Examining the efficacy of Joint Investigative Interviewing Training in Scotland through analysis of the quality of trainee interviewers' role-play interviews with adult actors

A thesis submitted for the degree of Doctor of Philosophy (PhD)

by

Annabelle Nicol

School of Social and Health Sciences
Abertay University.

April, 2016

Declaration

Candidate's declarations:

I, Annabelle Nicol, hereby certify that this thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy (PhD), Abertay University, is wholly my own work unless otherwise referenced or acknowledged. This work has not been submitted for any other qualification at any other academic institution.

Signed…………………………………………………
Date………………………………………………………………………..

Supervisor's declaration:

I, Christopher Watkins hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy (PhD) in Abertay University and that the candidate is qualified to submit this thesis in application for that degree.

Signed [Principal Supervisors signature]…………………………………...
Date………………………………………………………………………..

Certificate of Approval

I certify that this is a true and accurate version of the thesis approved by the examiners, and that all relevant ordinance regulations have been fulfilled.

Supervisor……………………………………………………….………
Date………………………………………………………………………..

Dedication

To Katie Anne Macleod (aka Granny)

&

Anne Nicol who somehow anticipated this journey even before I did

Acknowledgements

First and foremost, my sincere thanks go to my first supervisor Dr David La Rooy. It is certain that without his continual support, both academically and personally, this work could not have been completed, you have my eternal gratitude. Thank you for believing in me when I did not believe in myself. Thank you to my second supervisor, Professor Fiona Gabbert, for the academic feedback so sensitively (and hilariously) provided that it allowed me to laugh at myself and grow from my many mistakes and for the pastoral care and ever positive encouragement. Thanks also go to Professor Michael Lamb for providing me with academic advice, feedback and many professional development opportunities, Dr. Christopher Watkins for stepping in as an internal supervisor at short notice and for providing thorough and prompt feedback, and to Sam Andrews and Charlene Flynn for reliability coding. Thank you to all of the family members that supported me, special thanks go to my Granny, whose encouragement and never wavering support made this all possible, also Aunty Carol (a future interviewer herself I hope!), Aunty Marion for their interest and encouragement, and Uncle Donnie (and Granny again) for buying me the laptop that this thesis was written on. Steven Geddes, words cannot explain, you know that this would not have been possible without you, you turned everything around. My final thanks go to two wonderful ladies- Naomi Hodges who sadly passed away before the completion on this thesis but set me on the right track and Fiona Sturrock who kept me there. I will never forget any of these contributions- thank you all!

General abstract

Best-practice guidelines, grounded in psychological theory, are important for interviewers who engage with a vulnerable group such as children who allege abuse (reviewed in Chapters 1-6). Prior to conducting interviews, Scottish investigative interviewers must complete a JIIT (Joint Investigative Interviewing Training) course, where they interview an adult actor who plays the role of an abused child. Limited field work suggests child interviews in Scotland are of low quality. No empirical work on the quality of JIIT precludes an understanding of whether i) interviewers are trained to an appropriate standard, ii) gains made in training diminish over time (i.e. when compared against performance in the field) and iii) whether adults from acting companies respond in a manner that reinforces best practice questioning by the trainee.

This thesis examines the line of questioning and responses of actors during JIIT. Specifically, whether JIIT outcomes differ i) in two different police jurisdictions with separate untrained actors (Chapter 7), ii) according to whether training sessions are scribed or recorded (Chapter 8) and iii) according to an actor's expertise in the psychological literature on investigative interviewing (Chapter 9). Across studies, trainees were poor at using both 'ground rules' and closure rules. Lines of questioning did not differ between two forces and untrained actors/trainees responded-to/used invitations to the same extent as directives. Untrained actors provided *more informative* responses to focussed prompts and suggestive questions than they did to open prompts. Although a 'trained' adult actor provided a more authentic training opportunity, trainees were not sensitive to this feedback and it did not encourage better lines of questioning. Scribed versus recorded accounts of interviews captured fewer ground rules and closure rules, shorter responses to invitations, incomplete and incorrect records in response to invitations and, on average, omitted two details per interview.

By way of comparison, analyses of actual Scottish field interviews with children (Chapter 10) revealed no use of any ground/closure rules at levels greater than chance. Children provided more details in response to option-posing questions and relatively short but detailed responses to suggestive questions, only in comparison to 'other' questions in both instances. Of note in relation to earlier chapters, children gave longer average responses to invitations than directives. Overall the findings of this thesis indicate that interviewers' training is inconsistent across jurisdictions in Scotland and that many aspects of the training are inappropriate. For example, the adult actors hired to role-play abused children do not reinforce the use of best practice open prompts selectively as we would hope and the use of scribing as a method to record interviews is unreliable and resulted in both loss and incorrect recording of information. Further, the field interviews conducted with children in Scotland are of a low standard (6% open prompts), therefore, practical recommendations are made (Chapter 11) in light of these concerning findings.

Table of Contents

List of Figures

List of Tables

Chapter 1: The short-term effects of child abuse and maltreatment

Child abuse can have short- and long-term effects on children's development, functioning, and well-being (Maguire, et al., 2015; Mullen, Martin, Anderson, Romans & Herbison, 1996; Paolucci, Genuis & Violato, 2001). In order to effectively support maltreated children, it is of the utmost importance that professionals interacting with suspected child abuse victims are aware of these effects. Having a clear understanding of the potential difficulties that follow childhood maltreatment enables child protection workers to prepare strategies to counteract these problems. As the focus of this thesis is on obtaining accounts from children about their experiences, this chapter will focus on research findings of the short-term effects that abuse can have on children (i.e., those that present during childhood). The effects reviewed are divided into three developmental domains; social, emotional and cognitive, because skills in each of these domains contribute to children's ability to engage with interviewers and to recall and describe potentially upsetting experiences. Finally, the relationship between these experiences and children's performance in a forensic interview will be discussed, including strategies to counter these issues.

