.g., abuse (Lytle et al., 2019). Feelings of shame, guilt, and a misguided sense of loyalty can all make it very difficult for maltreated children to disclose their experiences (London et al., 2005). An assessment could

provide an additional opportunity to build rapport with a child (Anderson et al., 2014), as well as assess responsiveness, which may further increase the likelihood of a disclosure from reluctant children (e.g., Azzopardi et al., 2019).

Resistance to Suggestion

Predictions regarding the children's resistance to suggestion were equally poor regardless of whether they were based upon the findings of communication assessments or professional judgement. The researcher made six correct predictions in the assessment condition, underestimating the children's abilities on four occasions and overestimating on six occasions. Multiple explanations could account for the researcher's errors. During the communication assessment a practice narrative was used to assess the children's compliance or resistance to suggestion (see section 2.1.2). One explanation that could account for the researcher underestimating the children's abilities was that the event the children chose to recall during the practice narrative may have happened a long time ago, whereas the delay between the staged event and interview was only a week. Research has found that long delays can sometimes impair children's memory (Brubacher, Peterson, et al., 2019).

Conversely, the event chosen during the practice narrative could have led the researcher to overestimate some children's abilities. The children often chose to talk about a holiday or a birthday party they had been to. Children are likely to be familiar and therefore possess prior knowledge about such events. Prior knowledge can affect how information is encoded in memory (Ornstein & Haden, 2002). The more prior knowledge a child possesses the more likely the child is to understand the experience, attend to the salient features, and encode these fully in memory (Ornstein et al., 1997). This results in a stronger memory trace and, as discussed in section 2.1.2, the stronger the trace the less susceptible it is to suggestion (Ceci & Bruck, 1993). The children in the current study, may have possessed a stronger memory trace for their practice narrative event than the Mrs Science Germ Detective

Protocol, as the protocol was likely a far more novel experience. This may have made it more challenging for the children to make sense of the experience and thus encode it fully, leading to a weaker memory trace. This could account for some children being less resistant to suggestion at interview compared to during assessment.

An alternative explanation is that the children experienced greater levels of stress and anxiety at the interview, relative to the assessment. Every effort was made, during the interviews, to make the children feel at ease. However, the interviews did broach the topic of wrongdoing (i.e., Mrs Science broke the rule about touching the children's skin). This may have led to increased anxiety in some children. Anxiety is thought to impair cognitive functioning (Eysenck & Calvo, 1992) and potentially lead to increased suggestibility (Almerigogna et al., 2007). It is likely that this would be perpetuated in real world interviews where the repercussions for the child are much more salient.

However, further research is needed to examine whether increased levels of anxiety were in fact responsible for the null effect observed in the current study or whether this was related to the nature of the practice narrative. It may be that the practice narrative is most appropriate as an assessment tool when the event chosen is similar to the event that will be recalled at interview. Previous research has found incident specific recall (i.e., identified a repeated event and children described the time they remembered the best) to have a positive effect on children's memory for repeated events (Brubacher et al., 2011). As such when choosing an event for the practice narrative an interviewer should take into account how long ago and how many times the event occurred to make sure that this has similar features to the alleged abuse (if known).

Drawing Ability

Predictions regarding the children's ability to draw a person (that is sufficiently detailed to be submitted into evidence) were found to be more accurate when based upon the

findings of communication assessments as opposed to professional judgement alone. Some children may need additional aids (e.g., dolls, drawings, figures) in order to scaffold their communication. Previous research suggests that drawing may be the most effective and safest of these tools (see section 2.2.4 for a review). However, first an interviewer must establish whether a child can draw / produce a picture that is sufficiently detailed and accurate to be submitted into evidence (e.g., relevant body parts are clearly identified). In the current study not all children met age-related expectations (as can be seen from the predictions based on professional judgement). As children have limited attentional resources (Anderson et al., 2009) the interview is not the best forum in which to explore this. During the pre-interview communication assessment is more appropriate.

In the current study correct predictions were almost at ceiling, in all three conditions, for children's ability to use drawings to accurately map touches upon themselves. As such, there was no statistically significant difference in the predictions based upon the communication assessments or professional judgement. This is not particularly surprising given that children are believed to develop representational insight (see section 2.2.4) at around the age of 3 years old (Poole et al., 2011). Hence, representational insight is a cognitive skill that the majority of children, in the current study, will have already acquired and become proficient at leading to a lack of variability in the results. To be able to examine the ability of a communication assessment to accurately assess this area of cognition, future research would need to recruit a younger sample of participants (i.e., 2-year-old and 3-year-old children), with whom this skill is unlikely to be fully developed. Previous research has found that 95% of RIs use communication aids, such as body diagrams and drawings, in their work (Owen, 2016). However, in order for these tools to be effective the RI or interviewer must first ascertain that the child has developed representational insight. Without

representational insight, using these tools to help children to convey information about events is fraught with dangers.

Limitations and Areas for Future Research

There are a number of limitations to the current study. The first limitation is that being an exploratory study the sample size was relatively small and thus the findings need to be interpreted with caution. The sample size was adequate to detect a moderate effect size. Further research utilising a larger and more diverse sample is strongly recommended. A second limitation relates to the length of the interviews. In order to accommodate school timetables, the interviews were limited to 20 minutes. Unfortunately, it appears that this was not sufficient to fully test the predictions pertaining to attention thus it is recommended that future studies employ longer interviews which will be better able to determine whether a pre-interview communication assessment provides an accurate representation of children's attentional abilities. A further limitation of the current study relates to the age of the children. As noted above, predictions of the children's ability to use the drawing to identify body parts were almost at ceiling, making it impossible to determine whether an association truly exists between this and the assessment condition. Therefore, it is recommended that future studies include a cohort of younger children (i.e., 2 to 3 years old) to explore this further. A final limitation of the current study relates to the predictions based in their entirety on professional judgement. The judgements were made by the primary researcher. The researcher has considerable practical and research experience within this field. As such, the researcher's judgements may not be representative or reflect those of police interviewers. Had the predictions been made by police interviewers the benefits of communication assessments, over professional judgement, may well have been magnified. Although this is an interesting avenue for future research, arguably a more pressing issue is whether police interviewers can

be trained to both implement and apply the findings of pre-interview communication assessments.

Conclusions

The main purpose of a pre-interview communication assessment is to gain a better understanding of a child's communication abilities, which can then be used to plan the investigative interview. The current study found that communication assessments, using the 'Unpacking the Box' framework (Triangle, 2015) do provide a good indication of a child's abilities in all areas of cognition examined, other than resistance to suggestion. This may have produced a null effect because of a lack of similarity between the event chosen as part of the practice narrative and staged event. Overall, the findings suggest that a pre-interview communication assessment is better than professional judgement alone in ascertaining children's abilities. It is therefore hoped that assessments using 'Unpacking the Box' (Triangle, 2015) will become more commonplace. Although, extensive training is likely required in order for their true potential, as a tool to improve investigative interviews, to be realised.

Chapter Five: Individual Differences in Children's Recall and Use of Ground Rules in an Investigative Interview Context (Study 3)

It is not uncommon for a criminal case to involve multiple child witnesses all of whom observed the same event (Brubacher, Peterson, et al., 2019). Even when these children are of a similar age there is often a great deal of variability in how the children recall the event at interview (Brubacher, Peterson, et al., 2019; Chae & Ceci, 2005). Some children provide detailed and accurate accounts of past events whilst others recall very little information (Henry, Messer, et al., 2017). This variability has led to a growing body of research into the influence of individual difference factors on children's recall in an interview context. The current study examined three factors for which research has been sparse or previous studies have yielded inconsistent findings: language, visuospatial ability, and attention. The study also explored how these three cognitive factors impact upon children's understanding and use of the 'I don't know', 'I don't understand', and 'you got it wrong' ground rules. To the author's knowledge no previous study has examined the relationship between these cognitive abilities and children's understanding and / or use of ground rules. Having an awareness of how cognitive factors impact upon children's understanding and use of the ground rules is important, as empirical evidence shows that an increased propensity to use ground rules can improve the quality of children's accounts (e.g., Gee et al., 1999; Peters & Nunez, 1999, Saywitz et al., 1999). The practical and theoretical relevance of specific cognitive abilities to children's recall and use of grounds will now be considered in turn.

Language

Language is considered to be a critical component in the development, and subsequent recall, of autobiographical memories (Fivush & Nelson, 2004). Language not only affects how a child encodes and rehearses their past experiences (directly impacting the strength of the memory trace), it also affects the child's ability to understand an interviewer's questions

(and by proxy ground rules instructions), as well as structuring a coherent narrative (Henry, Messer, et al., 2017). Theoretically, these ‘roles’ of language could impact upon both the child’s understanding / use of a ground rule and the quality of information elicited from the child at interview. For example, children with higher language abilities are likely to develop stronger memory traces (Ornstein & Haden, 2002). Stronger memory traces are easier to access (Ornstein & Haden, 2002) and are more resistance to misinformation than weaker traces (Ceci & Bruck, 1993). This may not only result in the child providing a more detailed and accurate account, it could also increase the child’s use of the ‘you got it wrong rule’ as the child may be better able to appraise what has been said by the interviewer with their own memory for the event (as the memory is more accessible). As misleading questions are often synonymous with more complex syntax (e.g., tag questions) higher language abilities may also increase the use of the ‘you got it wrong’ rule simply through increased understanding.

Of the individual difference factors included in the current study, language ability has undergone the most exploration. However, previous research examining the relationship between children’s language abilities and event recall has elicited mixed findings, particularly in relation to the mediating effects of age. Chae and Ceci (2005) explored the effect of verbal intelligence (calculated by averaging scaled scores on two verbal subtests [Similarities and Vocabulary] of the Korean Educational Development Institute – Wechsler Intelligence Scale for Children) on the event recall of pre-schoolers (5- to 6-year-olds) and second graders (7- to 8-year-olds). Verbal intelligence was found to be related to open-ended recall. However, separate regression analyses for each age group showed that this relationship was largely driven by the older children. An earlier study by Burgwyn-Bailes et al. (2001) examined the impact of individual difference factors, including receptive vocabulary, on 3- to 7-year-old children’s recall of an emergency medical procedure. In contrast to Chae and Ceci (2005), Burgwyn-Bailes et al. (2001) found receptive vocabulary to be a significant predictor of the

younger children but not the older children's recall after 1-year. Interestingly receptive vocabulary was not a significant predictor, for either age group, at their initial or 6-week interviews. Lee (2013) also looked at the effects of various individual difference factors, including receptive language ability, on 4- to 9-year-old children's memory of a stressful event. Receptive language ability was again found to be an important predictor of free recall for the younger (4- to 6-year-olds) but not the older children (7- to 9-year-olds). This seems to indicate that language ability is an important indicator of memory recall particularly in younger children.

More recent studies have tended to focus on how language abilities impact upon event memory in younger children. For example, Klemfuss (2015) examined the relationship between two separate facets of language (receptive and expressive) and different types of episodic recall in children 2 to 5 years old. A strong relationship was found between children's language abilities and event recall. Expressive language abilities were associated with accurate free recall and receptive language abilities with resistance to misleading questions. Further, Chae et al., (2014) examined how the vocabulary skills of 3- to 5-year-old children affected their recall of a conflict event (the event involved two actors playing the role of student teachers. Whilst one of the actors was delivering a history lesson, the other actor entered and accused the actor delivering the lesson of taking some craft materials). Children with higher expressive and receptive language skills provided more correct information in free recall and made less errors when answering specific-closed and yes/no questions. A more recent study by Chae et al. (2016), also involving 3- to 5-year-old children, found several measures of language competence (i.e., narrative ability, adaptive language use, receptive vocabulary, and expressive vocabulary) to be related to event memory.

In contrast to most of the recent studies in this area Henry, Messer, et al. (2017) included a cohort of older children. Their study examined the relationship between children's

(6 to 11 years old) interview performance and several individual difference factors including multiple measures of language ability (i.e., receptive vocabulary, recalling sentences, formulated sentences, grammar, and syntax). Although the addition of the language variables improved the overall fit / prediction of their model, none of the variables were significant independent predictors of interview performance. An earlier study by Henry and Gudjonsson (2007) also explored whether there were relationships between standardised measures of cognitive ability, including the BPVSII (i.e., a measure of receptive language), and children's (aged 8 to 9 and 12 years old) eyewitness recall. Children with higher BPVSII scores provided more information in response to general questions (e.g., 'What did they look like?'). However, these children also provided more erroneous details during free recall. Given the variability in the above findings and the purported role of language within event memory further research is essential in order to disentangle the relationship between language and memory recall across both young and old child cohorts (e.g., 4 – 9 years old).

Visuospatial ability

Visuospatial ability refers to a person's capacity to process the location or orientation of objects in space (Irani, 2011). The more proficient a person is at interpreting visuospatial information, the more likely that the information will be encoded, stored, and subsequently retrieved from memory (Lewin et al., 2001). Thus, a child with high visuospatial ability should be able to recall more accurate information about relevant visuospatial information, such as the location of a past event or layout of an area where an incident took place. Additionally visuospatial ability is related to proprioception which refers to the sense of position and movement of the body within space (Renault, et al., 2018) which may be an important ability in determining interactions around and with the child's body. Increased visuospatial awareness may also strengthen the child's memory traces, empowering the child to refute an interviewer's incorrect suggestions pertaining to this aspect of the event (i.e., have

an increased propensity to use the ‘you got it wrong’) – weaker traces are linked to increased suggestibility (Ceci & Bruck, 1993).

Interpreting the current research regarding visuospatial ability and event memory is challenging. Standardised tests that have been found to measure visuospatial ability have also been found / were designed to measure other constructs. For example, Raven’s Progressive Matrices (RPM) measures both visuospatial ability and fluid intelligence – the ability to solve novel problems (Waschi et al., 2017). To the author’s knowledge, no previous studies have reported examining the link between visuospatial ability and event memory. However, Henry, Messer, et al. (2017) utilised a task that is similar to the RPM – the Matrix Reasoning subtest (MR) of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler & Zhou, 2011) – to measure and examine intelligence. MR score was found to be a significant predictor of interview performance (i.e., number of correct responses). That being said, it was not as important a predictor as either age or diagnostic status (i.e., TD or ASD) and was no longer a significant predictor when language and memory abilities were included within the model (Henry, Messer, et al., 2017). Block Design (Wechsler, 1981) is another subtest that is thought to measure visuospatial ability (Kaufman, 2001). Chae and Ceci (2005) included the Block Design subtest from the Korean Educational Development Institute – Wechsler Intelligence Scale for Children in their study, as a measure of visual intelligence. In Chae and Ceci’s (2005) study, scores on the Block Design test failed to predict both children’s open-ended and cued recall. Although Brown and Pipe (2003) also incorporated the Block Design test in their study of individual difference factors on event memory, the children’s scores on the test were summed with other subtests to create a composite intelligence score. Thus, the relative contribution of children’s visuospatial ability on event memory could not be determined. Due to the paucity of research in this area it is important that future research

examine the contribution of visuospatial ability, independent of other measures of intelligence, on children's event memory.

Attention

Attentional processes may influence how an event is encoded in memory (Henry, Messer, et al., 2017). High levels of selective attention abilities may mean children are less distracted by periphery perceptual information during an incident (Ackerman, 1987) and may be less prone to distraction when later interviewed (Lavie, 2005) allowing for more accurate information. In addition to this, loss of concentration / fatigue during an investigative interview can lead to a child responding to questions randomly (Anderson et al., 2009), jeopardising both the detail and accuracy of the child's account. By virtue of responding randomly the child will not be appraising each question against the ground rules instructions, also limiting their utility. Understanding a child's attentional strengths and weaknesses could potentially help interviewers to plan and schedule appropriate breaks during the interview hopefully preventing such undesirable behaviours.

Despite, the potential implications of attention on event memory very little research has been conducted examining the relationship between these two variables. Chae et al. (2016) included a measure of 'attentional focusing' (i.e., questions associated with the child's concentration) in their study of 3- to 5-year-old children's event memory. Attentional focus was found to be positively related to both the amount of correct information the children provided during free recall and the proportion of correct responses the children gave in response to specific-closed and yes/no questions. Measures of attention (i.e., focused attention, sustained attention, sustained-divided attention) were also included within Henry, Messer, et al. (2017). However, none of the attention variables, included in their study, emerged as significant predictors of interview performance. Due to limited research and

conflicting findings, further research is needed in order to determine the nature of the relationship between attention and event memory.

The Current Study

One aim of the current study was to examine demographic (e.g., age and gender) and cognitive (e.g., language, visuospatial ability, and attention) variables thought to be related to children's event recall. The purpose of the study was to identify predictors of children's recall (i.e., detail and accuracy) that could prove helpful for practitioners working within the CJS. Having an awareness of these predictors could allow practitioners to identify where these skills are lacking and when a witness may require additional support / scaffolding in order to provide their best evidence. This was achieved by analysing the data collected as part of study two of this thesis. Children's responses in study two were coded for the presence of IRI. Items of IRI were classified as correct, incorrect, or confabulations. Items were classified as correct if present in the staged event, items were classified as incorrect if inconsistent with the staged event (e.g., "Mrs Science was wearing a red coat." when it was actually white), or items were classified as confabulated if not present in the staged event (e.g., "we painted pictures"). Accuracy was calculated by dividing the number of correct items of IRI by the total number of items of IRI. Four stepwise regressions were conducted with the individual difference factors as the predictor variables and correct IRI, incorrect IRI, confabulated IRI, and accuracy of IRI as the outcome variables. It was hypothesised that:

- In line with previous research (e.g., Henry, Messer, et al., 2017), age would be a significant predictor of correct IRI (i.e., positive relationship), incorrect IRI (i.e., negative relationship), and confabulated IRI (i.e., negative relationship). Due to the current study incorporating misleading questions, it was also expected that age would be a significant predictor of the accuracy of IRI (i.e., positive relationship).

- As the existing findings in relation to language, visuospatial ability, and attention are inconsistent only very tentative predictions could be made, based on theoretical relevance. Thus, it was hypothesised that language, visuospatial ability, and attention would all be significant predictors of children's event recall, following a similar pattern as described for age.

Another aim of the current study was to examine demographic (e.g., age and gender) and cognitive (e.g., language, visuospatial ability, and attention) variables that may potentially be related to children's understanding and use of ground rules (i.e., 'I don't know', 'I don't understand', and 'you got it wrong'). Four stepwise regressions were conducted with the individual difference factors as the predictor variables and use of the 'I don't know' rule, use of the 'I don't understand' rule, use of the 'you got it wrong' rules, and ability to answer practise ground rules questions correctly as the outcome variable. It was hypothesised that:

- In line with previous research (e.g., Brown et al., 2019), age would be a significant predictor of all the outcome variables (i.e., positive relationship).
- As no previous research has examined the relationship between children's cognitive abilities and their understanding and use of the ground rules the following prediction is based solely on the theoretical discussions above. It was hypothesised that language, visuospatial ability, and attention may all be significant predictors of children's use of the ground rules with language also potentially predicting children's understanding of the ground rules (i.e., ability to answer ground rules questions correctly).

5.1 Method

This study involved the re-analysis of the data collected as part of study two as such see section 4.1 of this thesis for a detailed methodology.

Data Analysis

Eight stepwise regressions were conducted and all the relevant assumptions were tested. Linearity was assessed, and confirmed, by plots of studentized residuals against the predicted values. There was homoscedasticity, as indicated by plots of studentized residuals versus unstandardised predicted values. The assumption of normality was also met as assessed by Q-Q Plots. Durbin-Watson statistics demonstrated independence of residuals. An examination of correlations showed that age and receptive language (BPVS3 score), and visuospatial ability (Raven's score) and receptive language (BPVS3 score) were highly correlated. However, as the collinearity statistics (i.e., Tolerance and VIF) were all within the recommended limits, the assumption of multicollinearity was considered to have been met. For the majority of analyses there were no studentized deleted residuals greater than ± 3 indicating no significant outliers. There were no leverage values over 0.2 and no values for Cook's distance above 1 suggesting that there were no highly influential points.

5.2 Results

Predicting the Amount of Correct Investigation Relevant Information Children Provide from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with the amount of correct IRI as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. Stepwise regression was chosen, as opposed to other methods, due to the exploratory nature of the research. Stepwise regression has also been utilised in studies with similar aims (e.g., Roebbers and Scheider [2001] explored the influence of intelligence and shyness on children's eyewitness recall). The current study is considered to be exploratory given the small sample size and the lack of previous research regarding some of the variables

being examined. The correlations between the variables are reported in Table 5.1 and the regression statistics in Table 5.2.

Visuospatial ability entered into the regression model at step 1. The relationship between visuospatial ability and the amount of correct IRI was statistically significant, $R^2 = .38$, $F(1,49) = 29.67$, $p < .001$; adjusted $R^2 = .36$. None of the other predictor variables entered the model: age ($t = 1.41$, $p = .165$), gender ($t = 1.54$, $p = .130$), receptive language ($t = 0.80$, $p = .426$), expressive language ($t = 0.76$, $p = .449$), and attention ($t = 1.05$, $p = .297$).

Table 5.1*Correlations between Predictor Variables and the Outcome Variables*

Variable	Gender	Age	Receptive Language	Expressive Language	Visuospatial ability	Attention
Amount of Correct IRI	.19	.54**	.51**	.35**	.61**	.45**
Amount of Incorrect IRI	-.34**	-.12	-.10	-.39**	-.11	-.23
Amount of Confabulated IRI	-.04	-.19	-.34**	-.03	-.27*	-.21
Accuracy of IRI	.31*	.39**	.37**	.42**	.44**	.46**
Frequency of 'I Don't Know' Rule	.25*	-.11	-.18	.08	-.26*	.08
Frequency of 'I Don't Understand' Rule	.12	-.07	-.03	.17	.10	.18

Variable	Gender	Age	Receptive Language	Expressive Language	Visuospatial ability	Attention
Frequency of 'You Got It Wrong' Rule	.20	.13	.22	.10	.29*	.15
Gender	1.00	.13	.03	.35**	.04	.26*
Age	-	1.00	.74**	.39**	.69**	.48**
Receptive Language	-	-	1.00	.40**	.74**	.44**
Expressive Language	-	-	-	1.00	.44**	.54**
Visuospatial ability	-	-	-	-	1.00	.58**
Attention	-	-	-	-	-	1.00

*Correlation is significant at the .05 level.

** Correlation is significant at the .01 level.

Note. Receptive language relates to British Picture Vocabulary Scale Third Edition score, expressive language to Renfrew Action Picture Test score, visuospatial ability to Raven's Coloured Progressive Matrices score, and attention to Standard Dimensional Change Card Sort score.

Table 5.2*Stepwise Multiple Regression Predicting Amount of Correct Investigation Relevant**Information*

Variable	<i>B</i>	95% CI (LL, UL)	<i>SE B</i>	β	<i>p</i>
Step 1					
Intercept	27.39	17.16, 37.63	5.09	-	.000
Visuospatial ability	1.31	0.82, 1.79	0.24	0.61	.000

Note. Visuospatial ability relates to Raven's Coloured Progressive Matrices score.

Predicting the Amount of Incorrect Investigation Relevant Information Children Provide from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with the amount of incorrect IRI as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1 and the regression statistics in Table 5.3.

Expressive language entered into the regression model at step 1. The relationship between expressive language and the amount of incorrect IRI was statistically significant, $R^2 = .16$, $F(1,49) = 8.98$, $p = .004$; adjusted $R^2 = .14$. None of the other predictor variables entered the model: age ($t = 0.26$, $p = .795$), gender ($t = -1.65$, $p = .105$), receptive language ($t = 0.48$, $p = .636$), visuospatial ability ($t = 0.52$, $p = .606$), and attention ($t = -0.14$, $p = .892$).

Table 5.3

Stepwise Multiple Regression Predicting the Amount of Incorrect Investigation Relevant Information

Variable	<i>B</i>	95% CI (LL, UL)	<i>SE B</i>	β	<i>p</i>
Step 1					
Intercept	10.64	5.88, 15.40	2.37	-	.000
Expressive language	-0.12	-0.20, -0.04	0.04	-0.39	.004

Note. Expressive language relates to Renfrew Action Picture Test score.

Predicting the Amount of Confabulated Investigation Relevant Information Children Provide from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with the amount of confabulated IRI as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1 and the regression statistics in Table 5.4.

Receptive language entered into the regression model at step 1. The relationship between receptive language and the amount of confabulated IRI was statistically significant, $R^2 = .12$, $F(1,49) = 6.59$, $p = .013$; adjusted $R^2 = .10$. None of the other predictor variables entered the model: age ($t = 0.74$, $p = .464$), gender ($t = -0.18$, $p = .855$), visuospatial ability ($t = -0.21$, $p = .834$), expressive language ($t = 0.89$, $p = .379$), and attention ($t = -0.45$, $p = .653$).

Table 5.4*Stepwise Multiple Regression Predicting the Amount of Confabulated Investigation Relevant Information*

Variable	<i>B</i>	95% CI (LL, UL)	<i>SE B</i>	β	<i>p</i>
Step 1					
Intercept	2.46	1.20, 3.73	0.63	-	.000
Receptive language	-0.02	-0.03, -0.00	0.01	-0.34	.013

Note. Receptive language relates to British Picture Vocabulary Scale Third Edition score.

Predicting the Accuracy of Investigation Relevant Information Children Provide from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with the accuracy of IRI as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1 and the regression statistics in Table 5.5.

Attention entered into the regression model at step 1. The relationship between attention and the accuracy of IRI was statistically significant, $R^2 = .21$, $F(1,49) = 12.95$, $p = .001$; adjusted $R^2 = .19$. None of the other predictor variables entered the model: age ($t = 1.50$, $p = .140$), gender ($t = 1.56$, $p = .126$), receptive language ($t = 1.45$, $p = .154$), expressive language ($t = 1.62$, $p = .111$), and visuospatial ability ($t = 1.70$, $p = .096$).

Table 5.5*Stepwise Multiple Regression Predicting Accuracy of Investigation Relevant Information*

Variable	<i>B</i>	95% CI (LL, UL)	<i>SE B</i>	β	<i>p</i>
Step 1					
Intercept	82.08	76.70, 87.46	2.68	-	.000
Attention	3.99	1.76, 6.22	1.11	0.46	.001

Note. Attention relates to Standard Dimensional Change Card Sort score.

Predicting Children's Use of the 'I don't know' Rule from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with children's use of the 'I don't know' rule, during the substantive phase of the interview, as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1. None of the predictor variables entered the model.

Predicting Children's Use of the 'I don't understand' Rule from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with children's use of the 'I don't understand' rule, during the substantive phase of the interview, as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1. None of the predictor variables entered the model.

Predicting Children’s Use of the ‘You Got it Wrong’ Rule from Demographic and Cognitive Variables

A stepwise multiple regression was conducted with children’s use of the ‘you got it wrong’ rule, during the substantive phase of the interview, as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven’s score), and attention (SDCCS score) as predictor variables. The correlations between the variables are reported in Table 5.1 and the regression statistics in Table 5.6.

Visuospatial ability entered into the regression model at step 1. The relationship between visuospatial ability and the use of the ‘you got it wrong’ rule was statistically significant, $R^2 = .09$, $F(1,49) = 4.61$, $p = .037$; adjusted $R^2 = .07$. None of the other predictor variables entered the model: age ($t = -0.74$, $p = .465$), gender ($t = 1.36$, $p = .181$), receptive language ($t = 0.01$, $p = .989$), expressive language ($t = -0.22$, $p = .830$), and attention ($t = -0.14$, $p = .889$).

Table 5.6

Stepwise Multiple Regression Predicting Children’s Use of the ‘You Got It Wrong’ Rule

Variable	<i>B</i>	95% CI (LL, UL)	<i>SE B</i>	β	<i>p</i>
Step 1					
Intercept	1.79	1.00, 2.59	0.40	-	.000
Visuospatial ability	0.04	0.00, 0.08	0.02	0.29	.037

Note. Visuospatial ability relates to Raven’s Coloured Progressive Matrices score.

Predicting Children's Ability to Answer Practise Ground Rules Questions Correctly from Demographic and Cognitive Variables

A stepwise binomial logistic regression was conducted with children's ability to answer practise ground rules questions correctly on the first attempt as the outcome variable and age, gender, expressive language (RAPT score), receptive language (BPVS3 score), visuospatial ability (Raven's score), and attention (SDCCS score) as predictor variables. None of the predictor variables entered the model.