1.1 Defining and researching the effects of child abuse

Child abuse takes place in many forms; the most common four variants are physical abuse, sexual abuse, neglect and emotional abuse. Child abuse definitions vary according to legal contexts and child

14

protection guidelines between and within countries (for example, at the state-level in the Unites States of America) although most definitions converge on the same principles. The following definitions of child abuse were taken from the Scottish Government's National Child Protection Guidelines as this thesis was interested in practices and procedures of child protection workers in Scotland (see the "National Guidance for Child Protection in Scotland" (2014) for full details). *Physical abuse* involves causing physical harm to children for example by hitting, shaking, burning or suffocating. *Sexual abuse* involves forcing or enticing children to take part in sexual activities, including penetrative or non-penetrative touching, and non-contact activities such as looking at or being involved in producing pornography. *Child neglect* is failing to meet children's basic physical needs (e.g. for food, shelter and clothing) and/or psychological and emotional needs (e.g., for nurturing and stimulation)*. Emotional abuse* involves adversely affecting children emotionally, for example by conveying to children that they are worthless or unloved, by causing them to feel frightened or in danger, exploiting or corrupting them or imposing age - or developmentally-inappropriate expectations on them. Often children are victims of more than one type of abuse (Briere & Runtz, 1990; Finkelhor, Ormrod, & Turner, 2007; Radford, Corral, Bradley & Fisher, 2013).

Typically studies that have looked at the effects of child abuse have been retrospective studies of adults that experienced abuse in childhood (e.g., Aspelmeier, Elliott & Smith, 2007; Briere & Runtz, 1990; Cantón-Cortés, Cortés & Cantón, 2015; Frías, Brassard & Shaver, 2014;

Oshri, Sutton, Clay-Warner & Miller, 2015; Riggs & Kaminski, 2010; Styron & Janoff-Bulman, 1997; Unger & De Luca, 2014). Attributing difficulties in adulthood to childhood abuse is problematic as subsequent life experiences occurring between the abuse and time of reporting make it difficult to attribute sole cause to the abuse (Oshri, et al., 2015; Woon & Hedges, 2008). For example, a meta-analysis by Woon and Hedges (2008) showed that children with maltreatment-related Post Traumatic Stress Disorder (PTSD) did not exhibit the same brain changes (reduced hippocampal volume) as adults with maltreatment-related PTSD. While this delay in the appearance of symptomology could be attributed to a slow onset, it is likely that such changes are brought on by factors occurring in the time period between the abusive experience and the exhibition of symptoms in adulthood. For example, substance abuse is also related to reduced hippocampal volume (De Bellis et al., 2000) and adolescents and adults maltreated in childhood are at a higher risk of substance abuse than their non-maltreated counterparts (Clark, Thatcher & Maisto, 2004). Therefore, due to the difficulty in attributing sole cause of symptomology in adulthood to childhood maltreatment, the main focus of this chapter is on the short-term effects of child abuse that are already apparent in childhood.

There are also challenges in studying abused children themselves. For example, when selecting non-abused control groups to compare a particular behaviour or performance measure to a group of abused children, some researchers have noted that due to the secrecy involved in child abuse there is the potential that children assigned to the non-

16

abused group have in fact been abused (Bruck, Ceci, Francouer & Renick, 1995a; Chae, Goodman, Eisen & Qin; 2011). Further, often the abused groups of children involved in research are recruited from either clinical e.g. following PTSD related hospitalisation (Deblinger, McLeer , Atkins, Ralphe & Foa, 1989) or forensic (Chae, et al., 2011; Eisen, Qin, Goodman & Davis, 2002) settings which are themselves associated with particular child and case characteristics and may not be representative of abused children in general. Many abused children present no symptomology following abuse (e.g., Mannarino and Cohen (1986) reported 31% of the sexually abused children in their study were symptom free), therefore, clinical samples may represent children severely affected by abuse and ignore children that do not present with clinical symptoms. Further, the children in the Chae et al. (2011) and Eisen et al. (2002) studies were recruited during investigations into allegations that they had been abused. If sexual abuse, physical abuse or neglect was substantiated then the children were eligible to participate in the corresponding abused subgroup, if the allegations were unsubstantiated then the children were eligible to participate in the non-abused control group. Therefore, even though abuse of the control group was not substantiated there was still strong enough suspicion that abuse had occurred to warrant investigation, this factor alone may mean that the control group were more similar to the abused group than a non-abused group (Chamberland, Lacharité, Clément & Lessard, 2015). It is also entirely possible that some of the children may have in fact been abused but withheld from disclosing.

17

The dynamics of child abuse vary depending on a host of factors including, but not limited to, victim age, gender, type of abuse experienced, duration and frequency of abuse, individual personality characteristics, and the child's relationship to the perpetrator (Blanchard-Dallaire & Hébert, 2014; Manly, Cicchetti & Barnett, 1994). This means that no sample of abused children can ever be representative of the full range of circumstances involved in abuse, therefore, finding representative samples and appropriately matched comparison groups is an arduous task. These differing dynamics of individual cases of child abuse and the fact that children often experience multiple types of abuse (Briere & Runtz, 1990), add an additional challenge to attributing cause and effect to the study of the effects of child abuse.

Despite the difficulties of conducting research in this area, some well replicated findings on the effects of child abuse on children's short-term social, emotional and cognitive development have been found and they will be the focus of the following review. For this literature search, the author's familiarity with the literature, and literature cited by Dante Cicchetti, Gail Goodman and Seth Pollak were used as starting points. I then performed searches using Web of Science to find more recent and additional studies, based on the search term "physical abuse" in combination with "social development " (Chapter 1.2), "emotional development" (Chapter 1.3), and "cognitive development" (Chapter 1.4). This search process was repeated three times replacing "physical abuse" with "sexual abuse", "emotional abuse" and "neglect" in combination with the three developmental domains.

18

1.2 The relationship between childhood abuse and social development

Forming attachments with other people, such as family members, friends and romantic partners is important to well-being (Armsden & Greenberg, 1987; Greenberg, Siegel & Leitch, 1983). Attachment theory (Bowlby, 1958; 1988, see also Dozier & Bernard, 2015, for a more recent discussion) states that the very first relationship an infant forms in life is with his or her primary caregiver. When the caregiver is warm and responsive to the infant's needs they will develop a secure attachment to the adult that allows them to explore the world safe in the knowledge that comfort and security will be provided should they become frightened, fatigued or sick. If the caregiver fails to meet the child's needs by responding inconsistently or inappropriately, the child may develop an insecure attachment style. Ainsworth, Blehar, Waters and Wall (1978) classified such insecure attachment styles as being either an avoidant attachment style or an anxious-ambivalent attachment style, depending on the nature of the interactions with the caregiver. Avoidant attachments result from the caregiver rejecting or rebutting the child, causing the child to develop a coping strategy of reluctance to get close to the caregiver in order to protect themselves from further rejection (Karen, 1994). Inconsistent parenting where the caregiver is sometimes responsive and sometimes unresponsive result in anxious/ambivalent attachments, the child cannot predict how they will be treated leading to anxiety, even preceding separation, and inconsistent responses veering between anger and helplessness upon the caregiver's return (Karen, 1994). Abused children may perceive an abusive caregiver's behaviour as inconsistent

and/or rejecting depending on the nature of the abuse, which prevents abused children from forming secure attachments with caregivers. This in turn can result in disruption of the formation of close attachments with others both in the short-term and over the course of abused children's lives.

The specific dynamics of early relationships with caregivers can shape how close relationships are formed later in life, with attachment styles during childhood with parents reflecting attachment styles with romantic partners in adulthood (Bartholomew & Horowitz, 1991; Collins & Read, 1990; (experiment 3) Gleeson & Fitzgerald, 2014; Hazan & Shaver, 1987; Papaleontiou Louca & Kyriakides, 2016). Adults' perception of the quality of their relationships with both parents and their parents' relationship with each other were the best predictors of attachment style with their most significant romantic partner to date (Hazan & Shaver, 1987). Collins and Read (1990) found that participants' descriptions of their opposite-sex parent's attachment style predicted their own attachment style in their current romantic relationship, e.g. women that rated their fathers as warm and responsive (secure) were more likely to date men that reported they were comfortable with getting close to people (secure). While these findings relate to the long-term effects of childhood relationships, they also make an important contribution to the knowledge base on the short-term effects of child abuse because they provide insight into when and how relationship patterns begin in childhood.

20

These studies were also able to investigate how attachment processes beginning in childhood are maintained through questioning adults who have greater insight into their relationships and thought processes. Hazen and Shaver (1987) found that adults with avoidant attachment styles were more likely than adults with other attachment types to agree with statements such as "It is rare you can find someone you can really fall in love with", while anxiously attached were more likely to agree with "Few people are as willing and able as I am to commit themselves to a long-term relationship" than avoidant or securely attached participants. Similarly, Collins and Read (1990) found that participants who reported their parental style as warm and responsive (secure) were less likely to be anxious about being abandoned, while those that had experienced ambivalent parental style in childhood were less likely to report being able to depend on others in adulthood. These studies provide support for the idea that during childhood models of the self, others and relationships are formed through the dynamics of relationships with parents. Parents provide a model for behaviour that children internalise, during adulthood they seek relationships that mimic the previously experienced dynamics, thus both meeting their expectations of what relationships should be like based on their past experiences and reinforcing their attachment style (Sroufe, Egeland, Carlson & Collins, 2009). Therefore, it is likely that the effects of being abused or maltreated as a child have effects on attachment behaviour and relationship quality in both the long- and short-term following abuse.

21

Support for the early appearance of insecure attachment styles in maltreated children have been obtained through raters' objective observations of maltreated children as young as 12 months of age (Egeland & Sroufe, 1981). Mother-child dyads were recruited through a health centre that provided mother and infant health care. Within these dyads, a maltreatment group consisted of pairs in which the mother had been assessed as having either seriously abused or neglected the child and the comparison group consisted of mother-child dyads in which the mother had been assessed as providing excellent care for her child. Using Ainsworth's strange situation procedure (Ainsworth et al., 1978) the infant's behaviour is observed in a novel environment, in the presence of their mother, during a separation from her and upon her return. Securely attached infants typically explore the new environment in the presence of the mother, react with distress at her departure and seek comfort from her on her return. Avoidant infants ignore their mother on her return, even when distressed, while anxious infants are unwilling to explore the new environment even with their mother present and react to her return following separation by either resisting contact or failing to be comforted by her attempts on returning. Maltreated children were more likely to be rated during observations as having an insecure attachment style (68%) compared to their non-maltreated peers (25%).