Results Summary

Expressive language entered into the regression model for predicting incorrect IRI elicited from the children whilst receptive language entered into the model for predicting confabulated IRI. Visuospatial ability entered into both the regression model for predicting correct IRI and for the model predicting children's use of the 'you got it wrong' rule. Furthermore, attention entered into the model for predicting the accuracy of the IRI the children provided. No significant predictors emerged for children's ability to answer practise ground rules questions correctly or for their use of the 'I don't know' and 'I don't understand' rules.

5.3 Discussion

The current study sought to provide further insights into which variables (i.e., demographic and cognitive variables) relate to children's event recall and understanding / use of the ground rules during an investigative interview. Based on previous research and psychological theory, hypotheses one and three both expected age to enter the models as a significant predictor. Neither of these hypotheses were supported as age did not enter as a significant predictor in any of the stepwise regression models conducted as part of the current study. In contrast, hypotheses two and four, both of which related to the cognitive variables, were partially supported. None of the demographic nor cognitive variables entered the

regression models for children's ability to answer the practise ground rules questions correctly (a measure of understanding / comprehension of the rules), use of the 'I don't know' and use of the 'I don't understand' rules. However, visuospatial ability was found to be a significant predictor of correct IRI, expressive language a significant predictor of incorrect IRI, receptive language a significant predictor of confabulated IRI, attention a significant predictor of the accuracy of IRI, and visuospatial ability a significant predictor of the use of the 'you got it wrong rule'. These findings would suggest that cognitive factors may in fact be more important than age in predicting children's event recall and use of the 'you got it wrong' rule. Each of the cognitive predictors and the aforementioned outcome variables will now be considered in turn.

Expressive Language

The current study found expressive language to be a significant predictor of incorrect IRI, with expressive language accounting for 15.5% of the variability in incorrect IRI. As hypothesised, children with higher expressive language abilities gave less incorrect information during the investigative interview. This finding is important for practice as expressive language or responsiveness, as it is referred to in study two of this thesis, can be quickly assessed using the 'Unpacking the Box' framework. The framework provides a far more accurate representation of a child's expressive language abilities than professional judgement alone. The significant relationship between expressive language and incorrect IRI, found in the current study, bolsters the author's earlier suggestion that 'Unpacking the Box' (Triangle, 2015) could be used as an effective means of establishing the likelihood with which a child can provide an accurate account at an interview.

The relationship between expressive language and incorrect IRI could be explained by children with superior expressive language abilities being better able to articulate themselves and thus more likely to produce high-quality narrative accounts (in terms of length,

descriptive texture, and cohesion). Narrative quality has, in previous research, been linked to memory errors. For example, a study by Chae et al. (2016) found that children with more advanced expressive language abilities produced higher-quality narratives, and those children who produced higher-quality narratives made fewer errors.

Alternatively, the relationship between expressive language and incorrect IRI may be explained by children with higher expressive language abilities possessing more efficient memory strategies (Henry, Messer, et al., 2017). More specifically, children with higher expressive language abilities may be better able to encode information in a verbal format and more effective at rehearsing their past experiences (Henry, Messer, et al., 2017). Therefore, enabling more information to be encoded in memory, potentially resulting in more information being available for the children to access at the point of retrieval. This explanation, although viable remains tentative as the current study did not examine other factors (e.g., working memory; Vugs et al., 2014) that are known to be closely linked to both expressive language and memory. As such further research is needed to understand the relationship between expressive language, as part of a complex cognitive profile, and incorrect IRI.

Receptive Language

The current study found receptive language to be a significant predictor of confabulated IRI, with receptive language accounting for 11.8% of the variability in confabulated IRI. As hypothesised, children with more advanced receptive language abilities produced less confabulated information. This is in line with the findings of Klemfuss (2015) and Chae et al. (2014). Chae et al. (2014) found that three- to five-year-old children with higher receptive language abilities produced less commission errors (i.e., reporting that something occurred or was present during the event when it was not) when asked yes/no and specific-closed questions, than those children with lower receptive language abilities.

Similarly, Klemfuss (2015) reported that children, aged two to five years old, with higher receptive language abilities were more resistant to misleading questions. In the current study, confabulated information was most often elicited from misleading questions. Misleading questions are often synonymous with complex syntax (e.g., tag questions; Davies & Seymour, 1998). It appears in the current study, that children with more advanced receptive language abilities were more likely to be able to understand these syntactically complex questions. As such, the children with higher receptive language abilities were more likely to understand these questions and provide a correct response as opposed to ceding to the interviewer's incorrect suggestions. However, the converse could potentially be true when interviewing a child who is very eager to please or compliant. In this case, higher receptive language abilities may lead to the child providing more confabulated information due to the child understanding the question and, by virtue of this, the response the interviewer is looking for. Due to the perceived authority of the interviewer some children may actively seek to acquiesce, despite knowing the information to be incorrect (Bruck & Ceci, 1999). Children may adopt these behaviours out of fear of getting in trouble or due to negative encounters with the police in the past (Collins, 2012).

Visuospatial ability

The current study found visuospatial ability to be a significant predictor of correct IRI, with visuospatial ability accounting for 37.7% of the variability in correct IRI. In line with hypothesis two, children with better visuospatial abilities produced more correct IRI. This relationship may have been particularly pronounced in the current study due to the inclusion of a number of location-based and item-based questions, which could have potentially encouraged the children to mentally reinstate the physical context of the event. Mental reinstating the context of an event has long been established as a means of enhancing retrieval (Wheeler & Gabbert, 2017). As such, it is one of the techniques incorporated into the CI

(Fisher & Geiselman, 1992). Although, the current study, did not directly employ this technique, by asking the children to describe the room in which the event took place the children may have themselves mentally recreated the context. An interesting avenue for future research would be to explore whether techniques to mentally reinstate the physical context of an event are more effective with children who possess more advanced visuospatial abilities and whether visuospatial ability is related to other types of recall that may employ visuospatial skills (i.e., face identification).

Irrespective of whether the children were recreating the physical context of the event, it is likely that the children with higher visuospatial abilities were better able to process and subsequently encode location-based information, which will then have increased the likelihood of this information subsequently being recalled at interview (Lewin et al., 2001). Having access to this information, in memory, may have also cued further relevant information. For example, if a child could recall the locations and items present at each workstation this could have potentially cued further information such as what occurred at each station, what was said etc. The distinctive nature of the three workstations (in terms of what items were involved / layout) may have also facilitated the recall of the children with higher visuospatial abilities. This is due to cues which are distinct from one another, being more likely to lead to the retrieval of additional information (Wheeler & Gabbert, 2017).

The current study also found visuospatial ability to be a significant predictor of the use of the 'you got it wrong' rule, with visuospatial ability accounting for 8.6% of the variability in the use of the 'you got it wrong' rule. The children with higher visuospatial abilities used the 'you got it wrong' rule more frequently during the interviews, which could be related to the ability of these children to encode location-based information (Lewin et al., 2001). Their superior encoding skills may have resulted in stronger memory traces (Ornstein & Haden, 2002). Stronger memory traces are easier to access (see section 2.1.1) and are more resilient

to suggestion (Ceci & Bruck, 1993), potentially resulting in children with higher visuospatial abilities being more likely to correct the interviewer. The current study design may have amplified this relationship as a high proportion of the misleading questions required item-based knowledge.

Interestingly, the ‘you got it wrong rule’ was the only ground rule, examined in the current study, for which a significant predictor was found. None of the demographic or cognitive variables entered the regression model for use of the ‘I don’t know’ or use of the ‘I don’t understand’ rule. One explanation for the lack of significant predictors could be that the study did not include the cognitive skills (e.g., ToM, Brubacher et al., 2015) relevant to the acquisition and use of the ‘I don’t know’ and ‘I don’t understand’ rules (see section 2.2.3). A second explanation is that children with more advanced cognitive abilities may, as hypothesised, be more likely to alert the interviewer when they don’t know the answer or don’t understand a question. However, due to their more advanced cognitive abilities these children may need to use the rules less often (as they will be more likely to comprehend the questions and have access to the requested information). On the other hand, children with less advanced language abilities may be less likely to use the rules, but may need to use them more often thus negating any significant relationships. The ‘you got it wrong’ is not beholden to such caveats because if the interviewer asks a misleading question, the correct response, regardless of the of the child’s cognitive abilities, is ‘you got it wrong’.

Attention

The current study found attention to be a significant predictor of the accuracy of IRI, with attention accounting for 20.9% of the variability in accuracy of IRI. In line with hypothesis two, children with higher attentional abilities were more accurate in their recall of the staged event. This may be due to children with superior attentional abilities being better able to attend to and focus on the salient features of the staged event, resulting in more

information about the event being processed and encoded in memory (Ornstein & Haden, 2002). Alternatively, it could be due to children with higher attentional abilities being better able to attend to the questions asked at interview and less likely to respond randomly (Anderson et al., 2009) or rely on response biases (e.g., ‘yes’ bias; Moriguchi et al., 2008) which can reduce the accuracy of the information elicited. As such poor attentional abilities may be particularly problematic when young children are presented with a series of either yes/no or forced-choice questions. These questions require very little effort in order to generate a response as the response options are pre-determined by the interviewer (see section 2.2.2 for an overview of the retrieval effort / processes associated with different question types). Thus, the deleterious effects of low attentional abilities (i.e., less accurate accounts) may be more pronounced in interviews that deviate more markedly from best practice guidance (ABE; MoJ, 2011).

Limitations and Areas for Future Research

The main limitation of the current study, being exploratory, is the small sample size. Due to the small sample size the power of this study is low and, as such, some significant predictors may have been missed. This may explain why age did not enter as a significant predictor in any of the regression models in the current study despite having emerged as a significant predictor in previous research (e.g., Chae et al., 2016; Henry, Messer, et al., 2017). A further limitation of the current study is the omission of other potentially important cognitive variables, such as ToM and metacognitive abilities, that have been theorised to relate to the acquisition and use of ground rules (Brubacher et al., 2015). It is thus recommended that future research, along with using a larger sample size, should incorporate these cognitive variables in the analysis. Social factors, such as compliance, should also be considered in future research as an investigative interview, with its multiple parties, is by its

very nature a social context where social factors such as normalised behaviours, perceived authority, and social stereotypes come into play.

Conclusions

In conclusion, the current study would suggest that cognitive factors may be more important than age in predicting children's event recall and use of the grounds. Interestingly, different cognitive abilities appear to underpin different aspects of children's recall (i.e., correct, incorrect, confabulated, and accuracy of IRI) and use of the ground rules (i.e., 'you got it wrong' rule). This indicates that only relying on age as an indicator of a child's abilities is not appropriate and a more nuanced understanding of their differing cognitive profiles is required. Although explanations have been provided to account for these relationships further research, incorporating a larger sample size, is important to provide further insights into the mechanisms underpinning these relationships. Having an awareness of these mechanisms is important as it can help professionals, such as police officers and RIs, to decide which factors to prioritise during their pre-interview assessments. It is promising that the current training for RIs covers how to assess receptive language, expressive language, and attentional abilities (e.g., Collins & Krähenbühl, 2020). However, consideration may also have to be given to assessing visuospatial ability as the results of the current study would suggest that this is an important predictor of children's recall.

Chapter Six: The Role of Registered Intermediaries in Scaffolding Children's Communication in Investigative Interviews (Study 4)

Since the inception of the WIS in 2004 there has been very little research into the work of RIs (Collins et al., 2017). Much of the early research, as discussed in chapter two, focused upon legal practitioners' perceptions of the role (e.g., Plotnikoff & Woolfson, 2007) and the RIs' experiences (e.g., Cooper, 2009). Although these studies have continued (e.g., Agneswaran, 2018; Collins & Krähenbühl, 2020; Cooper, 2014), the research has also diversified. Studies have begun to examine how the presence of an RI impacts upon jury decision making (e.g., Collins et al., 2017; Ridley et al., 2015), along with RIs' perceptions of developmentally-appropriate questioning (e.g., Krähenbühl, 2011).

In addition, recent research has looked at the impact of the RI provision upon the ability of children to provide best evidence within the course of a mock criminal investigation. More specifically, studies have examined the effect of an RI on children's communication in an investigative interview (Henry, Crane, et al., 2017), on children's ability to correctly identify a perpetrator from an identification parade (Wilcock et al., 2018), and on children's propensity to refute misleading questions during cross-examination (Henry et al., 2021). Although all of these studies provide some insight into the efficacy of the RI role, all are limited due to their experimental nature. For ethical reasons, experimental research often lacks many of the salient features of a real-world criminal investigation, most notably the trauma associated with being a victim or witness to a crime (Hershkowitz et al., 2014). As such, research exploring the role of the RI in real-world contexts (i.e., during an ABE interview or in court) is essential. Without this research, it remains difficult to determine how RIs impact upon interview practice and whether they do succeed in scaffolding children's communication.

The Current Study

The current study sought to address this gap in the literature by analysing real-world investigative interviews with and without an RI. The study compared question types, communications aids, and levels of rapport across RI and no RI interviews. It also looked at how these factors, along with the presence of an RI (including any interventions), affect the amount of information children provide. The primary aims of the current study were to identify what works when interviewing children i.e., which techniques provide the most detailed accounts from child witnesses via police officers and / or RIs; and to provide an indication of the benefits of the WIS and examine the aspects that may need to be improved. It was hypothesised that:

- There would be a positive correlation between open-ended questions and IRI, and specific-closed questions and IRI. Open-ended questions and specific-closed questions are considered to be the best from an evidential perspective as these questions allow children to generate their own responses and have been found, in previous research, to elicit the most detailed accounts (e.g., Gagnon & Cyr, 2017; Luther et al., 2015).
- Open-ended questions would, as in previous research, constitute less than 25% of the total questions asked by the interviewer (e.g., Johnson et al., 2015; Luther et al., 2015). It is unlikely that the presence of an RI will increase the use of open-ended questions given that Krähenbühl (2011) found that the majority of alternatives provided by the RIs, in her study, were in a closed format. The findings of study one of this thesis also suggest that RIs lack knowledge pertaining to best practice guidance (ABE; MoJ, 2011), question types, and associated memory processes.
- Interviews in which communications aids are used will elicit more IRI from the children than those without aids. Previous research has found that both body diagrams

(e.g., Aldridge et al., 2004; Dickinson & Poole, 2017) and drawing (Butler et al., 1995; Woolford et al., 2015) can lead to children providing more detailed accounts.

- The RI interviews would be more likely to use communication aids compared to those without an RI. The use of communication aids is very prevalent amongst RIs, with 95% of RIs, in a recent survey, reporting using these in their work (Owen, 2016).
- There would be a positive correlation between both non-verbal rapport measures (attention and positivity) and IRI, with rapport and the related concept of social support having been found to increase children's informativeness in previous studies (e.g., Blasbalg et al., 2018; 2019).
- Higher levels of rapport would be observed in the interviews with an RI as opposed to those without. Part of the RI's role is to conduct a pre-interview assessment of the witnesses' communication (MoJ, 2020a). This often takes place in collaboration / alongside the interviewing officer. As such the officer and RI will have likely spent more time with the child prior to the interview and the child is thus likely to feel more at ease with both the interviewer and interview process.
- The children will provide significantly more IRI in the interviews with, as opposed to those without an RI. Henry, Crane, et al. (2017) found that TD children assisted by an RI reported 18.96 more items of correct information when interviewed than those children who were unassisted. Although similar benefits were not reported, in Henry, Crane, et al.'s (2017) study, for children with ASD, benefits may emerge in real-world interviews. In real-world interviews the RI's role may involve informing the police about ASD or other additional needs, facilitating rapport, familiarising the child with the investigative process and environment - all factors which could ultimately impact on the detail of the child's account.

6.1 Method

6.1.1 Sample

ABE interviews were obtained from two police forces in England. Forty-two ABE interviews were included in the final sample. Thirty were from a force in the North of England and 12 from a force in the South of England. All of the interviews were conducted since 2011, classified as closed, involved children under the age of 18, and related to allegations of sexual abuse, physical abuse, and / or neglect. An RI was present in half of the interviews ($n = 21$). The interviews with and without an RI were matched as closely as possible in terms of case (i.e., nature of the alleged offence, intra-familial / extra-familial abuse, single / repeated abuse), witness characteristics (i.e., age, gender), and interviewer characteristics (i.e., gender, force). The interviews were only matched within, not across, the two forces to control for differences in training, guidance, and practice in general.

Of the 21 interviews without an RI, 17 involved allegations of sexual abuse, two allegations of physical abuse, and two multiple forms of abuse. Sixteen interviews involved allegations of repeated abuse whilst five related to a single incident. Fourteen interviews involved allegations of intra-familial abuse and seven to allegations of extra-familial abuse. All of the interviews involved different children. The children ranged in age from 4 to 14 years old, with a mean age of 9.67. Seven of the children were male and 14 were female. None of the children were reported as having a concurrent vulnerability (e.g., ASD, cerebral palsy, learning difficulties). The interviews involved 14 different interviewers, each conducted a minimum of one and a maximum of three interviews. Six of the interviewers were male and eight were female. All of the interviewers were trained to at least PIP level 2. Nine of the interviews had a social worker present. For one of the interviews only a transcript was available, not the accompanying DVD.

All of the 21 interviews with an RI involved allegations of sexual abuse. Ten interviews involved allegations of repeated abuse whilst 11 related to a single incident. Eight interviews involved allegations of intra-familial abuse and 13 to allegations of extra-familial abuse. All of the interviews involved different children. The children ranged in age from 4 to 14 years old, with a mean age of 9.33. Five of the children were male and 16 were female. Five of the children were reported as having a concurrent vulnerability (e.g., ASD, cerebral palsy, learning difficulties). No additional needs were identified for the other 16 children. These children were eligible for RI assistance based upon their age (i.e., under 18). Statistics show that the largest proportion of RI referrals are based upon aged-based vulnerability¹⁰. The WIS' Annual Report 2019/20 shows that 35% of referrals in 2019/20 were made for children without a concurrent vulnerability. Therefore, the sample used in the current study can be considered as being fairly representative of RI referrals in 2019/20. The interviews involved 18 different interviewers, each conducted a minimum of one and a maximum of two interviews. Five of the interviewers were male and 13 were female. All of the interviewers were again trained to at least PIP level 2. None of the interviews had a social worker present. The interviews involved 14 different RIs, each assisted with between one and three interviews. Two of the RIs were male and 12 were female. Two had a background in speech and language therapy and one in teaching. No information was provided regarding the backgrounds of the remaining 11 RIs. A report was available for seven of the 21 RI interviews. For two of the interviews only a transcript was available, not the accompanying DVD.

¹⁰Not all children under the age of 18 require RI assistance. Whether a child requires assistance will depend upon many factors including their cognitive abilities and the level of trauma the child has experienced.

For context, at the start of 2018 there were 183 RIs on the MoJ Register (Plotnikoff & Woolfson, 2019). One hundred and fifty-five RIs were registered to accept cases involving witnesses under the age of 18. However, of these, only 91 (59%) were ‘active’ (available to accept cases; Plotnikoff & Woolfson, 2019). In terms of regional availability, in 2018, approximately 45 RIs were registered to accept cases in the two police forces involved in the current study (Newlove, 2018). Although a tentative estimation, this would suggest that the current study evaluated the practice of approximately 30% of RIs operating in these regions (the majority of interviews included in the current study were conducted from 2018 onwards [$n = 35$], see Appendix F). However, it is important to note that not all of the RIs in these areas will accept cases for children under 18, so the percentage is likely an underestimate for the population under investigation.

The sample size was determined by a priori power analysis (G* Power). Due to the applied nature of the research (e.g., Hoogesteyn et al., 2020) the study sought to achieve enough power to detect a large effect size. Using an alpha level of .05 and power of .8, to achieve a large effect size a total of 23 interviews were required for a correlation, 32 for a chi-square, and 42 for an independent-samples t-test (i.e., the main analyses in this study). The number of interviews included in the current study is also comparable to (exceeded) that used in some previous published research (e.g., Feltis et al., 2010; Hill & Davies, 2013; Korkman et al., 2006; Phillips et al., 2012).

6.1.2 Procedure

Ethical approval was granted from Teesside University ethics committee. A number of police forces, across England, were contacted. Of those that responded almost all declined to participate in the study due to a lack of resources. However, two police forces, one in the North and one in the South of England did agree to provide interviews for the study. Once all vetting procedures and agreements were drawn up and finalised, a designated police officer

from each of the forces identified interviews that met the inclusion criteria and compiled these for review and matching. The interviews / cases with an RI were selected / matched as closely as possible to cases without an RI. Cases were matched in terms of case (i.e., nature of the alleged offence, intra-familial / extra-familial abuse, single / repeated abuse), witness characteristics (i.e., age, gender), and interviewer characteristics (i.e., gender). A coding scheme was developed and applied to each interview. Each interview was coded for question types, communication aids, rapport, and IRI.

6.1.3 Coding

Interviewer Questions. These were coded according to ABE guidance (MoJ, 2011). ABE identifies five question types: open-ended, specific-closed, forced-choice, multiple, and leading. Following a discussion with a practitioner / researcher involved in the development of the ABE guidance (MoJ, 2011), yes/no questions were included as an additional category. This is because the structure, biases, and memory processes associated with yes/no questions, are not full encompassed within any of the question types outlined in the ABE guidance (MoJ, 2011). The number of indirect speech acts was also noted (it was felt that this would be important for the purposes of feedback and training that will be given to the police). Indirect speech acts directly ask if the child knows, whilst indirectly asking what they know and can thus serve to elicit a simple 'yes' or 'no' response (Evans et al., 2014). As such, indirect-speech acts can effectively be seen as a subcategory of yes/no questions, and were therefore included within the total for this question type. Descriptions and exemplars of each of the question types are provided in table 6.1. The total number of questions was calculated for each interview, along with the total number of questions within each category.

Table 6.1*Descriptions and Exemplars of the Different Question Types.*

Question	Description	Examples
Open-ended	Does not restrict the witness' response enabling the witness to control the flow of information.	'Tell me everything that happened'
Specific-closed	Specifies the nature of information required from the witness.	'Who', 'What', 'Where', 'When', 'Why' (5WH)
Forced-choice	Forced-choice questions present the witness with a small number of alternatives to choose from and may not include the correct response.	'Was the car red or blue, or don't you know?'
Yes/no	Yes/no questions demand a 'yes' or 'no' response.	'Was the car red?'
Indirect speech acts	Indirect speech acts directly ask if the child knows, whilst indirectly asking what they know.	'Can you remember what colour the car was?'
Multiple	Request multiple pieces of information from the witness at once.	'Did you see him? Where was he? 'What was he wearing?'

Question	Description	Examples
Leading	Imply answers and assume facts that are potentially in dispute. A question can be leading based on the question's structure, tone, or context (e.g., the question may incorporate previously undisclosed information).	'The car was red, wasn't it?'

RI's Questions. These were similarly coded as open-ended, specific-closed, forced-choice, multiple, leading, and yes/no (including in-direct speech acts). However, only questions directed from the RI specifically to the child were coded (not suggestions to the interviewer to rephrase).

Use of Communication Aids. These were categorised as follows: dolls¹¹, models, or figures (e.g., drawing mannequins); drawing / writing (e.g., sketch plans); body diagrams; ground rules cards; calming toys (e.g., colouring, fiddle toys); and 'other' (e.g., timeline of interview procedure, emotions scale to describe how the child felt during the alleged offence, cards depicting times of the year). Each type of communication aid was coded as being 'used' or 'not used' within each interview.

The function of the communication aids was also recorded. Functions were adapted from the Advocates Gateway Toolkit on the use of communication aids (Mattison, 2015) and the ABE guidance (MoJ, 2011). The function was categorised as:

- Information gathering - the aid was used as a means of eliciting additional information from

¹¹ Anatomical dolls were not used in any of the interviews included in the current study. Wooden drawing mannequins with moveable limbs were used.

the child.

- Clarification - the aid was used as a means of understanding information that the child had previously disclosed.
- State management - the aid was used as a means of managing the child's state (e.g., maintaining attention, reducing anxiety).
- Other - the aid was used for another purpose not defined above.

Rapport. A 15 second extract was taken from the beginning, middle, and end of each interview. These extracts were used to rate two non-verbal indicators of rapport – positivity and attention (Johnston et al., 2019). Coordination was not rated as the interviewers and RIs were often not visible on camera so it would have been difficult to determine how synchronous the interaction was. Each extract was viewed twice, both in silent mode. This is a practice endorsed by Johnston et al. (2019). The rationale behind it is that verbal information can interfere with the processing of non-verbal information. The extracts were rated, by the primary researcher, on a scale of 1-9: 1) how expressive the child was (i.e., indicative of positivity) 2) how attentive the child was. The scores from the beginning, middle, and end of each interview were summed to give a total score for expressivity and attention. Expressivity refers to the level of positivity in the interaction. It was exemplified by the following nonverbal behaviours: smiling, head nodding, affect, forward lean, and facial and hand gestures (Johnston et al., 2019). Attention refers to the level of interpersonal engagement with and interest in the interaction. Attention was inferred through the level of eye contact, body orientation, and proximity (Johnston et al., 2019). The maintenance of eye contact and gaze aversion were considered in the appraisal of attention. It is not uncommon for children to have eye contact whilst listening but to avoid eye contact when recalling a past event, due to the increased cognitive load (Doherty-Sneddon & Phelps, 2005).

RI Interventions. Five categories of intervention were identified. Interventions were coded as relating to breaks, questioning, communication aids, procedural / state management, and, understanding. Examples of each of the types of intervention are provided in table 6.2. The total number of interventions was calculated for each interview, along with the total number of interventions within each category. One intervention potentially involved the RI asking multiple questions. For example, if something the child said was unclear, it may have taken more than one question from the RI to resolve this misunderstanding.

Table 6.2*Description of the Different Interventions*

Intervention	Examples
Breaks	<ul style="list-style-type: none"> ● The RI asked the child if they needed a break. ● The RI recommended to the interviewer that a break was needed.
Questioning	<ul style="list-style-type: none"> ● The RI rephrased an interviewer's question. ● The RI suggested to the interviewer a line of questioning. ● The RI asked the child a question.
Communication Aids	<ul style="list-style-type: none"> ● The RI suggested the use of a communication aid. ● The RI facilitated or explained a communication aid to the child.
Procedural / State Management	<ul style="list-style-type: none"> ● The RI explained to the child about where to sit in order to be seen by the cameras. ● The RI encouraged the child to listen to the interviewer's questions. ● The RI explained that the child would not get in trouble and this was their opportunity to speak.
Understanding	<ul style="list-style-type: none"> ● The RI repeated back what the child had said if it was not intelligible. ● The RI checked or advised the officer to clarify what the child had said. ● The RI asked the officer to clarify the meaning of a question (e.g., which occurrence or who they were referring to).

Note. RI is the abbreviation for Registered Intermediary

Children's Responses. Children's responses were coded for the presence of items of IRI. The study used the coding scheme adopted by Philips et al. (2012), with two additional categories – conversation and body parts. The following categories of information were coded: Person, Action, Conversation, Location, Item, Body Parts, and Temporal (see table 6.3). Each item of IRI was counted once and repetitions were ignored. All items of IRI in each category were summed to give a total score for the category. All categories were then summed to give a total IRI score. Conversation was included as an additional category as previous research has shown that conversational details are an important factor to consider when establishing the veracity of a witness' claims (Hunt & Bull, 2011). Hunt and Bull (2011) found that more conversational details (i.e., offender utterances) are reported in genuine compared to false allegations of rape. False allegations were generally found to contain few or no offender utterances. During investigative interviews of alleged CSA children are often asked about conversations with the alleged perpetrator (to determine whether touching occurred as part of routine caregiving or had sexual intent) and others (to identify a previous disclosure; Lawson & London, 2015). As such, information about conversations can be highly influential in steering the direction of an investigation into potential CSA. Body parts was also included as an additional category as determining the specific body parts involved in an offence is often essential in securing a successful prosecution (Burrows et al., 2017).

Table 6.3*Investigation Relevant Information Coding Scheme (Phillips et al., 2012)*

IRI	Code	Description
Person	P	Child mentions information about the alleged perpetrator e.g., 'he has brown hair'.
Action	A	Child provides information that relates to an action e.g., 'he grabbed me'.
Conversation	C	Child describes what another has said e.g., 'he told me not to tell anyone'.
Location	L	Child gives relevant information about a location (before / during / after event).
Item	I	Child provides information about objects used in the commission of the event e.g., phone.
Temporal	T	Child provides information associated with the time of the event e.g., 'it was during the six-week holidays'.

IRI	Code	Description
Body Part	B	Child provides information associated with a body part e.g., 'hand'.