Finzi, Cohen, Sapir and Weizman (2000) attributed attachment style to groups of physically abused, neglected and a control group of non-abused/non-neglected 6-to-13-year-old children using their answers to self-report attachment questions. Their findings were consistent with

22

the attachment styles reported by adults from such categories (e.g. Unger & De Luca, 2014). For example, the physically abused children had the highest frequency of avoidant attachment styles (avoidant attachment styles are commonly found in adults that report being physically abused as children), the neglected children were mostly anxiously attached and the non-maltreated children were predominantly securely attached. Collectively, research has shown that insecure attachments are more likely in maltreated children, are already evident during childhood, and continue into adulthood.

As previously discussed, adults with insecure attachment styles have internal models that view other people as less dependable and less available (Collins & Read, 1990). There is evidence that maltreated children also hold internal models or views that other people will be unsupportive to the needs of others (Macfie et al., 1999). Macfie et al. (1999) demonstrated the presence of such expectations in maltreated 3- to 5-year-olds. Abused (either sexual, physical or both), neglected and non-abused children completed story stems that involved attributing reactions to the parent and peers of a child in distress (e.g. after burning their finger). Maltreated children (abused and neglected) depicted fewer parental and fewer peer responses that attempted to alleviate the distress of the protagonist than did their non-maltreated counterparts. When the maltreated groups responses were analysed separately both abused and neglected children displayed fewer parental helping responses than the non-maltreated children. However, the neglected children also displayed fewer peer attempts at alleviating distress than did non-abused or abused

23

children. Even in childhood, maltreated children develop internal models of others as being unresponsive to their needs, demonstrated by their expectations that the needs of the children in the stories would be ignored.

Maltreated children have been shown to have poor social skills and difficulty in making and maintaining friendships. Maltreated children were rated as less prosocial, more withdrawn and more disruptive/aggressive by their peers and camp counsellors (Alink, Cicchetti, Kim & Rogosch, 2012). Dodge, Pettit and Bates (1994) followed physically maltreated children from kindergarten to fourth grade, each year their peers, mother and teachers rated them more disliked, less popular and more socially withdrawn than their non-maltreated counterparts with this gap increasing over time. Kim and Cicchetti (2010) in a longitudinal study of 6- to 12-year-old children showed that neglect and emotional maltreatment were predictive of peer rejection as the children got older (emotional maltreatment particularly so). Bolger and Patterson (2001) found chronically maltreated children were at greater risk of being rejected by their peers, and Manly, et al., (1994) found that physically abused children aged 5-to 11-years-old were rated as having more behavioural problems than non-maltreated by their camp counsellors and peers. These findings show that internalised beliefs relating to a lack of social support can often turn out to be true.

The RS (rejection sensitivity) model of peer rejection proposed by Levy, Ayduk and Downey (2001), posits an explanation of how this

24

occurs. Rejection by valued others causes the rejected child to expect further rejection by others and to develop a hyper-vigilance to rejection, interpreting others' actions (both true rejecting and non-rejecting) as rejecting. Belief that people will be rejecting and unsupportive to the needs and emotions of the maltreated child, set off a self-fulfilling prophecy in that anxiety about being rejected leads to behaviours in social situations, such as avoidance, that undermine their social relationships, leading to the feared outcome of rejection (London, Downey, Bonica & Paltin, 2007; Purdie & Downey, 2000). Such social avoidance behaviours have been observed in abused 1-to 3-years-olds, who displayed avoidance (turning away, creeping away) towards attachment figures three times more often and to their peers four times more often than the control group of non-abused toddlers. They also appeared to become even more avoidant and inhibit approaching the caregivers when the caregivers attempted to initiate friendly approaches, such as offering them a toy (George & Main, 1979).

The authors suggest that the avoidant strategies used by the abused toddlers shift their attention back to maintaining proximity, enabling them to avoid negative emotions, such as fear or anger, which could result in emotional behaviours that in turn could affect the behaviour of the person attending to them. By avoiding interactions with others, abused children can prevent engaging in situations where behaviour expression could lead to further abuse. Given that the caregivers in this study were not the children's abusers, it is apparent that insecure attachment styles manifest cross-situationally, encompassing

25

not only the child's primary relationship, but also their interactions with other caregivers and peers. Caution must be exercised when interpreting the findings from this study due to its extremely small sample size (20 children in total). Presently there is not enough research to say for certain whether the model applies to younger children, though a larger more recent study by Valentino, Cicchetti, Toth & Rogosch (2011) assessed children's play dynamics and reported children from abusing families engaged in less child-initiated play than did children from neglecting and non-maltreating families.

The findings relating to the effects of sexual abuse on children's social development are less concrete. Sexually abused children have been reported to experience significant challenges in their social interactions with peers (Ozbaran et al., 2009; Daignault & Hébert, 2008; Hébert, Langevin & Daignault, 2016).For example, teachers rated 8-to 12-year-old sexually abused children as having a significantly higher number of social problems, being more withdrawn and less assertive than non-abused children. The sexually abused children themselves were also more likely than non-abused children to report reduced general interpersonal trust in people (Blanchard-Dallaire and Hébert (2014). However, although non-significant, sexually abused children tended to report a higher level of trust in their peers than the non-abused children. This is consistent with prior research finding that sexually abused children may turn to their peers for support (Feiring, Tasks & Lewis, 1998).

26

Further complicating the field, comparisons between sexually abused children and other subgroups of maltreated children have also produced mixed findings. Teisl, Rogosch, Oshri & Cicchetti (2012) found that maltreated children overall were more likely than non-maltreated children to be rated as dominant bullies by camp counsellors and their peers, and that the sexually abused children were just as likely as all other maltreated children (physically abused, emotionally abused and neglected) to be rated as such. In contrast, other studies have reported that sexually abused children exhibit enhanced social competence compared to physically abused and neglected children (Manly et al., 1994). It is not possible at this time to say for certain whether or not sexually abused children experience an overall deficit in social functioning compared to non-maltreated children or how their social skills compare to those of children that have experienced other forms of abuse. The literature on sexually abused children's social development and functioning is extremely sparse, particularly for the youngest children (Hébert, Langevin & Bernier, 2013), though some research suggests there is the potential for some degree of social resilience. The greater level of trust abused children place in their peers may enable them to seek social support from friends while the reduced interpersonal trust they report may be specifically towards adults, as a result of being sexually abused by one which may be taken as a form of betrayal (Finkelhor & Browne, 1985). More research in this area is needed before any firm conclusions can be drawn and it is likely that study methodology e.g. whether comparison groups are non-maltreated or children that have

27

experienced other forms of maltreatment, will affects the degree of social problems attributed to sexually abused children.

  1.2.1. Summary. Abused and maltreated children are more likely to have social difficulties, particularly in forming secure attachments to others, compared with their non-abused counterparts (Egeland & Sroufe, 1981; Finzi et al., 2000). Abuse status makes unique contributions to the types of social difficulties faced by maltreated individuals. Physical abuse may be interpreted as a form of rejection and predicts avoidant attachment styles (Finzi et al., 2000) and high rejection-sensitivity (Levy et al., 2001), characterised by a focus on maintaining detachment from others to avoid future rejection that persists even in the face of friendly overtures (George & Main, 1979). Neglected children are likely to exhibit an anxious-ambivalent style of attachment due to caregiver's inconsistent attitudes towards them (Finzi et al., 2000). As a result they have low expectations of people, expecting less support from their peers than do other maltreated subtypes (Macfie et al., 1999). A lack of consistent behavioural responses and appropriate feedback to their behaviour deprives them of the opportunity to learn appropriate social responses, and therefore puts them at a high risk of being rejected by their peers (Dadds & Salmon, 2003; Shipman, Edwards, Brown, Swisher & Jennings, 2005). The findings relating to sexually abused individuals are unclear, some research suggests, in childhood at least, they do not exhibit the same elevated levels of social difficulties that accompany other types of abuse (Manly et al., 1994). This may be due to the type of abuse, which may not be perceived as a rejection, but rather as a betrayal by a specific

28

person or persons (Finkelhor & Browne, 1985), allowing for better social integration with non-abusers than experienced by physically abused and neglected children (Blanchard-Dallaire & Hébert, 2014; Feiring, Tasks & Lewis, 1998).

Future research is needed to clarify the individual factors that exacerbate and alleviate the social dysfunction that results from experiencing child abuse, particularly within the different subtypes, in an effort to develop interventions for children. For example, social skills training may benefit neglected children that have not learned from parents how to interact with others, while more research on the social effects of child sexual abuse may pinpoint the factors that lead to social competence in some, but not other, sexually abused children. Further, educating caregivers and professionals that interact with abused children on a regular basis about how to form secure attachments with abused children will provide a secure base for abused children, allowing them to adapt their mental models of the self and others' in order to break the cycle of insecure attachment early in life before they repeat these attachment patterns in adulthood.

1.3   *The relationship between childhood abuse and emotional development*

The ability to recognise emotions and to understand the causes and consequences of emotions, along with situation-appropriate responding, together make up the multifaceted skill of emotional understanding (Parke, Cassidy, Burks, Carson, & Boyum, 1992). Disrupted emotional development, evidenced by dysregulated emotional

29

patterns in response to emotional arousal, is more likely to present in maltreated children than non-maltreated children (Maughan & Cicchetti, 2002). Deficits in the ability to regulate one's emotional arousal and subsequent behavioural responses can manifest as either extreme emotional control and lack of a behavioural response or under-controlled emotions and disorganised behaviour in response to arousal (Cummings, 1987). Emotional dysregulation has detrimental effects on the psychological adjustment (Cook, Greenberg, & Kusche, 1994) and emotional well-being of the abused individual as well as negatively impacting their social relationships with others (Maughan & Cicchetti, 2002; Rogosch, Cicchetti & Aber, 1995).

At the individual level, abused children are more likely to have lower self-esteem (Kim & Cicchetti, 2006) and to experience negative emotions (Shields & Cicchetti, 1998; Shipman et al., 2005) than their peers. The inability to regulate negative emotions and low self-esteem may play a part in abused children's greater likelihood of depression and anxiety (Hébert, Langevin & Bernier, 2013; Kim & Cicchetti, 2006; Maguire, et al., 2015; Maughan & Cicchetti, 2002; Shipman, Zeman, Penza & Champion, 2000; Valentino, Cicchetti, Rogosch, & Toth, 2008), and PTSD (McLeer, Callaghan, Henry & Wallen, 1994; Porter, Lawson & Bigler, 2005), than their non-abused peers. Finally, depressive symptoms and abuse (physical and sexual) predict repeated suicide attempts in 7- to 17-year-olds (Nelson, Faust, Doyle & Delucia, 2015).