Note. IRI is the abbreviation for Investigation Relevant Information.

The RIs' reports were coded for the following:

Communication being Assessed. The RI reports were categorised as either having or not having assessed the three main facets of communication (i.e., attention, anxiety, and behaviour; expressive communication; and receptive communication).

Assessment Tasks. The assessment tools / tasks were coded as standardised assessment tools (i.e., a test that is administered and scored in a consistent manner such as BPVS3), non-standardised assessment tools (i.e., examines an individual's performance and is not concerned with the production of a score), combination, or not stated.

Recommendations. The reports were coded as either having or not having made recommendations regarding each of the following: the use of communication aids (e.g., dolls, drawing), the type and structure of questions (i.e., length, complexity), language (i.e., simple, non-literal), the frequency of breaks / pace of questioning / signposting of topics, procedural factors (i.e., ground rules, truth and lies, introductions), state management (i.e., how to manage anxiety), and how to phrase questions related to time or distance.

6.1.4 Inter-Rater Reliability

One rater (other than the primary researcher) independently coded 10% of the interviews. The study achieved an inter-rater reliability (percent agreement) of .95 for

question types, .89 for communication aids, .86 for IRI, .80 for RI interventions, and .77 for rapport. Any differences between raters were resolved by discussion.

6.2 Results

Preliminary analyses were conducted to determine whether there were any significant differences between the interviews with and without an RI in terms of witness, interviewing officer, case, and interview characteristics.

Witness Characteristics

Age. An independent-samples t-test was run to determine if there were differences in the children's ages between the RI and no RI interviews. There were no outliers in the data as assessed by a boxplot. Age for each interview condition was normally distributed as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .456$). There was not a statistically significant difference in age between the RI ($M = 9.33$, $SD = 2.85$) and no RI interviews ($M = 9.67$, $SD = 3.09$), $t(40) = -0.36$, $p = .718$.

Gender. A chi-square test of association was conducted between witness gender and the interview condition. All expected cell frequencies were greater than five. There was no statistically significant association between witness gender and the interview condition: $\chi^2(1, N = 42) = 0.47$, $p = .495$.

Interviewer Characteristics

Gender. A chi-square test of association was conducted between interviewer gender and the interview condition. All expected cell frequencies were greater than five. There was no statistically significant association between interviewer gender and the interview condition: $\chi^2(1, N = 42) = 0.93$, $p = .334$.

Case Characteristics

Type of abuse. The analysis showed that 4 cells had an expected count less than 5, so an exact significance test was selected for Pearson's chi square. There was no statistically significant association between the type of abuse and the interview condition: $\chi^2(2, N = 42) = 4.42, p = .107$.

Repeated vs. Single Abuse. A chi-square test of association was conducted between repeated vs. single abuse and the interview condition. All expected cell frequencies were greater than five. There was no statistically significant association between whether the abuse was repeated vs. single and the interview condition: $\chi^2(1, N = 42) = 3.64, p = .057$.

Intra- vs. Extra-Familial Abuse. A chi-square test of association was conducted between intra- vs. extra-familial abuse and the interview condition. All expected cell frequencies were greater than five. There was no statistically significant association between whether the abuse was intra- vs. extra-familial and the interview condition: $\chi^2(1, N = 42) = 3.44, p = .064$.

Interview Characteristics.

Length of Interviews. An independent-samples t-test was run to determine if there were differences in length between the RI and no RI interviews. There were no outliers in the data as assessed by a boxplot. Length for each interview condition was normally distributed as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .587$). There was not a statistically significant difference in length between the RI ($M = 39.57, SD = 16.34$) and no RI interviews ($M = 37.62, SD = 14.67$), $t(40) = 0.41, p = .686$.

Total Number of Questions. An independent-samples t-test was run to determine if there was a difference in the total number of questions between the RI and no RI interviews. There were three outliers in the data as assessed by a boxplot. Removing the outliers had no

appreciable effect on the analysis so the decision was made to include the outliers. Total number of questions for each interview condition was normally distributed as assessed by Shapiro-Wilk test ($p > .05$). There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .345$). There was not a statistically significant difference in the total number of questions between the RI ($M = 98.33$, $SD = 43.80$) and no RI interviews ($M = 113.05$, $SD = 51.93$), $t(40) = -0.99$, $p = .327$ ¹².

DVD of ABE Available. The analysis showed that 2 cells had an expected count less than 5, so a Fisher's Exact Test was conducted. There was no statistically significant association between whether a DVD of the ABE was available and the interview condition, $p = 1.000$.

Exploring Practices to Facilitate Communication

A series of analyses were conducted to examine which factors (e.g., questions types, rapport, communication aids) may impact upon the elicitation of IRI in real-world child investigative interviews. As it has been suggested that concurrent vulnerabilities (i.e., ASD) could impact upon children's memory and communication (Almeida, 2018), the following analyses were conducted with both the full sample ($N = 42$) and with the children with a concurrent vulnerability (and their respective match) removed ($n = 32$). Only where the results differed are both analyses reported. Otherwise, the results are given for the full sample.

¹²No predictions were made regarding the total number of questions across the two interview conditions as this is likely to be highly dependent upon a number of situational factors. Although theoretically the use of an RI should result in less questions being asked of the child due to more effective scaffolding strategies, factors such as the police officer's adherence to the RIs recommendations (i.e., if the RI needs to regularly intervene and rephrase questions, the total number of questions will be higher) and the child's level of state management (i.e., the child may be better able to attend or manage their anxiety with the help of the RI meaning that they can be asked more substantive questions) could impact upon this.

Question Types. Pearson's product-moment correlations were run to assess the relationship between the different question types and the amount of IRI elicited from child witnesses. The assumptions of linearity and normality were assessed (Pearson's correlation is considered to be fairly robust to deviations in normality). When any of the assumptions were violated an equivalent non-parametric test was run. There was only one occasion where this had an appreciable effect on the analysis, thus the decision was made to report the results of the Pearson's correlation but alert the reader to the aforementioned discrepancy. There was a statistically significant positive correlation between the number of open-ended questions and IRI, $r = .53$, 95% BCa CI [.23, .77], $p < .005$; and the number of specific-closed questions and IRI, $r = .47$, 95% BCa CI [.13, .76], $p = .002$. There was not a statistically significant correlation between the number of forced-choice questions and IRI, $r = .12$, $p = .445$; multiple questions and IRI, $r = .22$, $p = .163$; leading questions and IRI, $r = .01$, $p = .964$; and yes/no questions and IRI, $r = .29$, $p = .066$ (although this correlation was significant when a Spearman's rank order correlation was run, this does not warrant the increased use of yes/no questions in investigative interviews. In the current study the author did not know the ground truth and therefore had no means of determining whether the information the children provided was accurate. However, yes/no questions utilise recognition memory processes and are therefore more likely to elicit incorrect information compared to open-ended or specific-closed questions. The quality of information children provide is just, if not more, important than the quantity).

Rapport. Spearman's rank order correlations were run to assess the relationship between the rapport measures (i.e., attention and positivity) and the amount of IRI elicited from child witnesses. There was a statistically significant correlation between attention score and IRI, $r_s = .42$, 95% BCa CI [.07, .71], $p = .008$. There was not a statistically significant correlation between expressivity score and IRI, $r_s = .03$, $p = .857$.

Communication Aids. Pearson's point-biserial correlations were run to assess the relationship between the use of different communication aids and the amount of IRI elicited from child witnesses. Outliers in the data were assessed by inspection of boxplots, the distribution was assessed by Shapiro-Wilk test, and homogeneity of variances was assessed by Levene's test for equality of variance. When any of the assumptions were violated, an equivalent non-parametric test was run. However, there were no incidences where this had an appreciable effect on the analysis, thus the decision was made to report the results of the Pearson's point-biserial correlations. There was not a statistically significant correlation between the use of dolls, models, and figures and IRI, $r_{pb}(40) = .16, p = .328$; drawing / writing and IRI, $r_{pb}(40) = .22, p = .169$; body diagrams and IRI, $r_{pb}(40) = .11, p = .491$; ground rules cards and IRI, $r_{pb}(40) = .25, p = .104$; calming tools and IRI, $r_{pb}(40) = .06, p = .722$; and communication aids classified as 'other' and IRI, $r_{pb}(40) = .23, p = .141$. However, when children with concurrent vulnerabilities were removed from the analysis there was a significant correlation for drawing: $r_{pb}(30) = .36, p = .041$. Interviews involving drawing elicited more IRI ($M = 174.18, SD = 82.32$) than those without ($M = 116.90, SD = 66.12$).

Differences in Practices between Registered Intermediary and No Registered Intermediary Interviews

A series of analyses were conducted to examine whether the presence of an RI impacts upon questioning practices, the use of communication aids, rapport, and IRI in real-world child investigative interviews.

Question Types. Independent-samples t-tests were run to determine if there were differences in question composition (i.e., questions asked by all practitioners) between the RI and no RI interviews. Outliers in the data were assessed by inspection of boxplots, the distribution of the dependent variables was assessed by Shapiro-Wilk test, and homogeneity of variances was assessed by Levene's test for equality of variance. When any of the

assumptions were violated, an equivalent non-parametric test was run. However, there were no incidences where this had an appreciable effect on the analysis, thus the decision was made to report the results of the t-test. There was not a statistically significant difference in the number of open-ended questions, $t(40) = -1.48, p = .147$; specific-closed questions, $t(40) = -0.55, p = .589$; forced-choice questions, $t(40) = 0.40, p = .690$; yes/no questions, $t(40) = -0.39, p = .699$; leading questions, $t(40) = -1.04, p = .303$; and in-direct speech acts, $t(40) = 0.08, p = .933$. There was a statistically significant difference in the number of multiple questions between the RI ($M = 6.95, SD = 4.78$) and no RI interviews ($M = 12.86, SD = 8.83$), $M = -5.91, 95\% \text{ CI } [-10.38, -1.43], t(40) = -2.69, p = .011$.

Table 6.4

Means and Standard Deviations of Question Types in Interviews With and Without a Registered Intermediary

Type of Question	Interview Condition	Mean	Standard Deviation	Percentage
Total Questions	RI	98.33	43.80	100.00%
	No RI	113.05	51.93	100.00%
	Total	105.69	48.03	100.00%
Open-ended Questions	RI	3.38	2.89	3.44%
	No RI	5.38	5.48	4.76%
	Total	4.38	4.45	4.14%
Specific-closed Questions	RI	39.95	19.41	40.63%
	No RI	43.90	27.01	38.83%
	Total	41.93	23.31	39.67%
Yes/no Questions	RI	40.38	21.12	41.07%
	No RI	42.76	18.39	37.82%
	Total	41.57	19.60	39.33%
Forced-Choice Questions	RI	5.05	4.66	5.14%
	No RI	4.52	3.74	4.00%
	Total	4.79	4.18	4.53%
Multiple Questions*	RI	6.95	4.78	7.07%
	No RI	12.86	8.83	11.38%
	Total	9.90	7.63	9.37%
Leading Questions	RI	2.62	2.99	2.66%
	No RI	3.62	3.22	3.20%
	Total	3.12	3.11	2.95%
Indirect Speech Acts	RI	9.05	5.85	9.20%
	No RI	8.90	5.11	7.87%
	Total	8.98	5.43	8.50%

Note. Total percentages do not total 100% as indirect speech acts are also included within yes/no questions.

* $p < .05$

Rapport. Mann-Whitney U tests were run to determine if there were differences in rapport between the RI and no RI interviews. The distribution of both the attention and expressivity scores for the RI and no RI interviews were similar as assessed by visual inspection. The median attention score was not statistically significantly different between the RI and no RI interviews, $U = 164.00$, $Z = -.74$, $p = .460$. The median expressivity score was not statistically significantly different between the RI and no RI interviews, $U = 185.50$, $Z = -.13$, $p = .899$.

Table 6.5

Median and Range of Rapport Scores in Interviews With and Without a Registered Intermediary

Rapport Measure	Interview Condition	Median	Range
Attention	RI	23	11 - 27
	No RI	22	7 - 24
Positivity	RI	12	6 - 21
	No RI	12	4 - 22

Communication Aids. Chi square tests were conducted to examine whether there were any significant associations between the use of different communication aids and the interview condition. Fisher's Exact Test is reported where the assumption related to cell counts is violated. There was a statistically significant association between the use of dolls, models, and figures and the interview condition as assessed by Fisher's Exact Test, $p = .009$. The association was of moderate strength (Cohen, 1988), $\Phi = .447$. Dolls were used in 33.3% of the RI interviews and none of the no RI interviews. There was no statistically significant association between the use of drawing / writing and the interview condition: $\chi^2(1, N = 42) = 0.10, p = .747$. The association between the use of body diagrams and the interview condition was approaching significance as assessed by Fisher's Exact Test, $p = .093$. There was no statistically significant association between the use of ground rules cards and the interview condition as assessed by Fisher's Exact Test, $p = .107$. There was a statistically significant association between the use of calming tools and the interview condition: $\chi^2(1, N = 42) = 4.71, p = .030$. The association was of moderate strength (Cohen, 1988), $\phi = .335$. Calming tools were used in 61.9% of the RI interviews compared to 28.6% of the no RI interviews. However, when children with concurrent vulnerabilities were removed from the analysis this was no longer significant: $\chi^2(1, N = 32) = 2.00, p = .157$ (calming tools were used in 62.5% of the RI interviews and 37.5% of the no RI interviews). There was a statistically significant association between the use of communication aids classified as 'other' and the interview condition as assessed by Fisher's Exact Test, $p = .001$. The association was strong (Cohen, 1988), $\phi = .522$. Tools classified as 'other' were used in 42.9% of the RI interviews and none of the no RI interviews (see table 6.6).

Further analyses were conducted to examine whether there were any significant associations between the purpose for which communication aids were used and the interview condition. All analyses violated the assumption related to cell counts, thus the results of

Fisher's Exact Test are reported. There was no statistically significant association between the use of communication aids for the purpose of information gathering and the interview condition as assessed by Fisher's Exact Test, $p = .115$; for the purpose of clarification and the interview condition as assessed by Fisher's Exact Test, $p = .419$; for the purpose of state management and the interview condition as assessed by Fisher's Exact Test, $p = .095$; and for 'other' purposes and the interview condition as assessed by Fisher's Exact Test, $p = .619$ (see table 6.7).

Table 6.6

Number of Registered Intermediary and No Registered Intermediary Interviews Using Each Communication Aid

Type of Communication Aid	Interview Condition	Used	Not Used
Dolls*	RI	7 (33.3%)	14 (66.7%)
	No RI	0 (0.0%)	21 (100.0%)
Drawing	RI	8 (38.1%)	13 (61.9%)
	No RI	7 (33.3%)	14 (66.7%)
Body Diagrams	RI	6 (28.6%)	15 (71.4%)
	No RI	1 (4.8%)	20 (95.2%)
Ground rules cards	RI	4 (19.0%)	17 (81.0%)
	No RI	0 (0.0%)	21 (100.0%)
Calming objects*	RI	13 (61.9%)	8 (38.1%)
	No RI	6 (28.6%)	15 (71.4%)
Other*	RI	9 (42.9%)	12 (57.1%)
	No RI	0 (0.0%)	21 (100.0%)

* $p < .05$

Table 6.7

The Purpose for which Communication Aids were Used in the Registered Intermediary and No Registered Intermediary Interviews

Purpose	Interview Condition	Employed for this Use	Not Employed for this Use
Information Gathering	RI	9 (52.9%)	8 (47.1%)
	No RI	2 (18.2%)	9 (81.8%)
Clarification	RI	11 (64.7%)	6 (35.3%)
	No RI	9 (81.8%)	2 (18.2%)
State Management	RI	14 (82.4%)	3 (17.6%)
	No RI	5 (45.5%)	6(54.5%)
Other	RI	4 (23.5%)	13 (76.5%)
	No RI	1 (9.1%)	10 (90.9%)

Note. Analysis only includes interviews in which a communication aids were used resulting in a small sample size. As such the results need to be interpreted with caution.

IRI. An independent-samples t-test was run to determine if there were differences in the amount of IRI elicited between the RI and no RI interviews. There were multiple outliers in the data as assessed by inspection of a boxplot. IRI for the RI interviews was normally distributed as assessed by Shapiro-Wilk test ($p > .05$), but the IRI for the no RI interviews violated this assumption ($p = .014$). Running an equivalent non-parametric test had no appreciable effect on the analysis thus the decision was made to report the results of the t-test. There was homogeneity of variances as assessed by Levene's test for equality of variances ($p = .599$). There was not a statistically significant difference in IRI between the RI ($M = 125.81$, $SD = 71.48$) and no RI interviews ($M = 142.38$, $SD = 71.00$), $t(40) = -0.75$, $p = .455$.

Descriptive Analyses of the RI Reports and Interview Interventions

The following analyses focus on the content of the RIs' reports and the number / purpose of the RIs interventions during the interview.

RI Reports. All of the RI reports ($N = 7$) assessed the children's attention, anxiety, and behaviour; expressive communication; and receptive communication. The use of standardised assessment tasks to assess the children's communication was described in none of the RI reports. Three of the reports described using non-standardised assessment tasks and four of the reports did not provide this information. All of the reports made recommendations regarding the structure of questions and language. Six of the reports made recommendations regarding the use of communication aids, procedural factors, and how to phrase questions related to time, distance, or frequency. Five made recommendations regarding the frequency of breaks, pace of questioning and / or signposting of topics; and four made recommendations regarding state management.

RI Interventions. The RIs intervened in 17 of the 21 interviews, with the total number of interventions ranging from 0 - 23 ($M = 7.29$, $SD = 7.11$). The RIs most frequently made interventions pertaining to questioning including rephrasing complex questions, advising on lines of questioning (i.e., topics), or directly asking the child a question ($M = 2.52$, $SD = 3.22$), followed by communication aids ($M = 1.62$, $SD = 2.13$), understanding ($M = 1.57$, $SD = 2.29$), procedural / state management ($M = 1.19$, $SD = 1.60$), and breaks ($M = 0.33$, $SD = 0.66$). Although, the number of interventions were comparable when the children with a concurrent vulnerability were removed ($M = 7.13$, $SD = 6.53$), interventions related to procedural / state management ($M = 1.44$, $SD = 1.75$) were the second most prevalent behind interventions related to questioning ($M = 2.75$, $SD = 3.36$). All other interventions remained in their respective positions. A Spearman's rank order correlation was run (as the assumptions of linearity and normality were violated, and outliers were also identified) to assess the

relationship between the number of RI interventions and the amount of IRI elicited from child witnesses. There was not a statistically significant correlation between the number of RI interventions and IRI, $r_s = -.086$, $p = .712$.

The RIs asked questions in 13 of the 21 interviews, with the number of questions ranging from 0 - 43 ($M = 6.67$, $SD = 11.08$). The RIs most frequently asked specific-closed questions ($M = 2.71$, $SD = 5.59$), followed by yes/no questions ($M = 2.29$, $SD = 3.96$), forced-choice ($M = 0.90$, $SD = 1.64$), multiple ($M = 0.43$, $SD = 0.75$), leading ($M = 0.24$, $SD = 0.89$), and open-ended questions ($M = 0.10$, $SD = 0.44$).

6.3 Discussion

The current study is the first to examine the impact of the RI provision on real-world investigative interviews. The first hypothesis was that there would be positive correlations between both open-ended questions and IRI, and specific-closed questions and IRI. This hypothesis was fully supported. Despite this, the use of open-ended questions was found to be low across both the RI and no RI interviews. It was also hypothesised that interviews in which communications aids were used would elicit more IRI from the children than those without aids, and that RI interviews would be more likely to use communication aids compared to those without an RI. The RI interviews were more likely to use dolls, calming tools, and communication aids classified as 'other'. However, there was only a relationship between the use of communication aids (i.e., drawing) and IRI when the children with a concurrent vulnerability were removed from the analysis. It was further hypothesised that there would be a positive correlation between both non-verbal rapport measures (i.e., attention and positivity) and IRI, and that higher levels of rapport would be observed in the interviews with an RI as opposed to those without. A positive relationship was found between attention and IRI, however levels of rapport did not differ across the RI and no RI interviews. Finally, it was hypothesised that the children would provide significantly more IRI in the interviews with, as

opposed to those without an RI. This hypothesis was not supported. Potential explanations for this and the other findings are discussed below.

Question Types

In the current study a positive correlation was found between open-ended question usage and IRI. This is in line with the findings of previous research (e.g., Gagnon & Cyr, 2017; Luther et al., 2015; Sternberg, Lamb, Davies, & Westcott., 2001). For example, Sternberg, Lamb, Davies, and Westcott (2001) found that on average children provide more details in response to open-ended questions than to the other question types (i.e., specific-closed, forced-choice, and leading) combined. The current study also found a positive correlation between the use of specific-closed questions and IRI. These findings add further credence to the recommendations, regarding question types, outlined in the ABE guidance (MoJ, 2011). The guidance recommends using open-ended questions predominantly during interviews with children, with specific-closed questions used as a means of eliciting further details.

Despite best practice guidance (ABE; MoJ, 2011) advocating the use of open-ended questions, the use of open-ended questions in the current study was low across both the RI and no RI interviews. Open-ended questions constituted approximately 4% of the total interviewer utterances. Although this finding is in line with the findings of previous research (e.g., Cederborg et al., 2000; Luther et al., 2015; Wolfman et al., 2016), it is in direct violation of the ABE guidance (MoJ, 2011). In order to address this continued deviation from best practice, one needs to determine what is responsible for it. It is likely that many of the reasons that were described in section 2.2.5 for failing to comply with the guidance, including the unusual nature of maintaining an open-ended discourse, can explain why the presence of an RI did not lead to an increase in the usage of open-ended questions. Two of the reasons discussed in section 2.2.5 appear to be particularly salient to the current study.

The first explanation is that there is a lack of awareness, amongst both police officers and RIs, as to what constitutes an open-ended question. Previous research has shown that the ability to identify open-ended questions is an important indicator of their use within an interview context (Yii et al., 2014). The findings of study one, of this thesis, would suggest that RIs have difficulty identifying open-ended questions. Of the RIs, 65.6% who participated in study one, stated that ‘Can you tell me what happened?’ is an example of an open-ended question. However, this is actually an example of an indirect speech act (see section 2.2.2 of this thesis). Indirect speech acts directly ask if the child knows, whilst indirectly asking what they know (Evans et al., 2014). As such some children will respond to these questions with a simple ‘yes’ or ‘no’ to indicate that they either do or do not know the answer - they will not recognise the indirect question. This was observed, particularly with some of the younger children, in the interviews reviewed as part of this study. However, this particular error can be easily rectified by both police officers and RIs being more conscious of not pre-facing open-ended questions (and other appropriate questions i.e., specific closed) with phrases such as ‘can you remember-’ or ‘do you know-’. A similar error, which was observed in the current study, was open-ended questions being followed with a yes/no question regarding whether the child remembers the information or knows the answer (e.g., ‘Tell me what happened? Can you remember?’). This results in the questions being coded as multiple as opposed to open-ended. An interesting avenue for future research would be to examine the effect of open-ended questions versus in-direct speech acts and the aforementioned multiple questions (e.g., ‘Tell me what happened? Do you remember?’) on children’s responses in real-world investigative interviews.

A second explanation is that, both police officers and RIs, fail to recognise the potential of open-ended questions. Of the RIs that participated in study one of this thesis, 30% were not aware that open-ended questions should be used predominantly during an

investigative interview, due to eliciting the most accurate accounts from child witnesses. This lack of awareness regarding the benefits of open-ended questions could account for the findings of Krähenbühl's (2011) study, in which many of the alternatives provided by RIs were not open-ended - a finding replicated in the current study. Only 1.5 % of the questions asked by the RIs in the current study were open-ended.

Previous research with police officers has found that open-ended questions are perceived, by many, as being less effective than specific-closed questions at eliciting specific details from children (Guadagno, et al., 2013; Powell et al., 2010; Wright & Powell, 2006). It is likely that this belief held by police officers, is perpetuated to some extent by the RIs. Inspection of the RI reports, revealed a number of incidents, whereby the RIs stated that the children would provide only limited detail in response to open-ended questions and would require additional, more specific questions to facilitate their account. Although this may be the case for some children, training needs to emphasise the variability in children's recall and that for many children open-ended questions are appropriate and will elicit the most detailed responses.

Training police officers and RIs in the effective use of open-ended questions should be considered a priority given the difficulties both appear to have in maintaining an open-ended discourse. This difficulty is likely perpetuated by the infrequent use of such questions in everyday conversation. As such, the training should cover the different types of open-ended questions and the importance of pairing (e.g., following a specific-closed question with an open-ended question). Interviewers can, for example, use the information elicited from a specific-closed question (e.g., 'where did you go?' 'the park') to structure and form the basis of either an open-ended breadth (e.g., 'you went to the park, then what happened?') or open-ended depth question (e.g., 'Tell me what happened at the park?'). These questions are less broad than open-ended invitations and have been found, in previous studies, to elicit more

detailed responses (e.g., Gagnon & Cyr, 2017). This is attributed to open-ended breadth and depth questions containing pre-disclosed details which can scaffold children's recall by focusing their attention (Orbach & Lamb, 2000).

A more promising characteristic of the interviews, reviewed as part of the current study, was the low number of leading questions. Leading questions constituted approximately 3% of the total interviewer utterances. Although this is slightly higher than the figure reported in Wolfman et al. (e.g., 0.5%; 2016), it is considerably lower than has been observed in other research (e.g., 17.8%; Verkampt et al., 2021). This would suggest that the police officers who conducted the interviews, analysed as part of the current study, do have some appreciation of the dangers of leading questions.

However, of the leading questions asked in the current study many were tag questions. It was noted in chapter two of this thesis that tag questions are considered to be the most suggestive form of leading question (ABE; MoJ, 2011). In the current study, tag questions often asked about topics that were unlikely to be in dispute (e.g., 'I came to visit you last week, didn't I?') or were used as a means to recap or clarify what the child had said (e.g., 'You said, didn't you?'). Although these questions are perhaps not the most damaging from an evidential perspective, they set a very poor precedent for the rest of the interview. The ABE guidance describes the interview as a "learning experience" (MoJ, 2011, p. 78), whereby the child will learn from the interviewer's behaviour what is expected of them and then act accordingly. Tag questions do not allow the child to control the flow of information, they also undermine the child's role as the expert, and can lead to the child becoming a passive participant in the interview process. As the child is expected to do the majority of the talking in an investigative interview (Malloy et al., 2015), anything that prohibits or restricts this could have a detrimental effect on the outcome of the interview and the case as a whole.

Therefore, tag questions should be avoided, throughout the entirety of the interview, even when discussing non-substantive topics.

The only significant difference that emerged in relation to question types across the two interview conditions, in the current study, was in the number of multiple questions. The RI interviews contained significantly fewer multiple questions than the interviews without an RI. There were incidences whereby the RIs advised interviewers to breakdown multiple questions into separate parts. Recommendations were also made in the RI reports to ask short manageable questions. Both of the above likely contributed to the interviewers using fewer multiple questions in the interviews with an RI. Given that multiple questions have been found to compromise the accuracy (e.g., Carter et al., 1996) and completeness of children's account (e.g., Katz & Hershkowitz, 2012), this can be considered a positive outcome of the RIs' presence.

Previous research has demonstrated that RIs have an awareness of the dangers of using multiple questions with child witnesses (e.g., Hanna & Henderson, 2018). A recent study by Hanna and Henderson (2018) found RIs to be more likely, than other criminal justice practitioners (i.e., lawyers), to identify subtle multiple questions. The RIs were also, in Hanna and Henderson's (2018) study, more successful at rephrasing and simplifying these questions to make them more developmentally appropriate. Furthermore, the RIs were reported as being more astute, than the lawyers, at identifying questions containing legal vocabulary, unclear references, the passive voice, and negation.

Krähenbühl (2011) similarly found RIs to have good awareness of developmental and emotional factors. However, it also emerged in Krähenbühl's study that RIs possess a limited knowledge of the best practice guidance. It may be that RIs are skilled at determining whether a child understands a question but, due to their limited knowledge of the guidance, lack awareness as to which questions elicit the best quality information. Multiple questions pose a

challenge to understanding whilst the other question types (e.g., open-ended, specific-closed, forced-choice) if phrased appropriately do not. This may account for why no significant differences were observed between the interview condition and the other question types.

As the current study only examined the impact of the RI provision on question types (i.e., open-ended, specific-closed) and not the structure or complexity of questions, it may be that the study has underestimated the role of the RI in improving the quality of questioning in investigative interviews with children. Comparing interviews with and without an RI in respect to the length and complexity of questions, along with vocabulary use should be considered as an important avenue for future research.

Communication Aids

The current study found no relationship, in the full sample, between the use of communication aids and the amount of IRI elicited from children during ABE interviews. This can be attributed to the aids, particularly in the interviews without an RI, being used more frequently for the purposes of clarification (e.g., establishing the child's understanding of sexual body parts) and state management (e.g., reducing the child's anxiety). Only two of the interviews without an RI used communication aids for the purpose of gathering new information. This may be due to police officers lacking confidence in using communication aids for this purpose or alternatively it may be that they are aware of the potential risks associated with this practice. Using communication aids to enable a child to give an account is deemed, in the ABE guidance, as being more controversial than for the purposes of clarification and state management (MoJ, 2011), and thus may be perceived by police officers as requiring a higher level of expertise. That being said, when the children with a concurrent vulnerability were removed from the analysis a significant relationship was found between the use of drawing / writing and the amount of IRI elicited from the children. Although this finding is in line with previous research which had demonstrated the positive effects of

drawing (e.g., Butler et al., 1995; Woodford et al., 2015), it is unclear as to why this relationship only emerged as significant when children with a concurrent vulnerability were removed.

Previous research (e.g., Mattison et al., 2015) has examined the effect of sketch reinstatement of context on the recall of children with ASD. Although this technique is not representative of how drawing was used in the interviews included in the current study (e.g., the timing and the instructions which accompanied drawing) it is important to acknowledge the beneficial effects reported by Mattison et al. (2015). Considerable mean differences were found, between the three retrieval conditions (i.e., sketch reinstatement of context, mental reinstatement of context, and control) for the amount of correct, erroneous, and confabulated information. These differences accumulated leading to a significant improvement in recall accuracy in the sketch reinstatement condition. Despite these positive findings further research is needed to examine how communication aids, including drawing, can be used most effectively to support recall in children with concurrent vulnerabilities.

Drawing was also shown to be the most frequently used communication aid in the interviews without an RI, with no difference in use across the two interview conditions (i.e., RI and no RI). This is in line with the findings of previous research (e.g., Mattison & Dando, 2020; Wolfman et al., 2018). Wolfman et al. (2018) in their analysis of 98 interviews, relating to CSA, found sketch plans to be the most commonly used tool by interviewers. Similarly, drawing was found to be a popular tool amongst police officers in a recent study by Mattison and Dando (2020). In Mattison and Dando's (2020) study 54% of police officers and 57% of RIs reported the frequent use of drawing in interviews with vulnerable witnesses. The benefits of drawing have been highlighted both in previous research (see section 2.2.4 of this thesis) and in best practice guidance (ABE; MoJ, 2011). ABE, for example, states that "drawing has significant benefits" and that "the symbolic nature of pictures and drawings is more easily

understood by young children than dolls and models” (MoJ, 2011, p. 91). This has likely led to practitioners, particularly police officers, favouring drawing over other communication aids. This is a positive finding as drawing has been found to facilitate children’s verbal communication without compromising accuracy (e.g., Barlow et al., 2011; Butler et al., 1995; Gross & Hayne, 1998). Research with dolls and body diagrams has, in contrast, elicited far more inconsistent research with some studies reporting a large number of commission errors (e.g., Bruck et al., 2000; Bruck et al., 2016) – these can have very serious repercussions in real-world investigative interviews.

Although in the current study, the use of drawing was comparable between the interviews with and without an RI, a greater disparity emerged in the use of other tools. The interviews with an RI were more likely to involve the use of dolls, models, and figures (not anatomical dolls); calming tools; and communication aids classified as ‘other’ (e.g., visual timelines and emotions cards). Given the amount of controversy which surrounds the use of dolls (Salmon et al., 2012), these tools may be perceived by police officers as more specialist aids to communication. Therefore, police officers may consider the use of these tools as being outside of their remit or area of expertise. That being said, dolls, models, and figures were used in a number of the RI interviews. It would not be surprising if RIs are perceived, by police officers, as having an elevated level of expertise within this area. The Advocates Gateway Toolkit, related to communication aids, recommends seeking the advice of an RI if it is believed that the use of aids would be beneficial to a witness’ communication (Mattison, 2015). This greater expertise and knowledge of communication aids could also account for the RI interviews, in the current study, involving more aids classified as ‘other’ and calming tools. A greater knowledge and awareness of communication aids would mean that RIs have a larger repertoire of tools at their disposal (i.e., visual timelines and emotions cards) and have a

greater appreciation of the diverse purposes with which the tools can be used (i.e., state management).

Rapport

The current study found a positive relationship between attention and the amount of IRI elicited from children during ABE interviews. This finding is in line with the belief that good rapport has the potential to improve communication (Saywitz et al., 2015). Although the current study found a positive relationship between attention and IRI, no similar relationship was found for positivity. This is in contrast to a previous study by Collins (2012). Collins' (2012) study explored the communicative effects of using play to build rapport. The study found collaborative play (i.e., with the researcher) to have communication benefits. Children who engaged in a collaborative play task were more communicative and expressive (a measure of positivity) in a subsequent interaction, than children who completed the task alone. As such positivity was found to have a positive relationship with productivity. However, Collins' (2012) study only involved a 3- to 5-minute discussion of neutral topics and thus cannot be considered representative of an investigative interview context. Nevertheless, failure to find a significant relationship in the current study could be attributed to how positivity was measured and is discussed in more detail below.

In the current study, no difference was found in levels of rapport (i.e., attention and positivity) between the RI and no RI interviews. This finding is unexpected given the additional time that will have likely been spent with the children, in the RI interviews, as a result of the communication assessment (it is rare that police officers will conduct these assessments; see study two of this thesis). The Registered Intermediary Procedural Guidance Manual (2020a) states that the assessment is an opportunity for the RI to build rapport with the child. However, any rapport that may have been built, between the RI and the child, during the assessment may not be as apparent at interview due to the less central role of the

RI. Rather than the problem being that the assessment is not successful in building rapport with a child, it may be that the problem lies in the level of involvement from the police officer. None of the reports, analysed as part of the current study, described the degree of the police officer's involvement in the assessment process. Perhaps rapport would better translate from the assessment to the interview if a more collaborative approach to the assessment was taken. For example, the police officer could conduct a practice narrative (see section 2.2.1) with the child. This may not only serve to increase rapport between the interviewer and child, it would also give the interviewer an opportunity to practise asking open-ended questions which could in turn improve their practice (Brubacher et al., 2011).

Alternatively, the lack of significant effect, pertaining to rapport, could be due to how the construct was measured in the current study. Fifteen second extracts were taken from the beginning, middle, and end of each interview and the children rated for two non-verbal indicators of rapport (i.e., attention and positivity). First, this approach was problematic in that some of the children's faces were not visible on camera, during large portions of the interview, due to the children being engaged with another task (e.g., drawing). This made it particularly difficult to give the children an accurate positivity rating. Arguably, for these ratings to be more accurate one would need to be able to see the child's face but also their drawing, as drawing is in itself a means of expression (Alford, 2015). In addition, the non-verbal indicators (i.e., attention and positivity; Tickle-Degnen & Rosenthal, 1990) and methods used to measure rapport in the current study were based on early research involving adult – adult dyads in benign situations (e.g., Bernieri et al., 1996). It maybe that the indicators of rapport are different for adult – child dyads in an investigative interview context - rapport is considered by Tickle-Degnen and Rosenthal (1990) to be a construct which is highly context dependent. Furthermore, the study did not measure the non-verbal rapport of the police officers and RIs (as they were not clearly visible on camera). Yet, rapport is considered

to be a construct which only possesses meaning as part of dyad or a group (Tickle-Degnan & Rosenthal, 1990) and as such is affected by one's interactional partner (Johnston et al., 2019). Some police officers may be more adept at developing rapport than others and failing to control for this could potentially have hidden any significant effect of the interview condition (as could failing to control for the child's reluctance; Johnston et al., 2019).

Investigation Relevant Information

In the current study, no significant difference was found in the amount of IRI between the interviews with and without an RI. This is despite the RIs intervening on average seven times per interview. The Registered Intermediary Procedural Guidance Manual (MoJ, 2020a) states that the role of the RI is to ensure that communication is as complete, coherent, and accurate as possible. It may be that the RI provision has a greater impact upon the accuracy and coherence, as opposed to the completeness of children's accounts in real-world interviews. Theoretically, many of the RIs' interventions (e.g., requesting a complex question be rephrased, asking the officer to clarify the meaning of their question, suggesting the use of a communication aid to check the location of a touch) could have impacted upon the accuracy of the child's account but not necessarily the detail. It is not unusual for children to answer questions that they do not understand (Waterman et al., 2000) or do not know the answer to (Rohwer et al., 2012). The RI interviews, in the current study, were also found to contain significantly fewer multiple questions than the interviews without an RI. The ABE guidance (MoJ, 2011) cautions that the use of multiple questions can lead to misunderstandings between the interviewer and child. These misunderstandings could potentially have an adverse effect on the coherence and accuracy of children's accounts, thus supporting the above assertion regarding the relative impact of the RI provision. Unfortunately, accuracy cannot be measured in real-world interviews as the researchers do not know the ground truth. However, future research could use Stein and Glenn's (1979) story-grammar framework to

look at how the RI provision impacts upon the coherence as well as the detail of children's narrative accounts in real-world investigative interviews (see section 7.5 of this thesis).

Alternatively, it could be that the benefits of the RI provision differ depending on the cognitive profile of the child. Study three found that different cognitive abilities appear to underpin different aspects of children's recall (i.e., correct, incorrect, confabulated, and accuracy of IRI). As such, for a child with poor attentional abilities, the presence of an RI may impact more upon the accuracy as opposed to the detail of their account (a significant relationship was found, in study three, between attentional abilities and the accuracy of IRI). As the cognitive profiles of the children, in the current study, were not examined it is also possible that the cognitive profiles of the children in the RI and no RI interviews were not comparable. This could have potentially hidden any positive effects of the RI provision.

Furthermore, the composition of the sample in the current study could account for there being no significant impact of the RI provision on elicitation of IRI. The sample comprised predominantly of TD children, with very few of the children under the age of 6 ($n = 3$). Had the sample included a greater number of young children, children with a concurrent vulnerability, or children who were highly traumatised a significant effect of the RI provision may have emerged¹³. More specialist skills are required when interviewing these groups of

¹³ To be able to evaluate these differences, the RI and no RI interviews would have to be evenly matched.

However, this presents a number of challenges. From discussions with police officers, it is now rare that very young children would be interviewed without an RI. Although anecdotal, it is also unlikely that children with a formal diagnosis (e.g., ASD) would be interviewed without an RI. Thus, in order to match the interviews with an RI to those without, older interviews would have to be used for the latter (e.g., either before or immediately preceding the introduction of 'special measures'). This creates its own difficulties as both research and practice has evolved since 2008. Alternatively, interviews could be used from a country which does not use RIs (e.g., Scotland). However, this is also problematic due to differences in practice and the use of different interview guidelines.

children (Marchant, 2013). Police officers have, for example, expressed a lack of confidence in interviewing very young child. This reluctance is captured in the following quote from a senior police officer: “if the child is under five, run a mile” (Marchant, 2013).

Given that RIs are considered to be specialists in communication (Plotnikoff & Woolfson, 2007) it is in these more complex cases (i.e., involving very young children or children with more pronounced communication difficulties) that the true potential of RIs may be fully realised. Importantly RIs only accept cases within their area of expertise (MoJ, 2020a). Police officers, on the other hand, have to be far more omniscient. It is therefore unlikely that police officers fully understand the nuances of all the vulnerabilities in which they come across, and thus where the experience and expertise of the RI maybe particularly beneficial. It is also important to recognise that the majority of the interviews included in the current study were conducted in the last three years ($n = 35$). As such, a lot of the interviewers, in the no RI interviews, may have still had prior experience of working with RIs. Through these experiences the police officers may have learnt strategies to scaffold communication from the RIs which they have then applied to their own practice (Plotnikoff & Woolfson, 2015). Thus, resulting in no significant difference between the amount of IRI elicited in the interviews with and without an RI.

RI Reports and Interventions

Given the paucity of research into the assessment process of RIs and more specifically the content of their reports, a brief discussion of this (although not the primary focus) was thought to be a pertinent addition to the current study. All of the RI reports, reviewed as part of the current study, addressed the three areas of communication which appeared to underpin the recommendations and advice given in the Registered Intermediary Procedural Guidance Manual (2020a): attention, expressive language, and receptive language. Although the reports rarely went into detail regarding the methods used to assess these areas of communication, the

reports did describe what they assessed within each of these areas. This included the child's attentional abilities; competency responding to different question types; ability to provide a free narrative account; understanding and use of language (including idioms); and accuracy in estimating time, frequency, and distance.

In addition to the factors described above, a number of the reports stated that the child's knowledge of sexual body parts was assessed (e.g., breasts, penis, vagina) prior to the interview. This is concerning as it could be interpreted as priming the child to discuss these topics. That being said, the researchers do not know the context or rationale behind the RIs' decisions to explore the children's knowledge of sexual body parts. Given the potential repercussions of this though, such decisions warrant further investigation. Future research, using a think aloud task (Wright & Powell, 2006), could prove valuable in understanding more about the decision-making process of RIs during their assessments of child witnesses.

The assessment is an important part of the RI role as it determines the recommendations for and potentially the outcome of the interview. Of the RI reports reviewed, recommendations generally related to seven aspects of communication: the use of communication aids; the type and structure of questions; language use; the frequency of breaks, pace of questioning, and signposting of topics; procedural factors; state management; and how to phrase questions related to time, distance, and frequency. All of these recommendations reflect the topics discussed in section 3.17 of the Registered Intermediary Procedural Guidance Manual (MoJ, 2020a), which outlines what should be considered when planning an ABE interview.

The reasons for the RI interventions, in the current study, largely related to the same overarching topics as their recommendations. However, the degree of the RI's involvement varied considerably across the interviews analysed as part of the current study. Some of the RIs did not intervene and took a very passive role, whilst others intervened on multiple

occasions and were far more active participants in the interview process. The degree of the RI's involvement is likely due, in part, to the success of the assessment and pre-interview planning. If this is done well and the interviewer adheres to the RI's recommendations, there should be little reason for the RI to intervene during the interview itself. However, that does not mean that multiple interventions from the RI are indicative of a poor assessment or an unreceptive interviewer. Instead, these could, for example, be related to the unfamiliar context of the interview, the child's anxiety and / or reluctance to make a disclosure, and difficulty putting their experiences (which depending upon the nature of the offence the child may not fully understand) into words.

In the current study, RIs made the most interventions in relation to questioning. The rephrasing of an interviewer's question, by an RI, is considered to be an entirely appropriate behaviour (MoJ, 2020a). However, some of the RIs' behaviours, which were encompassed within the questioning category, were arguably outside the remit of the role and could be interpreted as a violation of their Code of Practice (MoJ, 2020a). A number of the RIs, in the current study, advised the police officers on lines of questioning. There were also incidents of the RIs asking questions of the child unrelated to the interviewers' previous question (i.e., not rephrasing the question or in an attempt to resolve a miscommunication). This behaviour is concerning as it could be construed as the RIs aligning themselves with the police, thus compromising their perceived impartiality. Similar concerning behaviours (i.e., discussing inconsistencies in a witness' account) were reported in a recent study by Taggart et al. (2021) and indicate that further training is required in respect to the boundaries of the RI role.

The style, or how, the RIs intervened also varied. Some of the RIs would direct their interventions towards the interviewer (e.g., the RI would suggest to the interviewer that the question was too long and needed to be broken down into smaller parts), whereas others would bypass the interviewer and simply ask the child the rephrased question (e.g., the

question broken into smaller parts). The Registered Intermediary Procedural Guidance Manual (MoJ, 2020a) does not provide specific guidance as to how an RI should intervene but instead states that this should be discussed with the interviewing officer. It is therefore not surprising that different approaches have emerged. An interesting avenue for future research would be to examine the efficacy of RI interventions and which of the aforementioned approaches results in the best interview outcomes.

Limitations and Areas for Future Research

The first limitation of the current study is the modest sample size. The modest sample size precluded any statistical analysis of the RI reports as these were not available in all cases. In order to gain more of an insight into the RIs' assessments, recommendations, and the overall content of their reports, future research should aim to analyse a larger sample of interviews, from a greater number of police forces. A second limitation of the current study relates to how positivity was measured (i.e., expressivity score). Future research should try to develop a framework which takes into account other ways in which a child can express themselves (e.g., drawing; Alford, 2015). Due to the dyadic nature of rapport, an attempt should also be made to quantify the rapport of the police officers and RIs. The verbal indicators developed by Collins and Carthy (2019) may be a viable option as these would negate the issue of the police officer and RI not being clearly visible on the ABEs. A further limitation of the current study is that much of the analyses was correlational in nature. As such causation between the methods used to scaffold communication (i.e., question types, rapport, communication aids) and the elicitation of IRI cannot be established. It is therefore recommended that future research adopts an approach, such as behavioural sequence analysis (Keatley et al., 2017), which can determine whether a behaviour is directly influenced by the preceding behaviour. This would be particularly useful in exploring the efficacy and impact of the RI interventions.

An additional and important limitation of the current study relates to the matching of the interviews. First, whilst none of the interviews without an RI involved children with concurrent vulnerabilities, five of the interviews with an RI did. The reason for this disparity is discussed above, along with how this was controlled for in the analyses. Second, nine of the interviews without an RI had a social worker present; a social worker was not present in any of the interviews with an RI. As with the RIs, the involvement of the social workers varied considerably. Some of the social workers did not ask any questions (even when asked if they wished to add anything by the police officer), whilst others asked questions throughout the interviews. Thus, some of the social workers did contribute to scaffolding the children's communication – an area which needs to be explored in future research. Despite the current study seeking to eliminate differences between the RI and no RI interviews, it was not possible to control for this particular variable as the researcher had far fewer interviews to choose from without an RI (which is arguably positive as it demonstrates that RIs are being used frequently by the two forces included in the study). That being said, the fact that some of the interviews included a social worker is not considered a major concern as the role of the social worker has more parallels with the police officer as opposed to the RI. Finally, although there was no significant difference between the RI and no RI interviews in relation to whether the abuse was repeated vs. single or intra- vs. extra-familial, both were approaching significance. Both of these factors can impact upon a child's reluctance to disclose abuse. Previous research has found that children who experience repeated abuse or abuse by a family are the most reticent (e.g., Goodman-Brown et al., 2003; Herskowitz, Lanes, & Lamb, 2007; London et al., 2005). As such, future research should look to match these characteristics more closely than was achieved in the current study.

Conclusions

The current study is the first to examine the impact of the RI provision on practice in real-world investigative interviews. The findings provide further support for the use of open-ended questions and the importance of rapport, in eliciting detailed accounts from child witnesses. That being said, open-ended questions constituted less than 5% of the total interviewer utterances, with the number of open-ended questions comparable across the RI and no RI interviews. A more positive feature of the RI interviews was that significantly fewer multiple questions were asked compared to the interviews without an RI. The RI and no RI interviews also differed in relation to the use of communication aids. The interviews with an RI were more likely to include the use of dolls, models, and figures; calming tools; and communication aids classified, in the current study, as 'other'. As such, the RI provision does appear to have some influence on interview practice. However, future research is required to determine the impact of the RI provision on the complexity of questions (i.e., length, structure, language) in real-world investigative interviews, and to determine the efficacy of the RIs' interventions.

Chapter Seven: General Discussion

The overarching aim of this thesis was to expand upon the limited body of research regarding the RI provision and pre-interview communication assessments. This was achieved by conducting four novel research studies related to different aspects of the RI role namely the RIs' knowledge of children's memory, the efficacy of pre-interview communication assessments, the influence of individual differences on children's recall, and the impact of the RI provision on investigative interview practice. The general discussion begins by summarising the main findings of each of these studies. This is followed by a discussion of the theoretical, practical, and methodological implications of the research. Finally, the concluding section provides a holistic appraisal of the thesis and its findings.

7.1 Summary of Main Findings

Although the body of research regarding the work of RIs is expanding, there are still many facets of the role with which researchers, practitioners, and policy makers have very little knowledge. The current thesis sought to address some of these gaps in knowledge with the hope of enabling professionals to make more informed decisions regarding the direction of the WIS, along with areas for future training and CPD opportunities. A summary of the four studies conducted as part of this thesis is presented below.

Study one (chapter three) examined RIs' beliefs regarding children's memory and investigative interview practice, and compared these to the beliefs of lay people. Although there was no significant difference in the erroneous belief scores of RIs and lay people, there was a significant difference in knowledge scores with lay people scoring higher than RIs. The difference in knowledge scores appears to be due to the RIs more frequently airing on the side of caution when unsure of the correct response. Caution, amongst legal practitioners, has been observed in previous research (e.g., Kask, 2011) and is thought to be due to the highly consequential judgements which are characteristic of such positions. As a large part of the RI

role is to make recommendations to other legal practitioners (e.g., police and barristers), it is not surprising to see RIs demonstrating similar behaviour. Unfortunately, it is impossible to determine whether the RIs were being overly cautious in the current study or whether they simply lacked the requisite knowledge to provide a correct response.

Study two (chapter four) examined the efficacy of pre-interview communication assessments using the 'Unpacking the Box' framework (Triangle, 2015). Overall, pre-interview assessments were shown to provide a good indication of children's abilities in all areas of cognition, examined in the study, other than resistance to suggestion (predictive accuracy was equally poor across all three experimental conditions). The assessment was found to be superior to professional judgement in ascertaining the children's use of ground rules, responsiveness, and drawing ability; and equivalent to professional judgement in ascertaining children's attentiveness and ability to use a drawing to identify body parts (however the predictions pertaining to the latter two variables were almost at ceiling across all three experimental conditions). Having an awareness of this information should enable police officers (and RIs) to plan and conduct more effective, child-centred interviews. It is therefore recommended that more investigative interviewers are trained to conduct pre-interview assessments. This should not only improve the quality of child interviews but also alleviate concerns, amongst police officers, regarding what a communication assessment should involve, ultimately increasing the likelihood that a communication assessment will be conducted when an RI is not available.

Study three (chapter five) expanded upon previous research examining the variables (i.e., demographic and cognitive) associated with children's recall. It is also the first study, to the author's knowledge, to examine these cognitive factors in relation to children's understanding and use of the ground rules. Having an awareness of these factors is important in terms of allocating limited resources (e.g., RIs) to those children who would benefit most

from them, and also for further developing and improving the efficacy of pre-interview communication assessments. In the current study, visuospatial ability was found to be a significant predictor of correct IRI and use of the 'you got it wrong' rule, expressive language of incorrect IRI, receptive language of confabulated IRI, and attention of accuracy of IRI. Surprisingly age did not enter as a significant predictor in any of the regression models, thus suggesting that cognitive factors may be more important than age in predicting children's event recall and use of the 'you got it wrong' rule.

Study four (chapter six) is the first to examine the impact of the RI provision on practice in real-world investigative interviews. This research was essential in providing an indication of the Scheme's efficacy, along with any areas that may require improvement. The RI provision was found to have some impact upon investigative interview practice, with significantly fewer multiple questions asked in the RI interviews. The use of certain communication aids was also shown to be more prevalent in the interviews with an RI. Yet, the RI provision did not have a direct impact upon levels of rapport or the amount of information provided by the child witnesses, suggesting that there may be other indirect benefits of the RI role. Additional research and training are needed before the true impact and also the potential benefits of the role can be fully realised.

7.2 Theoretical Implications

Theoretically, this thesis is underpinned by Vygotsky's zone of proximal development (1978) and the related concept of scaffolding. The zone of proximal development is defined as the distance between what an individual (i.e., the child witness in the context of this thesis) can achieve without assistance and what they can achieve with scaffolding and support from a more experienced other (i.e., the police officer or RI). The degree and type of scaffolding required can vary from one individual to another.

Determining the Degree of Scaffolding Required

Although previous research (e.g., Chae et al., 2016; Henry et al., 2017) has found age to be a good predictor of children's communicative competence, study two of this thesis, found that some children did not meet age-based expectations. This, in addition to the findings of study three, would suggest that cognitive measures may be more robust predictors of children's event recall than age alone. As such, in order to scaffold a child's communication effectively an RI and / or police officer should assess a child to gain a greater understanding of their cognitive abilities. Training more police officers and RIs in the use of the 'Unpacking the Box' framework (Triangle, 2015) maybe an effective means of implementing this.

Study two of this thesis demonstrated that, overall, the framework provides a good indication of children's communication abilities. It also assesses many of the areas of cognition that were identified in study three of this thesis as being significant predictors of children's event recall (i.e., attention, expressive language, and receptive language). However, 'Unpacking the Box' (Triangle, 2015) does not assess children's visuospatial ability, which was found in study three to be an important predictor of both correct IRI and use of the 'you got it wrong' rule. As study three was exploratory (in relation to the sample size and chosen analyses) further research is essential to fully understand the relationship between visuospatial ability and children's event recall. This will allow researchers and practitioners to make an informed decision as to whether this needs to be assessed prior to an investigative interview and how best to go about this – current assessment frameworks and guidance may need be modified. Having a greater knowledge of the cognitive factors that underpin children's event recall should enable RIs and police officer to conduct more effective pre-interview assessments. This in turn should help them to plan investigative interviews which are more compatible with children's needs, encompassing an appropriate level of scaffolding.

The Efficacy of Verbal and Non-Verbal Scaffolds.

Study four of this thesis expands upon the current knowledge base regarding the efficacy of both verbal and non-verbal scaffolds. Previous research has found that open-ended questions elicit the most detailed accounts from child witnesses (e.g., Gagnon & Cyr, 2017; Luther et al., 2015). As such, the ABE guidance (MoJ, 2011) recommends that these questions be used predominantly during the interview. Although study four of this thesis found further support for the use of open-ended questions (i.e., a positive correlation between the use of open-ended questions and IRI), these questions, as in previous research (e.g., Cederborg et al., 2000; Luther et al., 2015; Sternberg, Lamb, Davies, & Westcott., 2001; Wolfman et al., 2016) were used infrequently.

A potential explanation for this finding is that police officers and RIs have a lack of awareness as to what constitutes an open-ended question. This has been demonstrated both in previous research (e.g., Yi & Lamb, 2018) and in study one of this thesis, in which 65.6% of the RIs misclassified an indirect speech act as an open-ended question. An alternative explanation is that police officers and RIs under-estimate the utility of open-ended questions for eliciting specific details from children (Guadagno et al., 2013). Again, this was demonstrated in study one of this thesis, with approximately a third of RIs unaware that open-ended questions should be used predominantly during an investigative interview, as they elicit the most accurate accounts.

The two examples given above show that not all RIs hold the same beliefs or possess the same level of knowledge regarding memory and investigative interview practice. This may be attributed to when the RIs completed their initial training. Approximately half of the RIs that participated in study one will have completed the revised training, which was introduced in 2018¹⁴. The revised training has a greater coverage of question types and memory which could account for the variability in knowledge scores observed in study one of

this thesis (i.e., knowledge scores ranged from 4 to 15). That being said, it is not clear whether the revised training teaches the RIs specifically about episodic and autobiographical memory (Collins and Krähenbühl [2020] only refer to working memory in their description of the training). It is therefore important that this is addressed in future training and CPD for both experienced RIs and those who have trained more recently (i.e., 2018 onwards). Depending upon the exact content of the revised training, it may be that the more experienced RIs will reap greater benefits from any additional training or CPD.

For three decades best practice guidelines, in England and Wales (e.g., MOGP, Home Office, 1992; ABE, Home Office, 2002; ABE, MoJ, 2011) have considered rapport to be an essential component of an investigative interview. The findings of study four provide support for the continued emphasis upon rapport within the guidelines and add credence to the assertion by Saywitz et al. (2015) that developing a good rapport with a witness can have a positive effect on their communication - with regards to the measure of attention. This is in line with the findings of experimental research with adult witnesses (e.g., Collins et al., 2002) and field research with adult suspects (e.g., Alison et al., 2013; Collins & Carthy, 2019; Holmberg & Christianson, 2002; Walsh & Bull, 2012). Although a significant relationship was found in study four between attention and IRI, no equivalent relationship was found for positivity. This may be due to how positivity was measured. Extracts were taken from the beginning, middle, and end of each interview and the children given an expressivity score (i.e., indicative of positivity). However, this score failed to take into account all of the ways in which children were expressing themselves, for example through drawing (Alford, 2015). This study, along with previous research (e.g., Owen, 2016; Wolfman et al., 2018), suggests

¹⁴The revised training is conducted face-to-face over a two-week period. The training involves classroom presentations, real-life case studies, and a visit to an interview suite and courtroom. See Collins and Krähenbühl (2020) for a more comprehensive description of the revised training.

that drawing is a popular tool amongst police officers and RIs. As such, future research may want to consider taking into account children's drawings when providing ratings of positivity.