30

As well as experiencing more negative emotion and mood disorders, abused children's emotional dysfunction causes social problems because emotional understanding is related to peer acceptance (Denham, McKinley, Couchoud & Holt, 1990). Even being at risk of being abused by a parent has been shown to negatively impact children's socio-emotional development (Chamberland, Lacharité, Clément & Lessard, 2015). Maltreated toddlers have been shown to react to peer distress (crying, screaming) with anger and withdrawal, and were more likely to cause the distress of their peers (physical pain, teasing) than non-abused children were (Klimes-Dougan & Kistner, 1990). Maltreated children are also more likely than non-maltreated children to act aggressively in both in situations of perceived threat (Cummings, Hennessy, Rabideau & Cicchetti, 1994) and in innocuous play activities with peers and teachers (Shields & Cicchetti, 1998). Sexually abused pre-schoolers were assessed as being less competent at regulating their emotions according to multiple informants compared to their non-abused counterparts, with sexually abused boys performing the poorest (Langevin, Cossette, & Hébert, 2016).Physically and sexually abused children are also more likely to show dissociative tendencies such as daydreaming and blank staring during play (Shields & Cicchetti, 1998). Abused children's inability to regulate their emotions and behaviour has social consequences. Aggressive behaviour and causing peer distress lead to poor peer relations and reacting with situation-inappropriate emotional such as withdrawing and dissociating from peers can serve to further isolate abused children socially (Dodge et al., 1994).

31

Such inappropriate behavioural responses may stem from the fact that abused children have difficulty recognising emotions in others. Maltreated children may not get the chance to learn to recognise emotions from parents as inconsistent and harsh parenting makes it difficult for children to make sense of their parents' emotions (Dadds & Salmon, 2003). Pollak, Cicchetti, Hornung and Reed (2000) read 3-to 5-year-old neglected, physically abused and control children story vignettes where the main character experienced happiness, sadness, disgust, fear or anger. The children were asked to choose from three pictures of a model expressing an emotion (one correct and two distractors) which emotion the character was feeling. The non-maltreated correctly attributed a greater percentage of emotions than did the two abused groups, with neglected children correctly recognising the fewest appropriate expressions. They also found that abuse group status was associated with problems identifying specific emotions; the physically abused were poorer than control children at recognising sadness and evidenced a bias toward choosing angry faces, while neglected children had trouble recognising anger and were more likely to choose sad faces.

A review of facial emotion processing and recognition in maltreated children found that maltreated children tended to exhibit reduced accuracy in facial tasks and showed greater reactivity, response bias, and electrophysiological activation of specific brain areas in response to faces expressing negative emotions, especially anger (Da Silva Ferreira, Crippa & De Lima Osorio, 2014). Physically abused children have been consistently shown in the literature to have a

32

propensity towards attributing others' emotions to anger. Pollak and

Kistler (2002), for example, created a continuum of faces expressing

different emotions, the faces then blended into the next emotion in

varying degrees, and asked children what emotions the faces displayed.

Physically abused 9-year-olds were more likely to attribute anger to faces

that displayed fear and sadness than the control group. In the case of

genuine anger, physically abused 8 to 10-year-olds detected this emotion

much more quickly and at a lower frequency than non-abused children

when viewing model faces depicting various emotions on computer

software that began with highly degraded images that became more

organised at regular intervals (Pollak and Sinha, 2002).

Physically maltreated children's hyper vigilance to hostile cues

may be an adaption brought about as a protective factor. Frequent

exposure to aggression and violence may increase the potential benefits

of recognising the initial signals of anger in others in order to be able to

flee or appease a situation that could cause them harm. However, this

heightened sensitivity to anger also may be accompanied by a reduced

ability to process neutral and positive emotions (Pollak & Tolley-Schell,

2003; Young & Wisdom, 2014).Unlike physically abused children,

sexually abused and neglected children may not be exposed to the same

levels of aggression, however, the particular emotions associated with

neglectful and sexually abusive environments may also lead to different

types of response biases. Pollak et al., (2000) found neglected children

had a response bias to sad faces and a mixed maltreatment group

33

(neglect, physical and sexual abuse) recognised fear quicker than control groups (Masten et al., 2008). Sexually abused and neglected children were both less able than controls to recognise positive emotions (Fishbein et al., 2009; Young & Wisdom, 2014) with the neglected children being the poorest at discriminating all emotions, possibly due to their neglectful environment prohibiting them from experiencing a range of emotions they could use to learn how to discriminate emotional states. However, some studies report no effect of sexual abuse on children's accuracy in matching facial expression to the correct emotion (e.g. Sullivan, Kirkpatrick & MacDonald, 1995).

If abused children have difficulty recognising emotions in others then they will have trouble knowing what the appropriate emotional response is and may choose a situation-inappropriate response. Abused children have indeed been shown to have difficulty understanding the cause and consequence of emotions (Dodge, Pettit, Bates & Valente, 1995; Edwards, Shipman, Brown, 2005). Meta analyses of the literature suggest that childhood physical abuse and neglect has a negative effect on measures of understanding others' emotions (d=-1.35; see Luke & Banerjee, 2013 for a meta-analytic review). Indeed, the extent of the effects of maltreatment on emotional understanding is moderated by age such that the effects are stronger for young people (aged 2 to 6) compared to adolescents, where the null effect in the latter group is almost zero. Dodge et al. (1995) showed kindergarteners cartoons of children's social interactions (e.g. a child's request to join a game being

34

refused) and asked children to explain why the peer(s) might have acted the way they did to the protagonist in the story, with explanation options ranging from benign to hostile. Children that had been physically abused chose more hostile attributional biases as to why they thought the peer(s) behaved in such a manner. When given three behavioural response options; aggressive (respond with violence), passive (cry) or assertive (direct verbal appeals) and to rate each option on a scale from being 'very good' – 'very bad'; the abused children selected more aggressive responses and evaluated the outcomes of aggressive behaviour more positively than did non-abused children. Choosing aggressive responses and evaluating them positively predicted later externalising outcomes in grades 4 through 5, with the abused group exhibiting significantly more conduct problems than non-abused group.

Weiss, Dodge, Bates, and Pettit, (1992) found physical punishment was associated with increased hostile attributional biases of others' hypothetical behaviour, increased tendency to generate aggressive hypothetical responses and to positively evaluate the likely outcomes of aggressing. Parents' and teachers' ratings of the children's aggression were also related to harsh discipline. Physically abused children perceive others' actions to be hostile, choose hostile reactive behaviours and evaluate these reactive behaviours positively. This in turn leads to aggressive behaviour in real-world situations. Edwards et al., (2005) also found deficits in understanding the causes and consequences of emotions in neglected children aged 5- to 12-years-old. Neglected and

non-neglected children's emotional understanding was assessed using an emotional understanding interview that involved viewing pictures of children exhibiting sadness and anger and then answering questions about the causes and consequences of expressing these emotions. Neglected children had lower emotional understanding scores than their peers. They also showed the children vignettes involving children depicting anger and sadness, and asked how their mother would react if they were to show these emotions, in order to measure expectation of maternal support, answers were coded as either supportive or non-supportive. Neglected children expected less maternal support to displays of emotion than their non-neglected counterparts, particularly in response to anger. These predictions turned out to be accurate, as the mother and child were subsequently prompted to engage in discussion of instances the child had been angry and sad the neglecting mothers provided less emotional support than the non-neglecting mothers.

Shipman et al., (2005), also used the emotional understanding interview with neglected and control children aged 6- to 12-years-old. The neglected children demonstrated lower understanding of both causes and consequences of emotion and appropriate responses to others' emotions, expected less maternal support and more conflict to displays of emotion and said they would be less likely to show emotion to their mothers. Mothers of neglected children reported lower levels of empathy in their children while the neglected children themselves confirmed this, reporting that they would feel happy in response to the negative display of others (e.g., "I'd feel happy that it wasn't me that was mad.") while non-

maltreated children indicated they would attempt to provide assistance or support to the other person. These findings indicate that neglected children are less aware of the causes and consequences of emotions than non-maltreated children, which may, in part, explain some of their inappropriate responses and lack of empathy to others' distress. In addition, the expectation of a lack of maternal support, which was subsequently confirmed when observing parent-child interactions, means neglected children have not learned how to appropriately support others emotionally which explains their avoidance of others in distress compared to their non-maltreated peers. Inability to interpret others' emotions, inhibiting one's own situation-appropriate emotional responses and not providing emotional support to peers may contribute to social difficulties that neglected children experience as a result of their early neglect.

Like the neglected children, Shipman et al., (2000) found sexually maltreated 6-to12-year-olds also achieved lower emotional understanding scores than non-maltreated controls when asked questions inquiring about what could cause specific emotions and the consequences of the emotions. With relation to the consequences of emotional displays, maltreated girls expected less emotional support and more conflict from both parents and from peers if they were to express anger and less support and more conflict from parents if they were to express sadness. They also reported inhibiting their anger more often and parental ratings of the abused children rated these children as less empathic than ratings of the control children by their parents of the controls. Neglected, sexually abused and physically abused children show deficits in understanding the

causes and consequences of emotions (e.g. Edwards et al., 2005; Shipman et al., 2000). Sexually abused and neglected children however expect less emotional support in response to negative emotions and report consequently inhibiting them, while physically abused children rate aggressive behaviour positively and are more likely to be rated as behaviourally aggressive than control children (Dodge et al. 1995; Shipman et al., 2000, 2005; Weiss et al., 1992).

1.3.1 Summary. Emotional development is affected in a variety of ways by experiencing child maltreatment. Maltreated children have difficulty recognising emotions compared to their non-maltreated peers (Fishbein et al., 2009; Pollak et al., 2000; Young & Wisdom, 2014). Physically abused children displayed response biases to angry faces (Pollak & Kistler, 2002) and detected anger at a lower frequency than controls (Pollak & Sinha, 2002). Physically abused children also have difficulty understanding the causes of emotions attributing innocuous behaviour as hostile (Dodge et al., 1995). Past experiences of inter and/or intra-personal violence and/or aggressive behaviour, may increase vigilance toward threatening cues as a means of self-protection from future hostility (Teisl & Cicchetti, 2008), but the stress experienced by maltreated children may become maladaptive, increasing them to perceive hostility and threat in situations where there is none. Physical maltreatment also affects emotion regulation and behaviour, with physically maltreated children more likely to exhibit reactive aggression (Dodge et al., 1995; Weiss et al., 1992) and inappropriate aggressive behaviour, for example, in response to a crying peer (Klimes-Dougan &

38

Kistner, 1990) or rating aggressive behaviour positively (Dodge et al., 1995; Weiss et al., 1992).

Unlike physically abused children, neglected and sexually abused children report inhibiting negative emotions such as anger and sadness as they expect less parental support and more conflict in response to such emotional displays (Edwards et al., 2005; Shipman et al., 2005; Shipman et al., 2000). This has often been the case for maltreated children, who find that emotion displays do not help achieve their immediate goals, and such children inhibit expressing anger and sadness more than a non-maltreated control group (Shipman & Zeman, 2001). Dissociation has been shown to be a common coping strategy in sexually abused children (Shields & Cicchetti, 1998). While dissociating provides protection from painful emotions associated with abuse, it reduces awareness of one's own emotions and may become a lifelong maladaptive coping behaviour (Rodriguez-Srednicki, 2002). Some also show sexually abused children have trouble recognising emotions (Young & Wisdom, 2014), but other studies have not replicated this (Sullivan et al., 1995). Neglected children may have had the least exposure to emotional learning opportunities due to their parental neglect (Dadds & Salmon, 2003) this may explain why they perform the poorest in tasks of recognising emotions (Pollak et al., 2000). They display response biases to sad faces (Pollak et al., 2000) indicating that instead of the hypervigilance to hostility they may be more attune to sadness, which may explain the increased risk for neglected children to become depressed and suffer from low self-esteem (Lynch & Cicchetti, 1998).