Alternatively, the current study may have failed to find a significant relationship due to positivity not being a relevant component of rapport in the context of an investigative interview. Rapport is thought to be context dependent, with the relationship between non-verbal behaviours (e.g., smiling) and positivity highly complex (Tickle-Degnen & Rosenthal, 1990). This is because the same non-verbal behaviours can occur in both positive or negative interactions with their interpretation dependent upon, amongst other things, the roles of the participants and the purpose of the interaction (Tickle-Degnen & Rosenthal, 1990). Tickle-Degnen and Rosenthal's (1990) framework and non-verbal correlates of rapport are based on friendly interactions wherein participants have cooperative goals. Yet in an investigative interview, the interviewer and the interviewee may not share a common goal. Although this maybe more apparent in suspect interviews (and could account for no relationship being found between positivity and IRI in the study by Collin and Carthy, 2019), it is still relevant in witness interviews where children may be reluctant to disclose abuse due to the perceived repercussions (e.g., Plastock, 2018). The findings of the current study suggest that further research is needed which examines the applicability of the three components of rapport (i.e., mutual attentiveness, positivity, and coordination), developed by Tickle-Degnen and Rosenthal (1990), within the context of a child investigative interview.

Study four also examined the efficacy and use of communication aids in child investigative interviews. The use of communication aids was shown to be prevalent in investigative interviews with children. Yet, only drawing / writing was found to impact upon the amount of information the children provided. Although this finding is in line with previous research (with TD children) which has demonstrated the positive effects of drawing (e.g., Butler et al., 1995; Woodford et al., 2015), this relationship only emerged as significant

when children with a concurrent vulnerability were removed from the analysis. This suggests that drawing may not have the same benefits for children with concurrent vulnerabilities as it does for TD children. However, due to a dearth of research very little is known about how drawing impacts upon the communication of children with additional needs within an interview context. To the authors' knowledge the only studies to have explored this have involved sketch reinstatement of context (e.g., ASD; Mattison et al., 2015). This technique differs (in respect to timing and instructions) from how drawing is ordinarily used in ABE (MoJ, 2011) interviews. As such, future research is needed to ensure that there is a solid evidence base supporting the use of drawing in interviews with children with concurrent vulnerabilities.

Registered Intermediaries' Knowledge of Memory and How this Underpins Effective Scaffolding

The success of an investigative interview relies upon the child's ability to remember and accurately report past events (Anderson et al., 2009). However, the competence of the child is highly dependent upon the competence of the interviewer and RI (Marchant, 2013). Competence (or effective professional practice) requires an understanding of the theoretical principles underpinning the best practice guidance – simply being familiar with the guidance is not enough (Ernberg, 2018). The ABE guidance (MoJ, 2011) is largely underpinned by memory research and theory. As such, in order to conduct effective interviews with children both police officers and RIs should have a good understanding of memory and the factors that can influence it. However, study one of this thesis demonstrated significant deficits in the RIs' knowledge of memory. Given that this is what underpins much of the best practice guidance (e.g., question types, communication aids) it was not surprising that study one also identified shortcomings in the RIs knowledge of ABE (MoJ, 2011).

These deficits in memory-related knowledge can be clearly mapped onto the shortcomings observed in the real-world interviews analysed as part of study four. For example, open-ended questions were relatively rare in study four (particularly when considering the questions posed by the RIs). One of the explanations given for this above is that the RIs do not appreciate the efficacy of open-ended questions. This could be due to RIs being unaware that different questions utilise different memory processes (e.g., recall or recognition), which have differential effects upon accuracy (La Rooy et al., 2011). This assertion is supported by the findings of study one, in which almost half of the RIs did not know which memory processes open-ended and forced-choice questions used. A further example, of the RI's deficits in knowledge relates to the inappropriate use of communication aids. Some of the RIs assessed children's knowledge of sexual body parts prior to the investigative interview (presumably using a body diagram). This could be seen as priming the child to discuss these topics at interview. Given that previous research has shown that the use of body diagrams can lead to serious commission errors (e.g., Bruck et al., 2016), this practice is concerning. As such, further training is needed for RIs in memory development. RIs also require further training in identifying appropriate and inappropriate verbal scaffolds (e.g., question types), and how to produce these appropriate scaffolds if required to rephrase. Furthermore, RIs need to recognise the risks associated with non-verbal scaffolds (e.g., communication aids) and how to utilise these in a safe and effective manner.

The Impact of the Registered Intermediary Provision on Interview Practice and Children's Communication

The RI provision can be seen as an overarching approach to scaffolding children's communication. Although study four of this thesis found the RI provision to have some influence on practice (e.g., the use of less multiple questions and increased use of certain communication aids), this did not result in the children providing more detailed accounts. This

is not to say that the RI provision did not scaffold children's communication, but instead may suggest that the provision has other benefits (e.g., improved accuracy and coherence). These benefits may differ dependent upon the needs of the child. For example, if a child has attentional difficulties, the RI's main role may be to keep the child focused and attentive. As such, the RI's presence may serve to scaffold the accuracy as opposed to the detail of the child's account (a significant relationship was found, in study three, between attentional abilities and the accuracy of IRI). An interesting avenue for future research would be to explore the relationship between children's cognitive abilities and the impact of the RI provision on different facets of communication. That being said, if RIs possessed a greater knowledge of memory and the best practice guidance (ABE; MoJ, 2011), specifically in relation to question types (as open-ended questions appeared to be the most effective method of scaffolding communication in study four of this thesis), the provision may have a significant impact upon the detail of children's accounts, irrespective of the children's cognitive abilities.

7.3 Practical Implications

This thesis has provided an insight into the efficacy of pre-interview communication assessments and the knowledge and work of RIs. As such a number of important implications have emerged in respect to current policies and practice.

Implications for Guidance

From discussions with police officers, it appears that age is a significant contributing factor in terms of whether a referral will be made for an RI. Given the findings of studies two and three of this thesis, basing this decision largely on age does not appear to be the most effective approach in terms of prioritising children who are the most in need of RI support. An approach which takes into account a child's cognitive profile is likely to be more effective. Although the ABE guidance (MoJ, 2011) acknowledges that other factors, beyond age, should

be considered in employing the services of an RI (e.g., whether a child can identify a problematic question) the guidance does not go into detail about how these factors should be assessed. This reiterates the importance of additional training for police officers in the administration of some sort of pre-interview assessment. ‘Unpacking the Box’ (Triangle, 2015) would be able to fulfil this role¹⁵. However, this tool is quite advanced and therefore more applicable to planning an interview. A simplified version or alternative tool may be more appropriate for determining whether RI support is needed. The researcher is aware that a police force in Northern England is currently working in collaboration with academics to create something in a similar vein. This assessment, however, should not be seen as a substitute for an RI but instead as a guide to make sure that resources are used where needed and children given the support they require. If this tool is found to be effective the ABE guidance (MoJ, 2011) should signpost police officers to this. The guidance should also consider referencing tools, such as ‘Unpacking the Box’ (Triangle, 2015), for those needing to conduct a more comprehensive assessment of a child’s abilities.

The findings of study two and study three of this thesis would suggest that ‘Unpacking the Box’ (Triangle, 2015) could provide an indication of how likely a child is to produce an accurate account at interview. Study two found that ‘Unpacking the Box’ (Triangle, 2015) was superior to professional judgement in determining a child’s use of ground rules and responsiveness to open-ended questions at interview. Both of these could be considered as factors associated with accuracy. For example, an increased propensity to use the ground rules would suggest that a child will be less likely to guess when they do not know the answer

¹⁵ Study two of this thesis found that ‘Unpacking the Box’ (Triangle, 2015) was superior to professional judgement in ascertaining children’s use of the ground rules, which encompasses within it the ability to identify a problematic question.

therefore, reducing the risk of inaccuracies. Longer responses to open-ended questions should also reduce the risk of inaccuracies or memory intrusions as open-ended questions utilise safer free recall processes (Orbach & Lamb, 2000). Furthermore, expressivity (conceptualised as responsiveness in study two) was found to be a significant predictor of incorrect IRI in study three. As such, having an awareness of these factors should help police officers to determine whether children are able to provide an accurate account of past events and what scaffolds need to be put in place during the interview to help them to do so. The assessment can be used as a forum to introduce and practise some of these scaffolds in order to determine their efficacy. However, the use of 'Unpacking the Box' (Triangle, 2015) should come with a caveat, similar to that included in the section of the guidance on the CI, that appropriate training is required.

Future revisions of the ABE guidance (MoJ, 2011) may also benefit from expanding the section on question types. Given the prevalence of indirect speech acts in study four of this thesis, it may be pertinent for the guidance to discuss and alert interviewers to the impact of pre-facing questions with 'can you...', 'do you know...' etc. In addition, it may be helpful to include the different types of open-ended questions and how these can be used to facilitate the recall of more reluctant children. The findings of studies one and four of this thesis would suggest that RIs may also benefit from the inclusion of a section on question types within their Procedural Guidance Manual (MoJ, 2020a). It may also be helpful to provide the RIs with more concrete examples of what they should and should not examine during a pre-interview assessment. Although, the Registered Intermediary Procedural Guidance Manual (MoJ, 2020a) rightly states that the assessment is dependent upon the child's communication needs and the skill set of the RI, some things are very important to assess (e.g., ability to respond to open-ended questions) whilst some are very important to avoid (e.g., knowledge of sexual body parts).

Elements of Communication that Should be Assessed

This thesis has identified a number of areas of communication which it appears are important to consider during a pre-interview assessment. Each of which will be discussed below:

Drawing Ability. In line with previous research (e.g., Butler et al., 1995; Woodford et al., 2015), study four of this thesis reported beneficial effects of drawing on children's event recall¹⁶. A significant relationship was found between the use of drawing and the amount of information the children provided. This suggests that drawing may be an effective method of scaffolding children's communication in an investigative interview context. However, study two of this thesis demonstrated that children do not always meet age-related expectations in respect to drawing. Thus, a pre-interview communication assessment is important in ascertaining whether children possess this skill.

Expressivity and Ability to Respond to Open-Ended Questions. Of the variables explored in study four, open-ended questions appeared to be the most effective in scaffolding children's communication. As such, examining children's ability to respond to open-ended questions along with which type of open-ended questions (e.g., invitation, breadth, depth) elicit the most detailed accounts, should be seen as a key element of a pre-interview communication assessment. The findings of study two suggest that professional judgement based on age-related norms is not sufficient in determining this. Furthermore, study three of this thesis found expressivity (arguably a related concept) to be a significant predictor of the amount of incorrect IRI children provided. Children with higher expressivity scores were shown to provide less incorrect information during the investigative interview. For children with poorer expressive language abilities and thus a greater propensity to provide incorrect

¹⁶ When the children with a concurrent vulnerability were removed from the analysis.

information, more extensive coverage of the ground rules may be required to reduce the risk of memory errors.

Receptive Language. Study three of this thesis found receptive language to be a significant predictor of children's confabulated IRI. Children with higher receptive language scores provided less confabulated information during the interview. For those children with lower receptive language abilities, greater consideration and planning will need to go into the structure and content of questions (e.g., language). A more extensive coverage of ground rules may also be required to reduce the risk of children ceding to any accidental suggestions made by the interviewer (e.g., recounting / clarifying what the child has said incorrectly – 'you said you were at home, didn't you?'). Misleading questions are often synonymous with complex syntax (e.g., tag questions; Davies & Seymour, 1998) and thus will prove challenging for children with lower receptive language abilities to process and refute.

Use of Ground Rules. Given what is discussed above, it is important to establish prior to an interview whether children are able to use ground rules effectively. Study two of this thesis demonstrated that a pre-interview communication assessment provides a better indication of this than professional judgement alone. Methods (e.g., practice narrative) should therefore be incorporated into the assessment to examine children's use of the rules. All of the rules should be assessed independently as study three suggested that different cognitive abilities may be related to the use of the different rules (visuospatial ability was found to be a significant predictor of the 'you got it wrong' rule; whereas none of the cognitive variables entered into the regression models for the 'I don't know' and 'I don't understand' rules, suggesting that these may be underpinned by other factors).

Attention. Study three of this thesis found attentional abilities to be a significant predictor of children's recall accuracy. Professional judgement (based on age-related norms) was found, in study two, to be equally as effective as a pre-interview assessment in

ascertaining children's attentional abilities. However, due to time constraints this facet of communication could not be fully explored, with predictions at ceiling across all three experimental conditions. Given the potential repercussions of a child providing inaccurate information, attentional abilities should thus still be considered an important area to assess prior to an investigative interview. If a child demonstrates poor attentional abilities, during the assessment, it will also be important to explore strategies to maintain attention - a loss of attention during the interview could compromise the quality of the child's evidence.

Additional Training to Improve Practice

The findings of this thesis would suggest that police officers and RIs require additional training in relation to pre-interview communication assessments, question types, and communication aids. In order to be effective, this training also needs to teach police officers and RIs about the theoretical principles which underpin best practice guidance (ABE; MoJ; 2011). It is an understanding of these principles which will enable them to achieve mastery (McMahon, 2006). As such, training for police officers and RIs needs to include extensive coverage of memory and the factors that can influence it. Ideally, this training should be delivered by a memory expert. It is not clear as to who is currently responsible for teaching this element of the revised RI training. However, the WIS' Annual Report 2018/19 would suggest that it is delivered by a more experienced RI or another criminal justice practitioner (MoJ, 2020b). Yet previous research has shown that criminal justice practitioners internationally, have a limited understanding of and hold incorrect beliefs regarding witness memory (e.g., Dodier, Melinder et al., 2019; Erens et al., 2020; Granhag et al., 2005; Jiang & Luo, 2016), which could potentially have an adverse impact upon the utility of this training.

To be successful the training program also needs to promote continuous development, involve regular feedback, and contain examples of good practice (Powell et al., 2005). It is promising that the revised training for RIs encompasses all of these features. For example,

during the two-week training program there are discussions of real-life cases and following the completion of the training those that are successful receive monthly peer group sessions (Collins & Krähenbühl, 2020). The MoJ has also recently reintroduced the WIS Annual Conference, offering an additional CPD and networking opportunity for RIs (MoJ, 2020b). Further training which focuses upon collaborative working between RIs and police officers may also be beneficial. It is important that both parties understand their respective roles within the interview. This will ensure more effective interviews but also that the RIs impartiality is maintained and cannot be called in to question.

7.4 Methodological Implications

Although this thesis offers an invaluable contribution to the growing body of research into the work of RIs, the findings of each study need to be interpreted in the context of their respective methodologies. As such, the findings and conclusion should be seen as tentative as opposed to definitive.

Participants and Sample Size

It is acknowledged that the sample sizes of all four studies, included within this thesis, are modest. However, in both studies one and four the potential pool of participants (i.e., RIs) / interviews were relatively small. For context, at the start of 2018 there were 183 RIs on the MoJ Register. One hundred and fifty-five RIs were registered to accept cases involving witnesses under the age of 18. However, of these, only 91 (59%) were 'active' (available to accept cases; Plotnikoff & Woolfson, 2019). Thus, although study one only included a sample of 32 RIs, this constitutes approximately one third of the entire target population. The potential pool of RIs was further restricted in study four as only two police forces provided access to interviews. It was decided that to ensure diversity a maximum of three interviews from each RI would be included in the final sample (none of which included the same police officer). This, along with the stringent matching criteria, adopted in study four, significantly

reduced the number of interviews that were eligible for inclusion. In order to provide a more representative overview of practice, across England and Wales, future research should attempt to gain access to a greater number of interviews across a wider geographical area. Due to studies two and three, of this thesis, being exploratory these similarly had modest sample sizes. As such, future research, including a larger sample of participants, may reveal stronger and / or additional significant effects by increasing the power of the studies.

Ecological Validity

A limitation of study two, of this thesis, is that certain aspects of the study's design may be seen to have low ecological validity. The communication assessments and interviews that were conducted, as part of study two, were shorter than is typically observed in a real-world investigation. This not only has implications for how well some of the predictions (e.g., attention) were able to be tested but also for how comprehensive an assessment was able to be conducted. In addition, although carefully selected, the staged-event cannot be considered as equivalent to child abuse in regards to the associated trauma. Trauma can have a profound impact upon how children store, maintain, and recall their experiences (Stakic & Illic, 2018) and as such could have affected the children's interview behaviour. That being said, it would be unethical to include such a manipulation (i.e., the assessment condition which has the potential to improve interview practice) in a real-world setting.

The analysis of real-world data also does not allow for as stringent control measures as laboratory research. For example, in study four, of this thesis, it was only possible to match the interviews according to broad case characteristics (i.e., type of abuse, repeated vs. single) along with the demographics of the witness and interviewer (i.e., age and gender). However, there an abundance of other interview variables which could not be controlled for including the context of the alleged abuse and the cognitive profile of the child. The findings of study three of this thesis suggest that cognitive variables may be more indicative than age of recall

in investigative interviews with children. Given the relative strengths and weaknesses of both laboratory and field research a combination of the two approaches is essential in order to establish a sound evidence base.

Questionnaire Items and Coding Frameworks

This thesis has involved the adaptation and development of a number of questionnaire items and coding frameworks. The questionnaire, utilised as part of study one of this thesis, incorporated both pre-existing and novel statements all based on the findings of previous research. It is important, if this questionnaire is to be used in future research, that any advancements in knowledge are taken into account and the statements adapted accordingly. The distinction should also be made regarding the accuracy (i.e., quality) and detail (i.e., quantity) of children's recall as children's abilities vary by proxy of this (Jack et al., 2014).

It is also recommended that future research, measuring IRI, use the adapted version of Philip et al.'s (2012) framework which was introduced in study four of this thesis. The adapted framework includes two additional categories: conversation and body parts. A considerable amount of the information children provided, in study four, fell into these categories. Most notably, in cases involving online grooming a large proportion of the information the children provided was classified under the category of conversation. With the prevalence of online abuse increasing¹⁷ (Bentley et al., 2016; Kasna & Kelly, 2021) it would be pertinent to include this category within future research to avoid the misrepresentation of the data.

¹⁷ In 2019/20 19% of child sexual offences, reported to the police in England and Wales, involved an online element.

7.5 Areas for Future Research

Throughout this thesis suggestions have been made regarding valuable areas for future research. That being said, the researcher felt it pertinent to draw the reader's attention to three areas believed to be the most important for future exploration.

Children with Concurrent Vulnerabilities and Vulnerable Adults

A recent report from the WIS showed that, in 2019/20, the most frequently cited vulnerability on the Request-for-Service forms was for children without a concurrent vulnerability (MoJ, 2020b). However, this group only constituted 35% of the total requests. Twenty-two percent of requests were made for children with learning disabilities, 4% for children with a mental disorder, and 1% for children with physical disabilities. The remaining requests were made for vulnerable adults (38%; MoJ, 2020b). Despite study four of this thesis including five children with a concurrent vulnerability, all of the other studies only included TD children. As such the conclusions cannot be applied to other populations. Future research with these populations is essential due to the increased propensity for children and adults with a concurrent vulnerability to be a witness or victim of a crime (e.g., Beadle-Brown et al., 2010; Hershkowitz, Lamb, & Horowitz, 2007). Some concurrent vulnerabilities are also associated with memorial difficulties (e.g., ASD; Almeida, 2018). It is therefore recommended that future research explore the knowledge of RIs in respect to concurrent vulnerabilities and memorial deficits, alongside the efficacy of pre-interview communication assessments with different populations of children.

Understanding the Decision-Making of the RI

A recent study by Taggart (2021) examined, via semi-structured interviews, how RIs conceptualise their role and scope within the CJS. During the interviews, the RIs described

some concerning behaviours and actions which could be considered as deviations from their Code of Practice. The Code of Practice states that RIs “must not enter into discussion, give advice, or express opinions concerning the evidence that the witness is to present or any aspect of the case” (MoJ, 2020a, p. 7). Yet, two of the RIs (from England and Wales), in Taggart’s (2021) study, described discussing inconsistencies in a witness’ account with the interviewing officer. This behaviour could be construed as the RIs aligning themselves with the police, thus threatening their perceived impartiality (Taggart, 2021). Similar behaviours were observed in study four, of this thesis, whereby a number of the RIs advised police officers on lines of questioning and asked questions of the child which did not directly relate to the interviewers’ previous question. This could arguably be construed as a violation of the aforementioned Code of Practice and the guidance provided in the procedural manual which states that “the RI is not a second interviewer” (MoJ, 2020a, p. 15). That being said, the manual does state that the RI may discuss with the interviewer during the planning meeting the opening question and how to get to the topic of the alleged abuse. It is therefore easy to see how the lines between what is within the RI role and what is not can become blurred, particularly given how inextricably linked communication and questioning practices are.

An interesting and important avenue for future research would be to examine why, in certain instances, RIs feel it necessary to ask questions of the child that are not in response to a breakdown or potential breakdown of communication, and could therefore potentially compromise perceptions of their impartiality. A think-aloud method (e.g., Wright & Powell, 2006) may be an appropriate approach to explore this, in which a mock interview is conducted and the RI is encouraged throughout to vocalise their thought processes. A similar approach could be adopted to examine the RIs’ decision-making processes during pre-interview communication assessments. Some of the RIs, in study four, of this thesis, reported assessing the children’s knowledge of sexual body parts. This could be seen as priming the

children and could have serious implications for not only the quality, but also the credibility of the children's evidence. Given these potentially serious repercussions, this should also be considered an essential area for further exploration.

The Interview as the Witness' Evidence in Chief

Section 27 of the YJCEA (1999) enables a child's ABE interview (MoJ, 2011) to be used as their evidence in chief in court. This creates additional responsibilities and considerations for police officers and RIs. Not only does the information elicited from the children need to be accurate and detailed, it also needs to be coherent and well-structured so that it may be interpreted by the listener as informative and meaningful. A coherent account is deemed important in allowing jurors to assess the credibility of children's stories (Raskin & Esplin, 1991) and establishing the 'essence of criminality' (i.e., the precise nature of the criminal acts; Guadagno et al., 2006). It is thus children's ability to 'tell their stories' that juries consider, above all else, in their decisions to convict (De Jong & Rose, 1991). Even in countries where juries are used infrequently (i.e., Sweden), judges have been shown to undertake similar evaluations of children's evidence. Evaluations most frequently focus on the richness of the child's testimony. Other factors considered include the length, coherence, and spontaneity of the evidence (Ernberg et al., 2018).

Myklebust and Bjørklund (2009) found that longer responses to open-ended questions were more likely to result in convictions, whilst shorter responses were more likely to result in acquittals or evidence which was deemed insufficient to proceed to court. In a more recent study, it was found that interviews conducted after the introduction of the NICHD (in which proportionally more information, compared to non-protocol interviews, is elicited from witnesses using open-ended questions) were more likely to result in charges being filed and higher conviction rates in the event that the case proceeded to trial (Pipe et al., 2013). The

findings of the aforementioned studies again demonstrate the importance of training police officers and RIs in the effective use of open-ended questions.

The “primary responsibility of the RI is to enable complete, coherent, and accurate communication to take place between the witness and the police or court” (MoJ, 2020a, p. 7). As such an important limitation of study four, of this thesis, is that it only measured the amount of detail provided by the children, not the coherence or overall structure of their narrative accounts. Examining this aspect of communicative competence is an important avenue for future research. It may be that RIs have more of an influence in how children ‘tell their stories’, as opposed to the amount of detail the children provide. Stein and Glenn’s (1979) story-grammar framework would be a useful tool in order to examine this. According to the framework a linguistically complete account should comprise of seven logically sequenced story-grammar elements: the setting, the initiating event, the protagonist’s internal response, the plan, the attempt / action, the consequences, and the resolution / outcome of the story. Previous research, which has adopted this framework, has unfortunately found the prevalence of story grammar in real-world investigative interviews to be low (Snow et al., 2009; Westcott & Kynan, 2004), with children’s stories often described as incomplete, ambiguous, and disordered (Westcott & Kynan, 2004).

7.6 Conclusion

This thesis has expanded upon the limited body of research regarding the RI provision and pre-interview communication assessments. Overall, the findings have demonstrated the efficacy of pre-interview assessments, using the ‘Unpacking the Box’ framework (Triangle, 2015). The framework is underpinned by three areas of cognition (e.g., receptive communication, expressive communication, and attention) which were shown in this thesis to be important predictors of children’s event recall. Pre-interview assessments, using the framework, were also found to be superior to professional judgement in determining

children's use of ground rules, responsiveness, and drawing ability; and equivalent in determining children's attentiveness and ability to use a drawing to identify body parts (which were both at ceiling across all experimental conditions). As such, pre-interview assessments can be considered an appropriate method for establishing the level of scaffolding a child requires, which can then be used to plan a more effective investigative interview.

Other findings within this thesis were not as positive. RIs were found to have poor memory-related knowledge, which appears to have an adverse impact upon other aspects of their practice. For example, the RIs demonstrated a lack of understanding and infrequent use of open-ended questions. This deviation from best practice (ABE; MoJ, 2011) may account for the RI provision, in the interviews analysed as part of this thesis, having no direct impact upon the detail of children's accounts. However, the provision did result in significantly fewer multiple questions and an increase in the use of certain communication aids (e.g., calming tools). This finding mirrors that of previous research (Krähenbühl, 2011) which has shown that RIs have an awareness of developmental and emotional factors, but lack an awareness of the best practice guidelines (ABE; MoJ, 2011). This is likely due to the fact that the guidelines are largely underpinned by memory research and theory, for which the RIs' knowledge is lacking. For the true potential of the RI provision to be realised, further training which covers memory, best practice guidelines (ABE; MoJ, 2011), and most importantly the inter-relationships between the two, is required.

References

- Abbe, A., & Brandon, S. E. (2013). The role of rapport in investigative interviewing: A review. *Journal of Investigative Psychology and Offender Profiling, 10*(3), 237-249.
- Ackerman, B. P. (1987). Selective attention and distraction in context-interactive situations in children and adults. *Journal of Experimental Child Psychology, 44*(1), 126-146.
- Adams, A., Bourke, L., & Willis, C. (1999). Working memory and spoken language comprehension in young children. *International Journal of Psychology, 34*(5), 364-373.
- Agneswaran, A. (2018). *Exploring the experiences of registered intermediaries and police officers in UK of working with adult witnesses with intellectual disabilities* [Doctoral Thesis, Manchester Metropolitan University].
- Agnew, T. (2006). Day in the life: As a witness intermediary, Lorna Pink helps people involved in police investigations and trials. Thelma Agnew reports. *Learning Disability Practice, 9*(2), 39-40.
- Ahern, E. C., Hershkowitz, I., Lamb, M. E., Blasbalg, U., & Winstanley, A. (2014). Support and reluctance in the pre-substantive phase of alleged child abuse victim investigative interviews: Revised versus standard NICHD protocols. *Behavioral Sciences & the Law, 32*(6), 762-774.
- Aldridge, J., & Cameron, S. (1999). Interviewing child witnesses: Questioning techniques and the role of training. *Applied Developmental Science, 3*(2), 136-147.
- Aldridge, J., Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Bowler, L. (2004). Using a human figure drawing to elicit information from alleged victims of child sexual abuse. *Journal of Consulting and Clinical Psychology, 72*(2), 304-316.

- Aldridge, M., Timmins, K., & Wood, K. (1997). Children's understanding of legal terminology: Judges get money at pet shows, don't they? *Child Abuse Review*, 6(2), 141-146.
- Aldridge-Waddon, M. (2019). Police delivery of the opt-out procedure for children's court evidence: evidence of inadequate language awareness. *Language Awareness*, 28(3), 166-185.
- Alford, C. (2015). Drawing: The universal language of children. *Teachers' Work*, 12(1), 45-62.
- Alison, L. J., Alison, E., Noone, G., Elntib, S., & Christiansen, P. (2013). Why tough tactics fail and rapport gets results: Observing Rapport-Based Interpersonal Techniques (ORBIT) to generate useful information from terrorists. *Psychology, Public Policy, and Law*, 19(4), 411-431.
- Allwood, C. M., Innes-Ker, Å., Holmgren, J., & Fredin, G. (2008). Children's and adults' realism in their event-recall confidence in response to free recall and focused questions. *Psychology, Crime & Law*, 14(6), 529-547.
- Almeida, T. (2018). *The impact of autism spectrum disorder on event memory and accuracy* [Doctoral Thesis, University of Cambridge].
- Almerigogna, J., Ost, J., Bull, R., & Akehurst, L. (2007). A state of high anxiety: How non-supportive interviewers can increase the suggestibility of child witnesses. *Applied Cognitive Psychology*, 21(7), 963-974.
- Anderson, J. (2013). The CornerHouse forensic interview protocol: An evolution in practice for almost 25 years. *APSAC Advisor*, 4, 2-7.
- Anderson, G. D., Anderson, J. N., & Gilgun, J. F. (2014). The influence of narrative practice techniques on child behaviors in forensic interviews. *Journal of Child Sexual Abuse*, 23(6), 615-634.

- Anderson, J., Ellefson, J., Lashley, J., Miller, A. L., Olinger, S., Russell, A., Stauffer, J., & Weigman, J. (2009). The CornerHouse forensic interview protocol: RATAc. *Thomas M. Cooley Journal of Practical. & Clinical Law.*, 12, 193-331.
- Andrade, J., & Donaldson, L. (2007). Evidence for an olfactory store in working memory? *Psychologia*, 50(2), 76-89.
- Andrews, S. J., Lamb, M. E., & Lyon, T. D. (2015). Question types, responsiveness and self-contradictions when prosecutors and defense attorneys question alleged victims of child sexual abuse. *Applied Cognitive Psychology*, 29(2), 253-261.
- Andrews, S. J., & Lamb, M. E. (2014). The effects of age and delay on responses to repeated questions in forensic interviews with children alleging sexual abuse. *Law and Human Behavior*, 38(2), 171-180.
- Association of Chief Police Officers. (2010). *Advice on the Structure of Visually Recorded Witness Interviews*.
- Azzopardi, C., Eirich, R., Rash, C. L., MacDonald, S., & Madigan, S. (2019). A meta-analysis of the prevalence of child sexual abuse disclosure in forensic settings. *Child Abuse & Neglect*, 93, 291-304.
- Baddeley, A. D., Eysenck, M. W., & Anderson, M. C. (2014). *Memory* (2nd edn). Psychology Press.
- Baker-Ward, L., Gordon, B. N., Ornstein, P. A., Larus, D. M., & Clubb, P. A. (1993). Young children's long-term retention of a paediatric examination. *Child development*, 64(5), 1519-1533.
- Barlow, C. M., Jolley, R. P., & Hallam, J. L. (2011). Drawings as memory aids: Optimising the drawing method to facilitate young children's recall. *Applied Cognitive Psychology*, 25(3), 480-487.

- Beadle-Brown, J., Mansell, J., Cambridge, P., Milne, A., & Whelton, B. (2010). Adult protection of people with intellectual disabilities: Incidence, nature and responses. *Journal of Applied Research in Intellectual Disabilities*, 23(6), 573-584.
- Behzadnia, A., & Mehrani, M. B. (2017). Young children's yes bias in response to tag questions. *Early Child Development and Care*, 188(12), 1665-1674.
- Bentley, H., O'Hagan, O., Raff, A., & Bhatti, I. (2016). *How safe are our children? The most comprehensive overview of child protection in the UK*. NSPCC.
- Berkowitz, S. R., Garrett, B. L., Fenn, K. M., & Loftus, E. F. (2022). Convicting with confidence? Why we should not over-rely on eyewitness confidence. *Memory*, 30(1), 10-15.
- Bernieri, F. J., Gillis, J. S., Davis, J. M., & Grahe, J. E. (1996). Dyad rapport and the accuracy of its judgment across situations: A lens model analysis. *Journal of Personality and Social Psychology*, 71(1), 110-129.
- Beuscher, E., & Roebbers, C. M. (2005). Does a warning help children to more accurately remember an event, to resist misleading questions, and to identify unanswerable questions? *Experimental Psychology*, 52(3), 232-241.
- Blasbalg, U., Hershkowitz, I., & Karni-Visel, Y. (2018). Support, reluctance, and production in child abuse investigative interviews. *Psychology, Public Policy, and Law*, 24(4), 518-527.
- Blasbalg, U., Hershkowitz, I., Lamb, M. E., Karni-Visel, Y., & Ahern, E. C. (2019). Is interviewer support associated with the reduced reluctance and enhanced informativeness of alleged child abuse victims? *Law and Human Behavior*, 43(2), 156-165.

- Boyle, W., Lindell, A. K., & Kidd, E. (2013). Investigating the role of verbal working memory in young children's sentence comprehension. *Language learning, 63*(2), 211-242.
- Brackmann, N., Sauerland, M., & Otgaar, H. (2019). Developmental trends in lineup performance: Adolescents are more prone to innocent bystander misidentifications than children and adults. *Memory & cognition, 47*(3), 428-440.
- Brainerd, C. J., Kingma, J., & Howe, M. L. (1985). On the development of forgetting. *Child Development, 56*(5), 1103-1119.
- Brainerd, C. J., & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review, 32*(3), 224-267.
- Brewer, N., & Day, K. (2005). The Confidence-Accuracy and Decision Latency-Accuracy Relationships in Children's Eyewitness Identification. *Psychiatry, Psychology and Law, 12*(1), 119-128.
- Brown, D. A., & Lamb, M. E. (2015). Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives, 9*(4), 250-255.
- Brown, D. A., Lamb, M. E., Lewis, C., Pipe, M., Orbach, Y., & Wolfman, M. (2013). The NICHD investigative interview protocol: An analogue study. *Journal of Experimental Psychology: Applied, 19*(4), 367-382.
- Brown, D. A., Lewis, C. N., Lamb, M. E., Gwynne, J., Kitto, O., & Stairmand, M. (2019). Developmental differences in children's learning and use of forensic ground rules during an interview about an experienced event. *Developmental Psychology, 55*(8), 1626-1639.
- Brown, D., & Pipe, M. E. (2003). Individual differences in children's event memory reports and the narrative elaboration technique. *Journal of Applied Psychology, 88*(2), 195-206.

- Brown, D., Pipe, M., Lewis, C., Lamb, M. E., & Orbach, Y. (2012). How do body diagrams affect the accuracy and consistency of children's reports of bodily touch across repeated interviews? *Applied Cognitive Psychology, 26*(2), 174-181.
- Brubacher, S. P., Peterson, C., La Rooy, D., Dickinson, J. J., & Poole, D. A. (2019). How children talk about events: Implications for eliciting and analyzing eyewitness reports. *Developmental Review, 51*, 70-89.
- Brubacher, S. P., Poole, D. A., & Dickinson, J. J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Review, 36*, 15-33.
- Brubacher, S. P., Poole, D. A., Dickinson, J. J., La Rooy, D., Szojka, Z. A., & Powell, M. B. (2019). Effects of interviewer familiarity and supportiveness on children's recall across repeated interviews. *Law and Human Behavior, 43*(6), 507-516.
- Brubacher, S. P., Roberts, K. P., & Powell, M. (2011). Effects of practicing episodic versus scripted recall on children's subsequent narratives of a repeated event. *Psychology, Public Policy, and Law, 17*(2), 286-314.
- Brubacher, S. P., Timms, L., Powell, M., & Bearman, M. (2019). "She Wanted to Know the Full Story": Children's Perceptions of Open Versus Closed Questions. *Child Maltreatment, 24*(2), 222-231.
- Bruck, M. (2009). Human figure drawings and children's recall of touching. *Journal of Experimental Psychology: Applied, 15*(4), 361-374.
- Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50*(1), 419-439.
- Bruck, M., Ceci, S. J., & Francoeur, E. (2000). Children's use of anatomically detailed dolls to report genital touching in a medical examination: Developmental and gender comparisons. *Journal of Experimental Psychology: Applied, 6*(1), 74-83.

- Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*(3), 520-554.
- Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology, 18*(8), 947-996.
- Bruck, M., Kelley, K., & Poole, D. A. (2016). Children's reports of body touching in medical examinations: The benefits and risks of using body diagrams. *Psychology, Public Policy, and Law, 22*(1), 1-11.
- Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B. N., & Ornstein, P. A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology, 15*(7), 25-48.
- Burrows, K. S., Bearman, M., Dion, J., & Powell, M. B. (2017). Children's use of sexual body part terms in witness interviews about sexual abuse. *Child Abuse & Neglect, 65*, 226-235.
- Burrows, K. S., & Powell, M. (2013). Prosecutors' recommendations for improving child witness statements about sexual abuse. *Policing and Society, 24*(2), 189-207.
- Burton, M., Evans, R., & Sanders, A. (2006). Implementing special measures for vulnerable and intimidated witnesses: The problem of identification. *Criminal Law Review, 229*-240.
- Butler, S., Gross, J., & Hayne, H. (1995). The effect of drawing on memory performance in young children. *Developmental Psychology, 31*(4), 597-608.
- Calado, B., Otgaar, H., & Muris, P. (2018). Are children better witnesses than adolescents? Developmental trends in different false memory paradigms. *Journal of Child Custody, 15*(4), 330-348.

- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28(2), 595–616.
- Carol, R. N., & Compo, N. S. (2017). Other people: A child's age predicts a source's effect on memory. *Legal and Criminological Psychology*, 22(1), 74-87.
- Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law and Human Behavior*, 20(3), 335-358.
- Cashmore, J., & Shackel, R. (2018). *Evaluation of the child sexual offence evidence pilot: Final outcome evaluation report*. University of New South Wales.
- Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113(3), 403-439.
- Cederborg, A., Alm, C., Lima da Silva Nises, D., & Lamb, M. E. (2013). Investigative interviewing of alleged child abuse victims: An evaluation of a new training programme for investigative interviewers. *Police Practice and Research*, 14(3), 242-254.
- Cederborg, A., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse & Neglect*, 24(10), 1355-1361.
- Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology*, 19(4), 383-407.
- Chae, Y., Goodman, M., Goodman, G. S., Troxel, N., McWilliams, K., Thompson, R. A., Shaver, P. R. & Widaman, K. F. (2018). How children remember the Strange Situation: The role of attachment. *Journal of Experimental Child Psychology*, 166, 360-379.

- Chae, Y., Kulkofsky, S., Debaran, F., Wang, Q., & Hart, S. L. (2014). Low-SES children's eyewitness memory: The effects of verbal labels and vocabulary skills. *Behavioral Sciences & the Law*, 32(6), 732-745.
- Chae, Y., Kulkofsky, S., Debaran, F., Wang, Q., & Hart, S. L. (2016). Low-SES preschool children's eyewitness memory: The role of narrative skill. *Behavioral Sciences & the Law*, 34(1), 55-73.
- Chaplin, C., & Shaw, J. (2016). Confidently wrong: Police endorsement of psycho-legal misconceptions. *Journal of Police and Criminal Psychology*, 31(3), 208-216.
- Cheit, R. E. (2014). *The witch-hunt narrative: Politics, psychology, and the sexual abuse of children*. Oxford University Press.
- Christianson, S. Å., & HübINETTE, B. (1993). Hands up! A study of witnesses' emotional reactions and memories associated with bank robberies. *Applied Cognitive Psychology*, 7(5), 365-379.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Collins, K. (2012). *Rapport building in child investigative interviews* [Doctoral Thesis, University of Stirling].
- Collins, K., & Carthy, N. (2019). No rapport, no comment: The relationship between rapport and communication during investigative interviews with suspects. *Journal of Investigative Psychology and Offender Profiling*, 16(1), 18-31.
- Collins, K., Doherty-Sneddon, G., & Doherty, M. J. (2014). Practitioner perspectives on rapport building during child investigative interviews. *Psychology, Crime and Law*, 20(9), 884-901.

- Collins, K., Harker, N., & Antonopoulos, G. A. (2017). The impact of the registered intermediary on adults' perceptions of child witnesses: Evidence from a mock cross examination. *European Journal on Criminal Policy and Research*, 23(2), 211-225.
- Collins, K., & Krähenbühl, S. (2020). Registered intermediaries' assessment of children's communication: An exploration of aims and processes. *The International Journal of Evidence & Proof*, 24(4), 374-395.
- Collins, R., Lincoln, R., & Frank, M. G. (2002). The effect of rapport in forensic interviewing. *Psychiatry, Psychology and Law*, 9(1), 69-78.
- Connolly, D. A., Price, H. L., Lavoie, J. A., & Gordon, H. M. (2008). Perceptions and predictors of children's credibility of a unique event and an instance of a repeated event. *Law and Human Behaviour*, 32(1), 92-112.
- Colloff, M. F., Flowe, H. D., Smith, H. M. J., Seale-Carlisle, T. M., Meissner, C. A., Rockey, J. C., Pande, B., Kujur, P., Parveen, N., Chandel, P., Singh, M. M., Pradhan, S., & Parganiha, A. (2022). Active exploration of faces in police lineups increases discrimination accuracy. *American Psychologist*, 77(2), 196-220.
- Cooper, P. (2009). *Tell me what's happening: Registered Intermediary Survey 2009*. City University.
- Cooper, P. (2011). *Tell me what's happening 2: Registered Intermediary Survey 2010*. City University.
- Cooper, P. (2012). *Tell me what's happening 3: Registered Intermediary Survey 2011*. City University.
- Cooper, P. (2014). *Highs and lows: The 4th Intermediary Survey*. Kingston University.
- Cooper, P., Dando, C., Ormerod, T., Mattison, M., Marchant, R., Milne, R., & Bull, R. (2018). One step forward and two steps back. The "20 Principles" for questioning vulnerable

- witnesses and the lack of an evidence-based approach. *The International Journal of Evidence & Proof*, 22(4), 392-410.
- Cordón, I. M., Saetermoe, C. L., & Goodman, G. S. (2005). Facilitating children's accurate responses: Conversational rules and interview style. *Applied Cognitive Psychology*, 19(3), 249-266.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural and Brain Sciences*, 24(1), 87-114.
- Criminal Justice Joint Inspection. (2014). *Achieving best evidence in child sexual abuse case—A joint inspection*. Crown Prosecution Service Inspectorate.
- Cyr, M., Dion, J., McDuff, P., & Trotier-Sylvain, K. (2012). Transfer of skills in the context of non-suggestive investigative interviews: Impact of structured interview protocol and feedback. *Applied Cognitive Psychology*, 26(4), 516-524.
- Danby, M. C., Brubacher, S. P., Sharman, S. J., & Powell, M. B. (2015). The effects of practice on children's ability to apply ground rules in a narrative interview. *Behavioral Sciences & the Law*, 33(4), 446-458.
- Danby, M. C., Brubacher, S. P., Sharman, S. J., & Powell, M. B. (2017). The effects of one versus two episodically orientated practice narratives on children's reports of a repeated event. *Legal and Criminological Psychology*, 22(2), 442-454.
- Danby, M. C., Sharman, S. J., Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2017). Differential effects of general versus cued invitations on children's reports of a repeated event episode. *Psychology, Crime and Law*, 23(8), 794-811.
- Dando, C. J., Ormerod, T. C., Cooper, P., Marchant, R., Mattison, M., Milne, R., & Bull, R. (2018). No evidence against sketch reinstatement of context, verbal labels or the use of Registered Intermediaries for children with Autism Spectrum Disorder: Response to

- Henry et al. (2017). *Journal of Autism and Developmental Disorders*, 48(7), 2593-2596.
- Dando, C., Wilcock, R., & Milne, R. (2009). The cognitive interview: The efficacy of a modified mental reinstatement of context procedure for frontline police investigators. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 23(1), 138-147.
- Daniel, T. A., & Katz, J. S. (2018). Primacy and recency effects for taste. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 399-405.
- Davies, G. M., Bull, R., Milne, R. J. (2016). Analysing and improving the testimony of vulnerable witnesses interviewed under the 'Achieving best evidence' protocol. In P. Radcliffe, G. H. Gudjonsson, A. Heaton-Armstrong, & D. Wolchover (Eds.), *Witness Testimony in Sexual Cases: Evidential, Investigative and Scientific Perspectives* (pp. 207-219). Oxford University Press.
- Davies, E., & Seymour, F. W. (1998). Questioning child complainants of sexual abuse: Analysis of criminal court transcripts in New Zealand. *Psychiatry, Psychology and Law*, 5(1), 47-61.
- Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, 6(2), 81-97.
- Davis, S. L., & Bottoms, B. L. (2002). Effects of social support on children's eyewitness reports: A test of the underlying mechanism. *Law and Human Behavior*, 26(2), 185-215.
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and human behavior*, 28(6), 687-706

- De Jong, A. R., & Rose, M. (1991). Legal proof of child sexual abuse in the absence of physical evidence. *Pediatrics*, *88*(3), 506–511.
- Department of Justice. (2015). *Northern Ireland Registered Intermediaries Schemes Pilot Project: Post-Project Review*. Dickinson, J. J., Brubacher, S. P., & Poole, D. A. (2015). Children's performance on ground rules questions: Implications for forensic interviewing. *Law and Human Behavior*, *39*(1), 87-97.
- Dickinson, J. J., & Poole, D. A. (2017). The influence of disclosure history and body diagrams on children's reports of inappropriate touching: Evidence from a new analog paradigm. *Law and Human Behavior*, *41*(1), 1-12.
- Dodier, O., & Denault, V. (2018). The Griffiths Question Map: A forensic tool for expert witnesses' assessments of witnesses' and victims' statements. *Journal of Forensic Sciences*, *63*(1), 266-274.
- Dodier, O., Melinder, A., Otgaar, H., Payoux, M., & Magnussen, S. (2019). Psychologists and psychiatrists in court: What do they know about eyewitness memory? A comparison of experts in inquisitorial and adversarial legal systems. *Journal of Police and Criminal Psychology*, *34*(3), 254-262.
- Dodier, O., & Otgaar, H. (2019). The forensic and clinical relevance of evidence-based investigative interview methods in historic sexual abuse cases. *Clinical Psychological Science*, *7*(6), 1244-1248.
- Dodier, O., & Payoux, M. (2017). Connaissances et croyances des psychologues et psychiatres experts judiciaires concernant le fonctionnement de la mémoire. *L'Année psychologique*, *117*(2), 139-171.
- Dodier, O., Tomas, F., Payoux, M., & Elissalde, B. (2019). Professional experience in investigative interviewing does not guarantee strong knowledge about memory. *Psychological Research on Urban Society*, *2*(1), 117-125.

- Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & cognition*, *33*(4), 727-733.
- Dunn, L. M., Dunn, D. M., & Styles, B. (2009). *British Picture Vocabulary Scale* (3rd ed.). GL Assessment.
- Earhart, B., La Rooy, D. J., Brubacher, S. P., & Lamb, M. E. (2014). An examination of 'don't know' responses in forensic interviews with children. *Behavioral Sciences & the Law*, *32*(6), 746-761.
- Eisen, M. L., Qin, J., Goodman, G. S., & Davis, S. L. (2002). Memory and suggestibility in maltreated children: Age, stress, arousal, dissociation, and psychopathology. *Journal of Experimental Child Psychology*, *83*(3), 167-212.
- Ellis, L. M., Powell, M. B., Thomson, D. M., & Jones, C. (2003). Do simple "Groundrules" reduce preschoolers' suggestibility about experienced and nonexperienced events? *Psychiatry, Psychology and Law*, *10*(2), 334-345.
- Erens, B., Otgaar, H., Patihis, L., & De Ruiter, C. (2020). Beliefs about children's memory and child investigative interviewing practices: A survey in Dutch child protection professionals from 'Safe Home'. *Frontiers in Psychology*, *11*, 1-9.
- Ernberg, E. (2018). There was nothing but her story: Prosecution of alleged child sexual abuse of preschoolers [Doctoral Dissertation, University of Gothenburg].
- Ernberg, E., Magnussen, M., Landstrom, S., & Tidefors, I. (2018). Court evaluations of young children's testimony in child sexual abuse cases. *Legal and Criminological Psychology*, *23*(2), 176-191.
- Evans, A. D., & Lyon, T. D. (2012). Assessing children's competency to take the oath in court: The influence of question type on children's accuracy. *Law and Human Behavior*, *36*(3), 195-205.

- Evans, A. D., Roberts, K. P., Price, H. L., & Stefek, C. P. (2010). The use of paraphrasing in investigative interviews. *Child Abuse and Neglect, 34*(8), 585-592.
- Evans, A. D., Stolzenberg, S. N., Lee, K., & Lyon, T. D. (2014). Young children's difficulty with indirect speech acts: Implications for questioning child witnesses. *Behavioral Sciences & the Law, 32*(6), 775-788.
- Everson, M. D., & Boat, B. W. (1994). Putting the anatomical doll controversy in perspective: An examination of the major uses and criticisms of the dolls in child sexual abuse evaluations. *Child Abuse and Neglect, 18*(2), 113-129.
- Eysenck, M. W. (2015). Memory in childhood. In A. Baddeley, M. W. Eysenck, & M. C. Anderson (Eds.). *Memory* (pp. 381-411). Psychology Press.
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion, 6*(6), 409-434.
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353.
- Eysenck, M. W., & Keane, M. T. (2015). *Cognitive Psychology: A Student's Handbook* (7th ed.) Psychology Press.
- Fairclough, S. (2017). 'It doesn't happen... and I've never thought it was necessary for it to happen' Barriers to vulnerable defendants giving evidence by live link in crown court trials. *The International Journal of Evidence and Proof, 21*(3), 209-229.
- Fairclough, S. (2018). Using Hawkin's surround, field, and frames concepts to understand the complexities of special measures decision making in crown court trials. *Journal of Law and Society, 45*(3), 457-485.
- Feltis, B. B., Powell, M. B., Snow, P. C., & Hughes-Scholes, C. H. (2010). An examination of the association between interviewer question type and story-grammar detail in child witness interviews about abuse. *Child Abuse and Neglect, 34*(6), 407-413.

- Fessinger, M. B., & McAuliff, B. D. (2020). A national survey of child forensic interviewers: Implications for research, practice, and law. *Law and Human Behavior, 44*(2), 113-127.
- Fischer, K. W., Bullock, D. H., Rosenberg, E. J., & Raya, P. (1993). The dynamics of competence: How context contributes directly to skill. In R. H. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments – The Jean Piaget symposium series* (pp. 275-305). Plenum.
- Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Charles Thomas Publishers.
- Fivush, R., & Hamond, N. R. (1989). Time and again: Effects of repetition and retention interval on 2 year olds' event recall. *Journal of Experimental Child Psychology, 47*(2), 259-273.
- Fivush, R., & Nelson, K. (2004). Culture and language in the emergence of autobiographical memory. *Psychological Science, 15*(9), 573-577.
- Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development, 37*(2), 283-299.
- Flin, R., Boon, J., Knox, A., & Bull, R. (1992). The effect of a five-month delay on children's and adults' eyewitness memory. *British Journal of Psychology, 83*(3), 323-336.
- Fritzley, V. H., & Lee, K. (2003). Do young children always say yes to yes-no questions? A metacognitive study of the affirmation bias. *Child Development, 74*(5), 1297-1313.
- Fritzley, V. H., Lindsay, R. C. L., & Lee, K. (2013). Young children's response tendencies toward yes-no questions concerning actions. *Child Development, 84*(2), 711-725.
- Frye, D., Zelazo, P.D. & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483-527.

- Gagnon, K., & Cyr, M. (2017). Sexual abuse and preschoolers: Forensic details in regard of question types. *Child Abuse & Neglect*, *67*, 109-118.
- Garven, S., Wood, J. M., Malpass, R. S., & Shaw III, J. S. (1998). More than Suggestion: The effect of interviewing techniques from the McMartin Preschool case. *Journal of Applied Psychology*, *83*(3), 347-359.
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. The structure of working memory from 4 to 15 years. *Developmental Psychology*, *40*(2), 177-190.
- Geddes, G. (2016). The price of justice: Can you hear me at the back? *Family Law*, *46*(7), 833-840.
- Geddie, L. F., Beer, J., Bartosik., S., & Wuensch, K. L. (2001). The relationship between interview characteristics and accuracy of recall in young children: Do individual differences matter? *Child Maltreatment*, *6*(1), 59-68.
- Gee, S., Gregory, M., & Pipe, M. (1999). 'What colour is your pet dinosaur?' the impact of pre-interview training and question type on children's answers. *Legal and Criminological Psychology*, *4*(1), 111-128.
- Geiselman, R. E., & Fisher, R. P. (2014). Interviewing witnesses and victims. In M. St-Yves (Ed.). *Investigative interviewing: Handbook of best practices*. Thomson Reuters Publishers.
- Gentle, M., Powell, M. B., & Sharman, S. J. (2014). Mental context reinstatement or drawing: Which better enhances children's recall of witnessed events and protects against suggestive questions? *Australian Journal of Psychology*, *66*(3), 158-167.
- Gerstadt, C., Hong, Y. & Diamond, A. (1994). The relationship between cognition and action: Performance of 3½ - 7 year old children on a Stroop-like day-night test. *Cognition*, *53*(2), 129-153.

- Ginet, M., & Verkampt, F. (2007). The cognitive interview: Is its benefit affected by the level of witness emotion? *Memory, 15*(4), 450-464.
- Goetzold, S. (2015). An open and shut case of closed questions: An exploration of joint investigative interview training in Scotland. *Child Abuse Review, 26*(2), 116-129.
- Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and Life History, 1*(1), 69-99.
- Goodman, G. S., Jones, O., & McLeod, C. (2017). Is there consensus about children's memory and suggestibility? *Journal of interpersonal violence, 32*(6), 926-939.
- Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1997). Children's reactions to and memory for a stressful event: Influences of age, anatomical dolls, knowledge, and parental attachment. *Applied Developmental Science, 1*(2), 54-75.
- Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P., & Gordon, D. S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect, 27*(5), 525-540.
- Gotzke, C., & Gosse, H. S. (2009). Interacting (37-60 months) – Increasingly social communicators. In the Canadian Language and Literacy Research Network (2009). *Handbook of language and literacy development: A roadmap from 0 to 60 months*. <http://www.theroadmap.ualberta.ca/interactions/research/37-60>.
- Guadagno, B. L., Hughes-Scholes, C. H., & Powell, M. B. (2013). What themes trigger investigative interviewers to ask specific questions when interviewing children? *International Journal of Police Science & Management, 15*(1), 51-60.

- Guadagno, B., Powell, M. B., & Wright, R. (2006). Police officers' and legal professionals' perception regarding how children are, and should be questioned about repeated abuse. *Psychiatry, Psychology and Law, 13*(2), 251–260.
- Granhag, P. A., Strömwall, L. A., & Hartwig, M. (2005). Eyewitness testimony: Tracing the beliefs of Swedish legal professionals. *Behavioral Sciences & the Law, 23*(5), 709-727.
- Griffiths, A., & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Eds.), *Investigative Interviewing* (pp.167-189). Willan Publishing.
- Gross, J., & Hayne, H. (1998). Drawing facilitates children's verbal reports of emotionally laden events. *Journal of Experimental Psychology: Applied, 4*(2), 163-179.
- Gruhn, M. A., & Compas, B. E. (2020). Effects of maltreatment on coping and emotion regulation in childhood and adolescence: A meta-analytic review. *Child Abuse & Neglect*. <https://doi.org/10.1016/j.chiabu.2020.104446>
- Hale, S., & Jansen, J. (1994). Global processing-time coefficients characterise individual and group differences in cognitive speed. *Psychological Science, 5*(6), 384-389.
- Hamlyn, B., Phelps, A., Turtle, J., & Sattar, G. (2004). *Are special measures working? Evidence from Surveys of Vulnerable and Intimidated Witnesses*. Home Office.
- Hanna, K., & Henderson, E. (2018). '[Expletive], that was confusing, wasn't it?' Defence lawyers' and intermediaries' assessment of the language used to question a child witness. *The International Journal of Evidence and Proof, 22*(4), 411-427.
- Hanway, P., Akehurst, L., Vernham, Z., & Hope, L. (2020). The effects of cognitive load during an investigative interviewing task on mock interviewers' recall of information. *Legal and Criminological Psychology, 26*(1), 25-41.

- Hardy, C. L., & Van Leeuwen, S. A. (2004). Interviewing young children: Effects of probe structures and focus of rapport-building talk on the qualities of young children's eyewitness statements. *Canadian Journal of Behavioural Science, 36*(2), 155-165.
- Hemmer, P., & Steyvers, M. (2009). A Bayesian account of reconstructive memory. *Topics in Cognitive Science, 1*(1), 189-202.
- Henderson, H. M., & Lamb, M. (2017). Pre-recording children's trial testimony: Effects on the criminal justice system. *Criminal Law Review, 5*, 345-356.
- Henry, L. A., Crane, L., Millmore, A., Nash, G., & Wilcock, R. (2021). Intermediaries and cross-examination resilience in children: The development of a novel experimental methodology. *Applied Cognitive Psychology, 35*(6), 1387-1404.
- Henry, L. A., Crane, L., Nash, G., Hobson, Z., Kirke-Smith, M., & Wilcock, R. (2017). Verbal, visual, and intermediary support for child witnesses with autism during investigative interviews. *Journal of Autism and Developmental Disorders, 47*(8), 2348-2362.
- Henry, L. A., & Gudjonsson, G. H. (2007). Individual and developmental differences in eyewitness recall and suggestibility in children with intellectual disabilities. *Applied Cognitive Psychology, 21*(3), 361-381.
- Henry, L. A., Messer, D. J., Wilcock, R., Nash, G., Kirke-Smith, M., Hobson, Z., & Crane, L. (2017). Do measures of memory, language, and attention predict eyewitness memory in children with and without autism? *Autism & Developmental Language Impairments, 2*, 1-17.
- Hershkowitz, I. (2011). Rapport building in investigative interviews of children. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 109-128). Wiley-Blackwell.

- Hershkowitz, I., & Lamb, M. E. (2020). Allegation rates and credibility assessment in forensic interviews of alleged child abuse victims: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, and Law*, 26(2), 176-184.
- Hershkowitz, I., Lamb, M. E., Blasbalg, U., & Karni-Visel, Y. (2021). The dynamics of two-session interviews with suspected victims of abuse who are reluctant to make allegations. *Development and Psychopathology*, 33(2), 739-747.
- Hershkowitz, I., Lamb, M. E., & Horowitz, D. (2007). Victimization of children with disabilities. *American Journal of Orthopsychiatry*, 77(4), 629-635.
- Hershkowitz, I., Lamb, M. E., & Katz, C. (2014). Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, and Law*, 20(3), 336-344.
- Hershkowitz, I., Lamb, M. E., Orbach, Y., Katz, C., & Horowitz, D. (2012). The development of communicative and narrative skills among preschoolers: Lessons from forensic interviews about child abuse. *Child Development*, 83(2), 611-622.
- Hershkowitz, I., Lanes, O., & Lamb, M. E. (2007). Exploring the disclosure of child sexual abuse with alleged victims and their parents. *Child Abuse & Neglect*, 31(2), 111-123.
- Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J., & Horowitz, D. (2006). Dynamics of forensic interviews with suspected abuse victims who do not disclose abuse. *Child Abuse & Neglect*, 30(7), 753-769.
- Hill, E. S., & Davies, G. M. (2013). Has the quality of investigative interviews with children improved with changes in guidance? An exploratory study. *Policing: A Journal of Policy and Practice*, 7(1), 63-71.
- Hlavka, H. R., Olinger, S. D., & Lashley, J. L. (2010). The use of anatomical dolls as a demonstration aid in child sexual abuse interviews: A study of forensic interviewers' perceptions. *Journal of Child Sexual Abuse*, 19(5), 519-553.

- HMICFRS. (2019). *National child protection post-inspection review: Metropolitan police service*.
- HMICFRS. (2018). *PEEL: Police effectiveness 2017: A national overview*.
- Hobbs, S. D., & Goodman, G. S. (2018). Self-representation: Pro se cross-examination and revisiting trauma upon child witnesses. *International Journal on Child Maltreatment: Research, Policy and Practice*, 1(1), 77-95.
- Holliday, R. E., Humphries, J. E., Brainerd, C. J., & Reyna, V. F. (2011). Interviewing Vulnerable Witnesses. In G. Davies, & A. Beech (Eds.), *Forensic Psychology: Crime, Justice, Law, Interventions* 2nd Edition (pp. 115-134). BPS Blackwell.
- Holliday, R. E., Reyna, V. F., & Hayes, B. K. (2002). Memory processes underlying misinformation effects in child witnesses. *Developmental Review*, 22(1), 37-77.
- Holmberg, U., & Christianson, S. Å. (2002). Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behavioral Sciences & the Law*, 20(1-2), 31-45.
- Home Office. (1992). *Memorandum of Good Practice for Video Recorded Interviews with Child Witnesses for Criminal Proceedings*. HMSO.
- Home Office. (2002). *Achieving the best evidence in criminal proceedings: Guidance for vulnerable and intimidated witnesses, including children*.
- Hoogesteyn, K., Meijer, E., & Vrij, A. (2020). Examining witness interviewing environments. *Journal of Investigative Psychology and Offender Profiling*, 17(3), 238-249.
- Houston, K. A., Hope, L., Memon, A., & Don Read, J. (2013). Expert testimony on eyewitness evidence: In search of common sense. *Behavioral Sciences & the Law*, 31(5), 637-651.
- Houwen, S., Visser, L., van der Putten, A., & Vlaskamp, C. (2016). The interrelationships between motor, cognitive and language development in children with and without

- intellectual and developmental disabilities. *Research in developmental disabilities*, 53, 19-31.
- Howie, P., Sheehan, M., Mojarrad, T., & Wrzesinska, M. (2004). 'Undesirable' and 'desirable' shifts in children's responses to repeated questions: Age differences in the effect of providing a rationale for repetition. *Applied Cognitive Psychology*, 18(9), 1161-1180.
- Hritz, A. C., Royer, C. E., Helm, R. K., Burd, K. A., Ojeda, K., & Ceci, S. J. (2015). Children's suggestibility research: Things to know before interviewing a child. *Anuario de Psicología Jurídica*, 25(1), 3-12.
- Huang, K. J., & Teoh, Y. S. (2019). Rapport building in suspect interviewing: A comparison of relationship-and procedure-based approaches in a laboratory setting. *Psychology, Public Policy, and Law*, 25(4), 253-265.
- Hudson, J. A., & Mayhew, E. M. Y. (2011). Children's temporal judgments for autobiographical past and future events. *Cognitive Development*, 26(4), 331-342.
- Hughes-Scholes, C. H., & Powell, M. B. (2008). An examination of the types of leading questions used by investigative interviewers of children. *Policing: An International Journal of Police Strategies & Management*, 31(2), 210-225.
- Hunt, L., & Bull, R. (2012). Differentiating genuine and false rape allegations: A model to aid rape investigations. *Psychiatry, Psychology and Law*, 19(5), 682-691.
- Innocence Project. (2015). *The causes of wrongful conviction*.
<https://innocenceproject.org/#causes>
- Irani F. (2011) Visual-Spatial Ability. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds). *Encyclopedia of Clinical Neuropsychology*. Springer.
- Iranzo, L. (2016). 'Unpacking the box': a novel tool to assess the development of working memory in children [Master's Thesis, University of Chester].

- Jack, F., Leov, J., & Zajac, R. (2014). Age-related differences in the free-recall accounts of child, adolescent, and adult witnesses. *Applied Cognitive Psychology, 28*(1), 30-38.
- Jiang, L., & Luo, D. (2016). Legal professionals' knowledge of eyewitness testimony in China: A cross-sectional survey. *Plos One, 11*(2), e0148116.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.
- Johnson, M., Magnussen, S., Thoresen, C., Lønnum, K., Burrell, L. V., & Melinder, A. (2015). Best practice recommendations still fail to result in action: A national 10-year follow-up study of investigative interviews in CSA cases. *Applied Cognitive Psychology, 29*(5), 661-668.
- Johnston, V., Brubacher, S. P., Powell, M., & Fuller-Tyszkiewicz, M. (2019). Patterns of nonverbal rapport behaviours across time in investigative interviews with children. *Journal of Nonverbal Behaviour, 43*(3), 411-434.
- Karni-Visel, Y., Hershkowitz, I., Lamb, M. E., & Blasbalg, U. (2019). Facilitating the expression of emotions by alleged victims of child abuse during investigative interviews using the Revised NICHD Protocol. *Child Maltreatment, 24*(3), 310-318.
- Karsna, K., & Kelly, L. (2021). *The scale and nature of child sexual abuse: Review of evidence*. Centre of expertise on child sexual abuse.
- Kask, K. (2011). Comparison of knowledge of law enforcement and lay people regarding eyewitness testimony. *Juridica International, 18*, 161-172.
- Kassin, S. M., Tubb, V. A., Hosch, H. M., & Memon, A. (2001). On the “general acceptance” of eyewitness research. A study of experts. *The American Psychologist, 56*, 405-416.
- Katus, T., Müller, M. M., & Eimer, M. (2015). Sustained maintenance of somatotopic information in brain regions recruited by tactile working memory. *Journal of Neuroscience, 35*(4), 1390-1395.

- Katz, C., & Hamama, L. (2013). "Draw me everything that happened to you": Exploring children's drawings of sexual abuse. *Children and Youth Services Review, 35*(5), 877-882.
- Katz, C., & Hershkowitz, I. (2010). The effects of drawing on children's accounts of sexual abuse. *Child Maltreatment, 15*(2), 171-179.
- Katz, C., & Hershkowitz, I. (2012). The effect of multipart prompts on children's testimonies in sexual abuse investigations. *Child Abuse and Neglect, 36*(11), 753-759.
- Kaufman, A. S. (2001). WAIS-III IQs, Horn's theory and generational changes from young adulthood to old age. *Intelligence, 29*(2), 131-167.
- Keatley, D. A., Barsky, A. D., & Clarke, D. D. (2017). Driving under the influence of alcohol: A sequence analysis approach. *Psychology, Crime & Law, 23*(2), 135-146.
- Kebbell, M. R., & Milne, R. (1998). Police officers' perceptions of eyewitness performance in forensic investigations. *The Journal of Social Psychology, 138*(3), 323-330.
- Kenny, M. C., & Wurtele, S. K. (2008). Toward prevention of childhood sexual abuse: Preschoolers' knowledge of genital body parts. In M. S. Plakhotnik, S. M. Nielsen (Eds.), *Proceedings of the seventh annual college of education research conference: Urban and international education section* (pp. 74-79). Florida International University.
- Klemfuss, J. Z. (2015). Differential contributions of language skills to children's episodic recall. *Journal of Cognition and Development, 16*(4), 608-620.
- Klika, J. B., & Conte, J. R. (Eds.). (2017). *The APSAC handbook on child maltreatment*. Sage Publications.
- Knowles, W., & Masidlover, M. (1982). *The Derbyshire Language Screen*.

- Knutsson, J., & Allwood, C. M. (2014). Opinions of legal professionals: Comparing child and adult witnesses' memory report capabilities. *The European Journal of Psychology Applied to Legal Context*, 6(2), 79-89.
- Korkman, J., Santtila, P., & Sandnabba, N. K. (2006). Dynamics of verbal interaction between interviewer and child in interviews with alleged victims of child sexual abuse. *Scandinavian Journal of Psychology*, 47(2), 109-119.
- Krackow, E., & Lynn, S. J. (2010). Event report training: An examination of the efficacy of a new intervention to improve children's eyewitness reports. *Applied Cognitive Psychology*, 24(6), 868-884.
- Krähenbühl, S. (2011). Effective and appropriate communication with children in legal proceedings according to lawyers and intermediaries. *Child Abuse Review*, 20(6), 407-420.
- Krähenbühl, S. (2019). Mock jurors' perceptions of a child witness: The impact of the presence and/or intervention of a registered intermediary during cross-examination. *Psychology, Crime and Law*, 25(7), 713-728.
- Krähenbühl, S., & Blades, M. (2006). The effect of question repetition within interviews on young children's eyewitness recall. *Journal of Experimental Child Psychology*, 94(1), 57-67.
- Krähenbühl, S., & Blades, M. (2009). Does the form of question repetition have an effect on children's recall accuracy and consistency? *International Journal of Police Science and Management*, 11(4), 460-475.
- Krähenbühl, S., Blades, M., & Westcott, H. (2010). 'What else should I say?' An analysis of the question repetition practiced in police interviews of 4-11-year-olds. *Police Practice and Research: An International Journal*, 11(6), 477-490.

- Lamb, M. E. (2016). Difficulties translating research on forensic interview practices to practitioners: Finding water, leading horses, but can we get them to drink? *American Psychologist, 71*(8), 710.
- Lamb, M. E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *Developmental Psychology, 24*(1), 215-234.
- Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradictions by children in the course of forensic interviews. *Applied Cognitive Psychology, 15*(5), 483-491.
- Lamb, M.E., Malloy, L.C., & La Rooy, D.J. (2011). Setting Realistic Expectations: Developmental Characteristics, Capacities and Limitations. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 15-48). Wiley-Blackwell.
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD investigative interview protocol. *Child Abuse & Neglect, 31*(11), 1201-1231.
- Lamb, M. E., Sternberg, K., J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child development, 71*(6), 1586-1596.
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews. *Journal of Consulting and Clinical Psychology, 71*(5), 926-934.

- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science, 6*(1), 35-41.
- La Rooy, D., Heydon, G., Korkman, J., & Myklebust, T. (2015). Interviewing child witnesses. In G. Oxburgh, T. Myklebust, T. Grant, & R. Milne (Eds.), *Communication in investigative and legal contexts* (pp. 57-68). John Wiley & Sons.
- La Rooy, D.J., Malloy, L.C. & Lamb, M.E. (2011). The Development of Memory in Childhood. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp. 49-68). Wiley-Blackwell.
- Lahtinen, H.M., Laitila, A., Korkman, J., & Ellonen, N. (2018). Children's disclosures of sexual abuse in population-based sample. *Child Abuse and Neglect, 76*, 84-94.
- Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences, 9*(2), 75-82.
- Lawson, M., & London, K. (2015). Tell me everything you discussed: Children's memory for dyadic conversations after a 1-week or a 3-week delay. *Behavioral Sciences & the Law, 33*(4), 429-445.
- Leahy-Harland, S., & Bull, R. (2017). Police strategies and suspect responses in real-life serious crime interviews. *Journal of Police and Criminal Psychology, 32*(2), 138-151.
- Leake, H. & Jeffels, K. (2017). *County Durham and Darlington Child Advocacy Centre Project: Baseline Assessment Project Report*.
- Lee, S. (2013). Explaining Variance in Children's Recall of a Stressful Experience: Influence of Cognitive and Emotional Individual Differences. *Korean Journal of Culture and Social Issues, 19*(3), 343-365.

- Lehman, E. B., McKinley, M. J., Thompson, D. W., Leonard, A. M., Liebman, J. I., & Rothrock, D. D. (2010). Long-term stability of young children's eyewitness accuracy, suggestibility and resistance to misinformation. *Journal of Applied Developmental Psychology, 31*(2), 145-154.
- Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*(4), 568-578.
- Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but not in visuospatial episodic memory. *Neuropsychology, 15*(2), 165-173.
- Loftus, E. F., Schooler, J. W., & Wagenaar, W. A. (1985). The fate of memory: Comment on McCloskey and Zaragoza. *Journal of Experimental Psychology, 114*(3), 375-380.
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy and Law, 11*(1), 194-226.
- London, K., Hall, A. K., & Lytle, N. E. (2017). Does it help, hurt, or something else? The effect of something else response alternative on children's performance on forced-choice questions. *Psychology, Public and Law, 23*(3), 281-289.
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source-monitoring. *Journal of Experimental Child Psychology, 52*(3), 297-318.
- Lum, J. A., Powell, M., & Snow, P. C. (2018). The influence of maltreatment history and out-of-home-care on children's language and social skills. *Child Abuse & Neglect, 76*, 65-74.
- Luther, K., Snook, B., Barron, T., & Lamb, M. E. (2015). Child interviewing practices in Canada: A box score from field observations. *Journal of Police and Criminal Psychology, 30*(3), 204-212.

- Lyon, T. D. (1998). New wave in children's suggestibility research: A critique. *Cornell Law Review*, *84*(4), 1004-1087.
- Lyon, T. D. (2012). Twenty-five years of interviewing research and practice: Dolls, diagrams, and the dynamics of abuse disclosure. *APSAC Advisor*, *24*(1-2), 14-19.
- Lyon, T. D. (2014). Interviewing children. *Annual Review of Law and Social Science*, *10*, 73-89.
- Lyon, T. D., & Ahern, E. C. (2011). Disclosure of child sexual abuse. In J. Myers (Ed.), *The APSAC handbook on child maltreatment* (3rd ed., pp. 233-252). Sage.
- Lytle, N.E., Dickinson, J. J., & Poole, D.A. (2019). Techniques for Interviewing Reluctant Child Witnesses. In J. J. Dickinson, N. S. Compo, R. N. Carol, B. L. Schwartz, & M. R. McCauley (Eds). *Evidence-Based Investigative Interviewing: Applying Cognitive Principles* (pp. 193-216). Routledge.
- Lytle, N., London, K., & Bruck, M. (2015). Young children's ability to use two-dimensional and three-dimensional symbols to show placements of body touches and hidden objects. *Journal of Experimental Child Psychology*, *134*, 30-42.
- Macleod, E., Gross, J., & Hayne, H. (2013). The clinical and forensic value of information that children report while drawing. *Applied Cognitive Psychology*, *27*(5), 564-573.
- Magnussen, S., & Melinder, A. (2012). What psychologists know and believe about memory: A survey of practitioners. *Applied Cognitive Psychology*, *26*(1), 54-60.
- Magnussen, S., Melinder, A., Stridbeck, U., & Raja, A. Q. (2010). Beliefs about factors affecting the reliability of eyewitness testimony: A comparison of judges, jurors and the general public. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *24*(1), 122-133.

- Malloy, L. C., Katz, C., Lamb, M. E., & Mugno, A. P. (2015). Children's requests for clarification in investigative interviews about suspected sexual abuse. *Applied Cognitive Psychology, 29*(3), 323-333.
- Marchant, R. (2013). How young is too young? The evidence of children under five in the English criminal justice system. *Child Abuse Review, 22*(6), 432-445.
- Marchant, R. (2016). Age is not determinative: The evidence of very young children in the English justice system. *Criminal Law and Justice Weekly, 180*(12-13).
- Marche, T. A., (1999). Memory strength affects reporting of misinformation. *Journal of Experimental Child Psychology, 73*(1), 45-71.
- Marr, C., Sauerland, M., Otgaar, H., Quaedflieg, C. W., & Hope, L. (2021). The effects of acute stress on eyewitness memory: An integrative review for eyewitness researchers. *Memory, 29*(8), 1091-1100.
- Matheson, I., & Hutchinson, N. (n.d.). <https://www.ldatschool.ca/working-memory-and-cognitive-load/>
- Matthias, C. R., & Zaal, F. N. (2011). Intermediaries for child witnesses: Old problems, new solutions and judicial differences in South Africa. *The International Journal of Children's Rights, 19*(2), 251-269.
- Mattison, M. (2015). *Using communication aids in the criminal justice system: Toolkit 14*. The Advocate's Gateway.
- Mattison, M. L., Dando, C. J., & Ormerod, T. C. (2015). Sketching to remember: Episodic free recall task support for child witnesses and victims with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(6), 1751-1765.
- Mattison, M. L., & Dando, C. J. (2020). Police officers' and Registered Intermediaries' use of drawing during investigative interviews with vulnerable witnesses. *Psychology, Crime & Law, 26*(2), 167-185.

- McCullough, A. (2017). *Examining police officers' perceptions of Achieving Best Evidence Language Screen for interviewing children* [Unpublished Dissertation, University of Portsmouth].
- McMahon, T. (2006). Teaching for more effective learning: Seven maxims for practice. *Radiography*, *12*(1), 34-44.
- Mehrani, M. B. (2011). What is biased? Children's strategies or the structure of yes/no questions? *First language*, *31*(2), 214-221.
- Mehrani, M. B., & Peterson, C. (2015). Recency tendency: Responses to forced-choice questions. *Applied Cognitive Psychology*, *29*(3), 418-424.
- Mehrani, M. B., & Peterson, C. (2017). Children's recency tendency: A cross-linguistic study of Persian, Kurdish and English. *First Language*, *37*(4), 360-367.
- Melinder, A., & Magnussen, S. (2015). Psychologists and psychiatrists serving as expert witnesses in court: what do they know about eyewitness memory? *Psychology, Crime & Law*, *21*(1), 53-61.
- Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental psychology*, *24*(4), 470-476.
- Memon, A., Meisser, C. A., & Fraser, J. (2010). The Cognitive Interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law*, *16*(4), 340-372.
- Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology*, *87*(3), 403-415.
- Memon, A., Wark, L., Bull, R., & Koehnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology*, *88*(2), 179-197.
- Merritt, K. A., Ornstein, P. A., & Spicker, B. (1994). Children's memory for salient medical procedure: Implications for testimony. *Pediatrics*, *94*(1), 17-23.

- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Wiley.
- Milne, R., & Bull, R. (2006). Interviewing victims of crime, including children and people with intellectual disabilities. In M. R. Kebbell & G. M. Davies (Eds.), *Practical Psychology for Forensic Investigations and Prosecutions* (pp. 7-24). Wiley.
- Ministry of Justice. (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures*.
- Ministry of Justice. (2015). *The Registered Intermediary Procedural Guidance Manual*.
- Ministry of Justice. (2020a). *The Registered Intermediary Procedural Guidance Manual 2020*.
- Ministry of Justice (2020b). *The Witness Intermediary Scheme: Annual Report 2019/20*.
- Morgan, K., Dorgan, K., & Hayne, H. (2013). Body maps do not facilitate older children's report of touch. *Scandinavian Journal of Psychology*, 54(1), 51-55.
- Moriguchi, Y., Okanda, M., & Itakura, S. (2008). Young children's yes bias: How does it relate to verbal ability, inhibitory control, and theory of mind? *First Language*, 28(4), 431-442.
- Morton, J., Hammersley, R. H., & Bekerian, D. A. (1985). Headed records: A model for memory and its failures. *Cognition*, 20(1), 1-23.
- Moston, S., & Engelberg, T. (1992). The effects of social support on children's eyewitness testimony. *Applied Cognitive Psychology*, 6(1), 61-75.
- Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect*, 20(7), 623-631.

- Myklebust, T., & Bjørklund, R. A. (2009). The child verbal competence effect in court: a comparative study of field investigative interviews of children in child sexual abuse cases. *Journal of Investigative Psychology and Offender Profiling*, 6(2), 117-128.
- Newlove, H. (2018). *A voice for the voiceless: The victims' commissioner's review into the provision for Registered Intermediaries for children and vulnerable victims and witnesses*. Victims' Commissioner.
- NHS. (n.d.). *Child Development: Ages and Stages*.
<https://www.cambscommunityservices.nhs.uk/what-we-do/children-young-people-health-services-cambridgeshire/specialist-services/childrens-speech-and-language-therapy/activities-ideas-and-info/child-development-ages-and-stages>
- Niveau, G. (2020). Sensory information in children's statements of sexual abuse. *Forensic Sciences Research*, 6(2), 97-102.
- NSPCC. (2021, March). *Statistic Briefing: Child Sexual Abuse*.
<https://learning.nspcc.org.uk/media/1710/statistics-briefing-child-sexual-abuse.pdf>
- Odegard, T. N., Cooper, C. M., Lampinen, J. M., Reyna, V. F., & Brainerd, C. J. (2009). Children's eyewitness memory for multiple Real-Life events. *Child Development*, 80(6), 1877-1890.
- Odinot, G., Wolters, G., & van Koppen, P. J. (2009). Eyewitness memory of a supermarket robbery: A case study of accuracy and confidence after 3 months. *Law and Human Behavior*, 33(6), 506-514.
- Ofen, N., Kao, Y. C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience*, 10(9), 1198-1205.

- Office for Criminal Justice Reform. (2007). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Using Special Measures*.
- Office of the Director of Public Prosecutions (ACT) and Australian Federal Police (2005). *Responding to Sexual Assault: The challenge of change*.
- Okanda, M., & Itakura, S. (2010). When do children exhibit a “yes” bias? *Child Development*, 81(2), 568-580.
- O'Mahony, B. M., Smith, K., & Milne, B. (2011). The early identification of vulnerable witnesses prior to an investigative interview. *British Journal of Forensic Practice*, 13(2), 114-123.
- Orbach, Y., & Lamb, M. E. (2000). Enhancing children's narratives in investigative interviews. *Child Abuse & Neglect*, 24(12), 1631-1648.
- Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect*, 25(3), 323-333.
- Orbach, Y., Shiloach, H., & Lamb, M. E. (2007). Reluctant disclosers of child sexual abuse. In M. E. Pipe, M. E. Lamb, Y. Orbach, & A. C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay, and denial* (pp. 115-134). Lawrence Erlbaum Associates Publishers.
- Ornstein, P.A. & Haden, C. A. (2002). The Development of Memory: Toward an Understanding of Children's Testimony. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.). *Memory and Suggestibility in the Forensic Interview* (pp. 29-62). Lawrence Erlbaum Associates, Inc.
- Ornstein, P. A., Merritt, K. A., Baker-Ward, L., Furtado, E., Gordon, B. N., & Principe, G. (1998). Children's knowledge, expectation, and long-term retention. *Applied*

Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 12(4), 387-405.

- Ornstein, P. A., Shapiro, L. R., Clubb, P. A., Follmer, A., & Baker-Ward, L. (1997). The influence of prior knowledge on children's memory for salient medical experiences. In N. L. Stein, P. A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp.83-112). Psychology Press.
- Otgaar, H., Horselenberg, R., van Kampen, R., & Lalleman, K. (2012). Clothed and unclothed human figure drawings lead to more correct and incorrect reports of touch in children. *Psychology, Crime & Law*, 18(7), 641-653.
- Otgaar, H., van Ansem, R., Pauw, C., & Horselenberg, R. (2016). Improving children's interviewing methods? The effects of drawing and practice on children's memories for an event. *Journal of Police and Criminal Psychology*, 31(4), 279-287.
- Oxburgh, G. E., Myklebust, T., & Grant, T. (2010). The question of question types in police interviews. *International Journal of Speech Language and the Law*, 17(1), 45-66.
- Owen, R. S. (2016). *Impact on law but what about practice? Intermediaries and how they aid vulnerable people to access the criminal justice system* [Master's Thesis, University of Chester].
- Pathman, T., Samson, Z., Dugas, K., Cabeza, R., & Bauer, P. J. (2011). A "snapshot" of declarative memory: Differing developmental trajectories in episodic and autobiographical memory. *Memory*, 19(8), 825-835.
- Patterson, T., & Hayne, H. (2011). Does drawing facilitate older children's reports of emotionally laden events? *Applied Cognitive Psychology*, 25(1), 119-126.
- Paulo, R. M., Albuquerque, P. B., & Bull, R. (2013). The enhanced cognitive interview: Towards a better use and understanding of this procedure. *International Journal of Police Science & Management*, 15(3), 190-199.

- Paulo, R. M., Albuquerque, P. B., Vitorino, F., & Bull, R. (2017). Enhancing the cognitive interview with an alternative procedure to witness-compatible questioning: Category clustering recall. *Psychology, Crime & Law*, 23(10), 967-982.
- Perry, N. W., McAuliff, B. D., Tam, P., Claycomb, L., Dostal, C., & Flanagan, C. (1995). When lawyers question children: Is justice served? *Law and Human Behavior*, 19(6), 609-629.
- Peters, W. W., & Nunez, N. (1999). Complex language and comprehension monitoring: Teaching child witnesses to recognise linguistic confusion. *Journal of Applied Psychology*, 84(5), 661-669.
- Peterson, C., & Grant, M. (2001). Forced-choice: Are forensic interviewers asking the right questions? *Canadian Journal of Behavioural Science*, 33(2), 118-127.
- Phillips, E., Oxburgh, G., Gavin, A., & Myklebust, T. (2012). Investigative interviews with victims of child sexual abuse: The relationship between question type and investigation relevant information. *Journal of Police and Criminal Psychology*, 27(1), 45-54.
- Pichler, A. S. (2018). *The relationship between investigative interview quality, trial process, and outcome in cases of child sexual abuse* [Doctoral Dissertation, Deakin University].
- Pipe, M. E., Orbach, Y., Lamb, M. E., Abbott, C. B., & Stewart, H. (2013). Do case outcomes change when investigative interviewing practices change? *Psychology, Public Policy, and Law*, 19(2), 179.
- Plastock, H. (2018). *"I deserved better than that": Survivors' decision making around legal disclosure of historic childhood sexual abuse: An interpretative phenomenological analysis and clinical research portfolio* [Doctoral Dissertation, University of Glasgow].

- Plotnikoff, J., & Woolfson, R. (2007). *The “go-between”*: Evaluation of intermediary pathfinder projects. NSPCC.
- Plotnikoff, J., & Woolfson, R. (2009). *Measuring up? Evaluating implementation of Government commitments to young witnesses in criminal proceedings*. NSPCC.
- Plotnikoff, J. & Woolfson, R. (2013). *‘Registered Intermediaries in action’ Messages for the CJS from the Witness Intermediary Scheme SmartSite*. Lexicon Limited.
- Plotnikoff, J., & Woolfson, R. (2015). *Intermediaries in the criminal justice system: Improving communication for vulnerable witnesses and defendants*. Policy Press.
- Plotnikoff, J., & Woolfson, R. (2019). *Falling short? A snapshot of young witness policy and practice*. NSPCC.
- Poole, D. A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review, 32*(3), 165-180.
- Poole, D. A., Bruck, M., & Pipe, M. (2011). Forensic interviewing aids: Do props help children answer questions about touching? *Current Directions in Psychological Science, 20*(1), 11-15.
- Poole, D. A., & Dickinson, J. J. (2011). Evidence supporting restrictions on uses of body diagrams in forensic interviews. *Child Abuse and Neglect, 35*(9), 659-669.
- Porter, M. L. (2018). From on the stand to on tape: Why recorded child victim testimony is safer, more effective, & fairer. *UC Davis Journal of Juvenile Law & Policy, 22*(1), 38-70.
- Powell, M. B. (2000). PRIDE: The essential elements of a forensic interview with an Aboriginal person. *Australian Psychologist, 35*(3), 186-192.
- Powell, M. B., Bowden, P., & Mattison, M. (2015). Stakeholders' perceptions of the benefit of introducing an Australian intermediary system for vulnerable witnesses. *Australian & New Zealand Journal of Criminology, 48*(4), 498-512.

- Powell, M. B., Fisher, R. P., & Wright, R. (2005). Investigative interviewing. In N. Brewer & K. Williams (Eds.). *Psychology and law: An empirical perspective* (pp. 11–42). Guilford Press.
- Powell, M. B., & Guadagno, B. (2008). An examination of the limitations in investigative interviewers' use of open-ended questions. *Psychiatry, Psychology and Law, 15*(3), 382-395.
- Powell, M. B., Roberts, K. P., Ceci, S. J., & Hembrooke, H. (1999). The effects of repeated experience on children's suggestibility. *Developmental Psychology, 35*(6), 1462-1477.
- Powell, M. B., Wright, R., & Clark, S. (2010). Improving the competency of police officers in conducting investigative interviews with children. *Police Practice and Research, 11*(3), 211-226.
- Price, P. C., Jhangiani, R., & Chiang, I. A. (2015). *Research Methods in Psychology – 2nd Canadian Edition*. BCcampus.
- Price, H. L., & Roberts, K. P. (2011). The effects of an intensive training and feedback program on police and social workers' investigative interviews of children. *Canadian Journal of Behavioural Science, 43*(3), 235-244.
- Price, H. L., Roberts, K. P., & Collins, A. (2013). The quality of children's allegations of abuse in investigative interviews containing practice narratives. *Journal of Applied Research in Memory and Cognition, 2*(1), 1-6.
- Quas, J. A., & Lench, H. C. (2007). Arousal at encoding, arousal at retrieval, interviewer support, and children's memory for a mild stressor. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21*(3), 289-305.
- Quas, J. A., Rush, E. B., Yim, I. S., & Nikolayev, M. (2014). Effects of stress on memory in children and adolescents: Testing causal connections. *Memory, 22*(6), 616-632.

- Quas, J. A., Stolzenberg, S. N., & Lyon, T. D. (2018). The effects of promising to tell the truth, the putative confession, and recall and recognition questions on maltreated and non-maltreated children's disclosure of a minor transgression. *Journal of Experimental Child Psychology, 166*, 266-279.
- R. v. Barker (2010) EWCA Crim 4.
- Raskin, D. C., & Esplin, P. W. (1991). Statement validity assessment: Interviewing procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment, 13*(3), 256–291.
- Raven, J. (2008). *Raven's Coloured Progressive Matrices*. Pearson.
- Renault, A. G., Auvray, M., Parseihian, G., Miall, R. C., Cole, J., & Sarlegna, F. R. (2018). Does proprioception influence human spatial cognition? A study on individuals with massive deafferentation. *Frontiers in psychology, 9*.
- Renfrew, C. E. (1997). *Action Picture Test* (4th ed.). Winslow Press.
- Ridley, A. M., Clifford, B. R., & Keogh, E. (2002). The effects of state anxiety on the suggestibility and accuracy of child eyewitnesses. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 16*(5), 547-558.
- Ridley, A. M., Van Rheede, V., & Wilcock, R. (2015). Interviews, intermediaries and interventions: Mock-jurors', police officers' and barristers' perceptions of a child witness interview. *Investigative interviewing: Research and practice, 7*(1), 21-35.
- Riggs, K. J., Jolley, R. P., & Simpson, A. (2013). The role of inhibitory control in the development of human figure drawing in young children. *Journal of Experimental Child Psychology, 114*(4), 537-542.
- Rispens, I. & van der Sleen, J. (2017). Interviewing witnesses in the Netherlands. In D. Walsh, G. E. Oxburgh, A. D. Redlich, & T. Mykleburst (Eds.), *International*

- developments and practices in investigative interviewing and interrogation. Volume 1: Victims and witnesses* (pp.174-188). Routledge.
- Roberts, K. P., Brubacher, S. P., Powell, M. B., & Price, H. L. (2011). Practice narratives. In M. E. Lamb, D. J. La Rooy, L. C. Malloy, & C. Katz (Eds.), *Children's Testimony: A Handbook of Psychological Research and Forensic Practice* (pp.129-145). Wiley-Blackwell.
- Roberts, K. P., Lamb, M. E., & Sternberg, K. J. (2004). The effects of rapport-building style on children's reports of a staged event. *Applied Cognitive Psychology, 18*(2), 189-202.
- Rocha, E. M., Marche, T. A., & Briere, J. L. (2013). The effect of forced-choice questions on children's suggestibility: A comparison of multiple-choice and yes/no questions. *Canadian Journal of Behavioural Science, 45*(1), 1-11.
- Roebers, C. M., & Schneider, W. (2001). Individual differences in children's eyewitness recall: The influence of intelligence and shyness. *Applied Developmental Science, 5*(1), 9-20.
- Roeper, T. (2007). *The prism of grammar: How child language illuminates humanism*. MIT Press.
- Rohwer, M., Kloos, D., & Perner, J. (2012). Escape from meta ignorance: How children develop an understanding of their own lack of knowledge. *Child development, 83*(6), 1869-1883.
- Salmon, K., Pipe, M., Malloy, A., & Mackay, K. (2012). Do non-verbal aids increase the effectiveness of 'best practice' verbal interview techniques? An experimental study. *Applied Cognitive Psychology, 26*(3), 370-380.
- Santtila, P., Korkman, J., & Sandnabba, K. N. (2004). Effects of interview phase, repeated interviewing, presence of a support person, and anatomically detailed dolls on child sexual abuse interviews. *Psychology, Crime & Law, 10*(1), 21-35.

- Sauerland, M., Brackmann, N., & Otgaar, H. (2018). Rapport little effect on children's, adolescents', and adults' statement quality, accuracy and suggestibility. *Journal of Child Custody, 15*(4), 268-285.
- Saywitz, K. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology, 59*(5), 682-691.
- Saywitz, K. J., Larson, R. P., Hobbs, S. D., & Wells, C. R. (2015). Developing rapport with children in forensic interviews: Systematic review of experimental research. *Behavioral Sciences & the Law, 33*(4), 372-389.
- Saywitz, K. J., Lyon, T. D., & Goodman, G. S. (2011). Interviewing children. In J. E. B. Myers (Ed.), *The APSAC Handbook on Child Maltreatment* (pp. 337-360). Sage.
- Saywitz, K. J., Snyder, L., & Nathanson, R. (1999). Facilitating the communicative competence of the child witness. *Applied Developmental Science, 3*(1), 58-68.
- Saywitz, K. J., Wells, C. R., Larson, R. P., & Hobbs, S. D. (2016). Effects of interviewer support on Children's memory and suggestibility: Systematic review and meta-analyses of experimental research. *Trauma, Violence, & Abuse, 20*(1), 22-39.
- Seale-Carlisle, T. M., Wetmore, S. A., Flowe, H. D., & Mickes, L. (2019). Designing police lineups to maximize memory performance. *Journal of Experimental Psychology: Applied, 25*(3), 410-430.
- Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: a meta-analysis and integrative review. *Psychological Bulletin, 143*(6), 636-675.
- Smethurst, A., & Collins, K. (2019). Mock Jury Perceptions of Vulnerable Defendants Assisted in Court by Intermediaries - Are Jurors' Expectations Violated? *Applied Psychology in Criminal Justice, 15*(1), 23-40.

- Smith, P. K., Cowie, H., & Blades, M. (2015). *Understanding children's development*. John Wiley & Sons.
- Smith, K., & Milne, R. (2017). Vulnerable witnesses: The investigation stage. In H. Norton, & P. Cooper (Eds.). *Vulnerable people and the criminal justice system: A guide to law and practice* (pp. 23-55). Oxford University Press.
- Snow, P. C., Powell, M. B., & Murfett, R. (2009). Getting the story from child witnesses: Exploring the application of a story grammar framework. *Psychology, Crime and Law*, 15(6), 555-568.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93(24), 13515-13522.
- Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, 8(3), 205-211.
- Stakic, D., & Ilic, Z. (2018). Help me tell you what really happened: Forensic interview with children victims and witnesses. In D. Simovic (Ed.), *International Scientific Conference "Archibald Reiss Days" Thematic Conference Proceedings of International Significance* (pp. 3-14). Academy of Criminalistic and Police Studies.
- Stein, N., & Glenn, G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 32-120). Ablex.
- Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The memorandum of good practice: Theory versus application. *Child Abuse & Neglect*, 25(5), 669-681.
- Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., & Sunshine, N. (1996). The relation between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, 17(3), 439-451.

- Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology, 86*(5), 997-1005.
- Stolzenberg, S. N., McWilliams, K., & Lyon, T. D. (2017). Spatial language, question type, and young children's ability to describe clothing: Legal and developmental implications. *Law and Human Behavior, 41*(4), 398-409.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662.
- St-Yves, M., Griffiths, A., Cyr, M., Gabbert, F., Carmans, M., Sellie, C., Bruneau, G., & Powell, M. (2014). Training in investigative interviewing: Observations and challenges. In M. St-Yves (Ed.), *Investigative interviewing – the essentials* (pp. 245-282). Carswell.
- Taggart, J. (2018). *Defendant intermediaries in the criminal justice system: Is Northern Ireland leading the way?* Policy Briefing Paper No. 30.
<https://dx.doi.org/10.2139/ssrn.3219987>
- Taggart, J. (2021). 'I am not beholden to anyone... I consider myself to be an officer of the court': A comparison of the intermediary role in England and Wales and Northern Ireland. *The International Journal of Evidence & Proof, 25*(2), 141-162.
- Tang, C. M., Dickey, S., & Samuelson, D. (2017). Young children's reports of when events occurred: Do event type and assessment method matter? *Infant and child development, 26*(1), e1963.
- Teoh, Y., & Lamb, M. E. (2010). Preparing children for investigative interviews: Rapport-building, instruction, and evaluation. *Applied Developmental Science, 14*(3), 154-163.

- Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1*(4), 285-293.
- Tomasello, M. (2003). *Constructing a good language: A usage-based theory of language acquisition*. Harvard University Press.
- Tomasello, M. (2010). *Origins of human connection*. Harvard University Press.
- Triangle. (2015). *Unpacking the Box*.
- Triangle. (n.d.). *Welcome!* Retrieved December 9, 2021, from <https://triangle.org.uk/>
- Van der Linden, M. (1998). The relationships between working memory and long-term memory. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie, 321*(2-3), 175-177.
- Verkampt, F., Dodier, O., Milne, R., & Ginet, M. (2021). An analysis of the quality of French investigative interviews with children: Age of the witnesses does matter. *Police Practice and Research: An International Journal, 22*(2), 1130-1154..
- Vrij, A., Hope, L., & Fisher, R.P. (2014). Eliciting Reliable Information in Investigative Interviews. *Policy Insights from the Behavioral and Brain Sciences, 1*(1), 129-136.
- Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities, 35*(1), 62-74.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children, 23*(3), 34-41.
- Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology, 27*(1), 73-84.

- Warren, A. R., Woodall, C. E., Thomas, M., Nunno, M., Keeney, J. M., Larson, S. M., & Stadfeld, J. A. (1999). Assessing the effectiveness of a training program for interviewing child witnesses. *Applied Developmental Science, 3*(2), 128-135.
- Waschi, N. A. (2017). *Seeing reason: Visuospatial ability, sex differences and the Raven's Progressive Matrices* [PhD Thesis, University of Adelaide].
- Waterhouse, G. F., Ridley, A. M., Bull, R., La Rooy, D. J., & Wilcock, R. (2019). Mapping repeated interviews. *Journal of Police and Criminal Psychology, 34*(4), 392-409.
- Waterman, A. H., & Blades, M. (2011). Helping children correctly say "I don't know" to unanswerable questions. *Journal of Experimental Psychology: Applied, 17*(4), 396-405.
- Waterman, A. H., Blades, M., & Spencer, C. P. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology, 18*(2), 211-225.
- Waterman, A. H., Blades, M., & Spencer, C. P. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's 'don't know' responses. *British Journal of Developmental Psychology, 22*(3), 335-348.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523-541.
- Wechsler, D. (1981). *WAIS-R manual: Wechsler Adult Intelligence Scale-Revised*. Psychological Corporation.
- Wechsler, D., & Zhou, X. (2011). *Wechsler abbreviated scale of intelligence (2nd ed.)*. Pearson.
- Westcott, H. L., & Kyan, S. (2004). The application of a 'story-telling' framework to investigate interviews for suspected child sexual abuse. *Legal and Criminological Psychology, 9*(1), 37-56.

- Westcott, H. L., & Kynan, S. (2006). Interviewer practice in investigative interviews for suspected child sexual abuse. *Psychology, Crime & Law, 12*(4), 367-382.
- Westcott, H. L., Kynan, S., & Few, C. (2006). Improving the quality of investigative interviews for suspected child abuse: A case study. *Psychology, Crime & Law, 12*(1), 77-96.
- Wheeler, R. L., & Gabbert, F. (2017). Using self-generated cues to facilitate recall: A narrative review. *Frontiers in psychology, 8*.
- Whiting, B. F., & Price, H. L. (2017). Practice narratives enhance children's memory reports. *Psychology, Crime & Law, 23*(8), 730-747.
- Willcock, E., Morgan, K., & Hayne, H. (2006). Body maps do not facilitate children's reports of touch. *Applied Cognitive Psychology, 20*(5), 607-615.
- Wilcock, R., Crane, L., Hobson, Z., Nash, G., Kirke-Smith, M., & Henry, L. A. (2018). Supporting child witnesses during identification lineups: Exploring the effectiveness of registered intermediaries. *Applied Cognitive Psychology, 32*(3), 367-375.
- Wise, R. A., & Safer, M. A. (2004). What US judges know and believe about eyewitness testimony. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 18*(4), 427-443.
- Wise, R. A., & Safer, M. A. (2010). A comparison of what US judges and students know and believe about eyewitness testimony. *Journal of Applied Social Psychology, 40*(6), 1400-1422.
- Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest, 18*, 10-65.
- Wolfman, M., Brown, D., & Jose, P. (2016). Taking stock: Evaluating the conduct of forensic interviews with children in New Zealand. *Psychology, Crime & Law, 22*(6), 581-598.

- Wolfman, M., Brown, D., & Jose, P. (2018). The use of visual aids in forensic interviews with children. *Journal of Applied Research in Memory and Cognition*, 7(4), 587-596.
- Woolford, J., Patterson, T., Macleod, E., Hobbs, L., & Hayne, H. (2015). Drawing helps children to talk about their presenting problems during a mental health assessment. *Clinical Child Psychology and Psychiatry*, 20(1), 68-83.
- Woolnough, P. S., & MacLeod, M. D. (2001). Watching the birdie watching you: Eyewitness memory for actions using CCTV recordings of actual crimes. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 15(4), 395-411.
- Wright, R., & Powell, M. B. (2006). Investigative interviewers' perceptions of their difficulty in adhering to open-ended questions with child witnesses. *International Journal of Police Science & Management*, 8(4), 316-325.
- Yarbrough, J., Herve, H. F., & Harms, R. (2013). The sins of interviewing: Errors made by investigative interviewers and suggestions for redress. In B. S. Cooper, D. Griesel & M. Ternes (Eds.), *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment* (pp. 59-98). Springer.
- Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *The Journal of Applied Psychology*, 71(2), 291-301.
- Yuille, J. C., Hunter, R., Joffe, R., & Zaparnuik, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child Victims, child witnesses: Understanding and improving testimony* (pp. 95-115). Guildford Press.
- Yi, M., & Lamb, M. E. (2018). How accurately do police officers identify the types of questions used in investigative interviews with child victims? *The Korean Journal of Forensic Psychology*, 9(3), 117-135.

Yi, S. L. B., Powell, M. B., & Guadagno, B. (2014). The association between investigative interviewers' knowledge of question type and adherence to best-practice interviewing. *Legal and Criminological Psychology, 19*(2), 270-281.

Appendices

Appendix A – Empirical Evidence Supporting Questionnaire Items

Questionnaire Item	Empirical Evidence / Supporting Research / Item Utilised in Previous Study
Multiple Witnesses	See Brubacher, Peterson, et al. (2019) for a discussion.
Effect of Post-Event Information	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004).
Minor Details	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004).
Confidence-Accuracy	Adapted from an item included in Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004). Supported by the findings of Kassin et al. (2001).
Impact of Stress	Christiansen and Hubinette (1993), Odinot et al. (2009), Woolnough and MacLeod, (2001), Yuille and Cutshall (1986).
Attitudes and expectations	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004). Supported by the findings of Kassin et al. (2001).
Weapon Focus	Item included in Dodier, Tomas, et al. (2019), adapted from an item included in Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004).
Forgetting Curve	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2012), Melinder and Magnussen (2015), Wise and Safer (2004).

Questionnaire Item	Empirical Evidence / Supporting Research / Item Utilised in Previous Study
Children's Recall	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2011), Melinder and Magnussen (2015).
Dramatic Events	Item included in Dodier, Tomas, et al. (2019), Magnussen and Melinder (2012), Melinder and Magnussen (2015).
Immediate Acceptance of Suggested Information	Adapted from an item included in Dodier, Tomas, et al. (2019).
Response Bias	Research has found that young children can possess a 'yes' bias (e.g., Fritzley & Lee, 2003; Fritzley et al., 2013). See section 2.2.2 of this thesis for a more in-depth discussion.
False Memories	This is referred to as a reverse developmental trend (Calado et al., 2018). A review by Brainard and Reyna (2012) found that false memories were more prevalent amongst older children. See section 2.1.2 of this thesis for a more in-depth discussion.
Anatomical Dolls	Research has found that the use of dolls can lead to false reports of anal and vaginal touching (Bruck et al., 2000).
Communication Aids	ABE (MoJ, 2011) states that children find it easier to understand the symbolic nature of drawing. Research has shown that drawing does not appear to compromise children's accuracy (e.g., Barlow et al., 2011; Patterson & Hayne, 2011). See section 2.2.4 of this thesis for a more in-depth discussion.
Question Types	"... this type of question [open-ended] should be used predominantly during the interview... This questioning style also minimises the risk that the interviewer will impose their view of what happened on the witness." (MoJ, 2011, p. 78).

Questionnaire Item	Empirical Evidence / Supporting Research / Item Utilised in Previous Study
Open-Ended Questions	The question is an example of an indirect speech act (Evans et al., 2014). See section 2.2.2 of this thesis for a more in-depth discussion.
Recall vs. Recognition Memory	See section 2.2.2 of this thesis for a more in-depth discussion.
Repeating Questions	“Specific-closed questions should not be repeated ‘word for word’ because the witness may feel that their first answer was incorrect and change their response accordingly.” (MoJ, 2011, p. 79).
Children’s Narratives	See Brubacher, Peterson, et al. (2019) for a discussion.

Appendix B – Reflection

Throughout the PhD process I have reflected upon my practice. Given my previous employment, as an NRI, I have had to be aware and address any unconscious biases I may have regarding the role. To safeguard against potential criticism and to prevent these biases impacting my results I sought to conduct quantitative research, with clearly defined parameters, which offer as little room for interpretation as possible. I also had two other researchers inter-rate my results. For study two, the inter-rater was blind to the condition (i.e., assessment vs. colouring vs. no assessment) thus any biases they may have possessed were negated. As for study four, long discussions were had with my supervisors (one of whom also acted as an inter-rater) as to how best to match and subsequently code the interviews. This is reflected in the relatively high levels of agreement between raters. Overall, I do not feel my previous employment has, in any way, influenced the findings of this PhD thesis. Without this experience I do not feel that the PhD thesis would have been possible. The research required an in-depth knowledge of communication assessments and understanding of the RI role which I would have otherwise not possessed.

Appendix C - Rule Cards



If I get it wrong, tell me



I don't know



I don't understand

Appendix D - Interview Protocol

Note: The interview is semi-structured. Children may not be asked all of the questions.

Whether a child is asked a question will be dependent upon their previous response.

Interviewer Questions (Introductions)

Notes

Hi (child's name). Yesterday you helped me with some jobs. Today I would like us to do some talking is that ok? Thank you, yesterday we had some talking rules (Shows child each rule card) What's this one? And this one? And this one? And what can you use this one for? (practises rule with the child)

OR

Hi (child's name). Yesterday we did some colouring. Today I would like us to do some talking, is that ok? When we do talking we have some rules (explains and practises each rule with the child).

OR

Hi (child's name). My name is Alex and today I would like us to do some talking, is that ok? When we do talking we have some rules (explains and practises each rule with the child).

We have one more rule. Let me show you this video clip. You need to pay close attention to it as I'm going to ask you some questions about it (Shows child truth and lies clip). What happened? What did the girl say? What did the boy say? Did he tell a truth, a lie or don't you know? What should he have said? It is really important that in this room we only tell the truth. Is that ok?

Interviewer Questions (Substantive Phase)

Notes

1. Thank you. I need your help. Last week Mrs Science came to your school. But I was late so I didn't see her. Can you tell me what happened?

2. Then what happened?

3. Tell me more?

Interviewer Questions (substantive phase)	Notes
4. Tell me about Mrs Science.	
5. What was Mrs Science wearing?	
6. What colour were Mrs Science's socks? (Note: child did not see socks)	
7. I heard you played (or you said you played) some games with Mrs Science. Did you play one game, more than one game, or don't you know?	
8. Tell me about the first game you played?	
9. What did you need to play the game?	
10. Did you need a bottle?	
11. What were its contents?	
12. I think I have played this game before. You also need a plate, don't you?	
13. Now I want to talk about another game. I found some glitter in the classroom. Did you play a game with glitter?	
14. Tell me about the game?	
15. To prevent your hands from getting dirty, what did Mrs Science ask you to do before you played with the glitter?	
16. Did you put gloves on?	
17. Then what happened?	
18. When you were playing the game, did Mrs Science ask you to shut your eyes? Why did she ask you to do that?	

Interviewer Questions (substantive phase)	Notes
19. Did you have to find where Mrs Science had put the glitter?	
20. What did you use to help you look for the glitter?	
21. What did it look like?	
22. Did you find the glitter? Where did you find it?	
23. Was the glitter green or blue? (Note: it was pink)	
24. Tell me about the last game you played?	
25. Did Mrs Science put something on her hands?	
26. What did Mrs Science put on her hands?	
27. What did it look like?	
28. Was it transparent or translucent?	
29. Mrs Science had a light that was red, did she not?	
30. What did Mrs Science do with the light?	
31. Then what happened?	
32. Did Mrs Science wash her hands?	
33. How many times did Mrs Science wash her hands?	
34. How long did you play the games with Mrs Science?	
35. Where did you play the games?	
36. Did you play the games in the same place or different places?	
37. Describe the layout of the classroom	

Interviewer Questions (substantive phase)	Notes
38. If I was stood at the door of the classroom when you were playing the games with Mrs Science, what would I have been able to see? Could I have seen you?	
39. Who was there when you played the games with Mrs Science?	
40. Before you played the games with Mrs Science, did (name of assistant) tell you a rule?	
41. What was the rule?	
42. The rule was that Mrs Science should not put her skin on your skin. Why was that?	
43. Did Mrs Science touch your skin?	
44. Tell me what happened when Mrs Science touched your skin?	
45. I would like you to draw a picture of yourself.	
46. Thank you. Please make this (gingerbread man) look like you.	
47. I want to make sure that I understand what happened. Show me on here (gingerbread man) where Mrs Science touched you? Has that part got a name?	
OR	
Did Mrs Science touch you?"	
48. Tell me about when Mrs Science touched you?	
49. Was the touching intentional?	
50. How do you know?	
51. What did Mrs Science do when she touched your skin?	

Interviewer Questions (substantive phase)

Notes

52. You said Mrs Science touched you here. Tell me about that? (Note: interviewer indicates to wrong place)

Appendix E – Empirical Evidence Supporting Predictions

Age-Based Predictions

Ground Rules

1) Child does not employ ground rules during interview.	2) Child employs ground rules on one occasion during interview.	3) Child employs ground rules on two or more occasions during interview.
	4-year-olds 5-year-olds 6-year-olds	7-year-olds 8-year-olds 9-year-olds
	Ground rules have been shown to be effective with young children (e.g., 4-year-olds, Krackow & Lynn, 2011; 5-year-olds, Waterman & Blades, 2011). Thought to be dependent on the development of cognitive skills – knowledge access and false belief (see section 2.2.3 of this thesis). These skills typically develop between the ages of 4 and 6 years old.	Children’s ability to apply the ground rules correctly during an interview improves with development level (Brown et al., 2019).

Attention

0-4 minutes	5-9 minutes	10-14 minutes	15-19 minutes	20 + minutes
		4-year-olds	5-year-olds 6-year-olds	7-year-olds 8-year-olds 9-year-olds

The general standard is that children are able to attend for 3 to 5 minutes per year of age (Schmitt, 1999 as cited in Anderson et al., 2009). From professional judgement I opted to estimate attention at the more conservative end of this range.

Responsiveness

1) Child responds with a single word or phrase.	2) Child responds with a full sentence.	3) Child gives an extensive narrative (multiple sentences).
4-year-olds 5-year-olds	6-year-olds 7-year-olds	8-year-olds 9-year-olds
Triangle's developmental milestones: <ul style="list-style-type: none"> - 4-year-olds mean length of utterance is 3.75-4.50 words. - 5-year-olds mean length of utterance is 4.50+ words. 	Informativeness has been found to increase with age (Hershowitz et al., 2012; Lamb et al., 2000; Lamb et al., 2003). Developed knowledge access (i.e., an understanding of how individuals acquire knowledge, Wellman & Liu, 2004) – provide more information.	

Resistance to Suggestion / Compliance

1) Child does not acquiesce to suggestion.	2) Child acquiesces to a misleading suggestion once during the interview.	3) Child acquiesces to a misleading suggestion on two or more occasions during the interview.
	8-year-olds 9-year-olds	4-year-olds 5-year-olds 6-year-olds 7-year-olds
	Reverse developmental trends (see Calado et al., 2018) demonstrate how older children may also cede to incorrect suggestions.	False belief typically develops between 4 to 6 (Wellman & Liu, 2004). If this is not fully developed the child may have difficulty correcting the interviewer. Younger children often have weaker memory traces (see section 2.1.1 of thesis). Younger children are more suggestible than older children (Ceci & Bruck, 1993).

Drawing

1) Unusable (i.e., cannot distinguish body parts)	2) Usable (i.e., can distinguish body parts)
4-year-olds	5-year-olds 6-year-olds 7-year-olds 8-year-olds 9-year-olds
At around 3 years old children tend to represent people as “tadpoles” with a large head on a small body with extended arms (Riggs et al., 2013).	Children can produce “conventional” human figures from around the age of 5 (Riggs et al., 2013).
1) Able to locate body parts	2) Unable to locate body parts
4-year-olds 5-year-olds 6-year-olds 7-year-olds 8-year-olds 9-year-olds	
Children are thought to develop representational insight around 3 years old (ABE; MOJ, 2011).	

Appendix F – Date of Real-World Interviews

Year	Frequency
2011	1
2012	4
2014	1
2017	1
2018	16
2019	6
2020	12
2021	1