Future research should focus on specific childhood interventions for emotional dysregulation. Sexually abused and neglected children suffer from internalising symptoms such as depression and anxiety and inhibit expressing negative emotions; interventions should focus on helping children express negative emotions in an appropriate manner and on enhancing self-esteem. As physically abused children interpret innocuous behaviour as threatening (Dodge et al. 1995; Weiss et al., 1992), cognitive interventions that help them generate alternative explanations for other's behaviour may reduce their perceptions of threat and the subsequent aggression that follows such attributions. Adults that have been abused as children evidence maladaptive coping strategies and greater levels of psychopathology than non-abused adults, therapeutic interventions during childhood that teach children emotional coping strategies and break the cycle of using maladaptive coping strategies, can reducing the likelihood of later risk taking behaviours, symptomology and maladaptive coping strategies in adulthood. Children that evince externalising behaviours, such as violence and aggression, should be taught to manage anger as aggression can hamper goal achievement and impact their social relationships. Many abused children have not learned the causes and consequences of emotions and developing methods of educating such populations on emotion recognition and emotional understanding should be a priority.

1.4   The relationship between childhood abuse and cognitive development

In addition to abused children's socio-emotional functioning, children's cognitive performance has also been of interest to researchers (Barahal, Waterman, & Martin, 1981; Mills et al., 2011). Cognitive processes or mental abilities, also known as executive functions, include, but are not limited to, memory, reasoning, problem solving, decision making and language. Such abilities are important in academic environments and abused children have been shown to evince poorer academic performance than control groups (Ayoub et al., 2006; Hostinar, Stellern, Schaefer, Carlson & Gunnar, 2012; Kinard, 1999, 2001; Mills et al., 2011). Kinard (1999), for example, measured 7- to 11-year-old abused (physical, sexual and neglected) and non-abused children's academic achievement using a test that encompassed reading, writing and arithmetic, at two testing sessions spaced one year apart. Mean scores for the non-abused group were comparable to national norms while the abused group scored at least 10 points below the non-abused group on all three measures and at both time points. When the scores were turned into grade-level equivalents and compared with actual grade in school at time one, more than two-fifths of the abused group were functioning below grade level compared to less than one-fifth of the non-abused group. One year later nearly half the abused group were functioning below grade level, compared to about one quarter or less of the non-abused group.

Interpreting these differences is difficult as there could be a range of explanations for the poorer academic achievement of abused children. For example, Kinard (1991) reports that at Time 1 abused children were absent a maximum of 53 days compared to 35 for their non-abused peers, while the maximum number of days absent at time 2 was 81 for abused children compared to a maximum of 39 days absent for the non-abused children. Therefore, rather than a lack of academic skill, abused children's poorer performance could be the product of social problems such as a lack of familial interest in nurturing and supporting their children's academic achievement. There is a correlation between Socio Economic Status (SES) and intellectual functioning (Sattler, 1992). Low SES families have parents with lower educational achievement themselves, who may be less able to help their children academically (Kinard, 1999) and SES has also been correlated with child abuse, particularly neglect (Sedlak & Broadhurst, 1996). As child abuse can increase the likelihood of psychopathology (Porter et al., 2005), difficulties associated with various pathologies may also prevent abused children from achieving their full academic potential.

With regards to the specific cognitive skills that have the potential to affect children's performance in a forensic interview, Eigsti and Cicchetti (2004) found language delays in both vocabulary and production of syntactic structures in maltreated children. In this sample the maltreatment occurred before they were 2 years old yet at 6 years old, the maltreated children still exhibited an overall 3-month delay compared to their peers. Coster, Gersten, Beeghly and Cicchetti (1989) found

42

delays in acquisition of key linguistic abilities in maltreated toddlers at 2 ½ years-old. This was the case for both language production, which hampers the ability to communicate clearly, and skill in using language as a means of positive social exchange. Maltreated children engaged in shorter episodes of sustained dialogue, they were more likely to terminate a topic after a briefer length of exchange, offered fewer utterances that were relevant to the ongoing dialogue when they were interacting and were less likely to seek information from their partners. It is possible that delays in language production could stem from a lack of interest in social communication and are not necessarily reflective of a lack of comprehension. A lack of willingness to engage in conversation could exacerbate a lack of language production over time by adjustments conversational partners may make in response to children who appear to be unskilled at language production, this unchallenging communication may eventually contribute further to language delays.

Child maltreatment has been associated with lower IQ and vocabulary scores (e.g.,Armsworth & Holaday, 1993; Friedrich, Einbender & Leucke, 1983; Tarter, Hegedus, Winsten, & Alterman, 1984), both of which have been shown to be associated with memory performance (Alloway & Alloway, 2010; Kane & Engle, 2002). Researchers have tested whether childhood abuse has a detrimental effect on memory specifically and numerous studies have found that maltreatment does not appear to uniquely affect children's memory performance (e.g. McWilliams, Harris & Goodman; 2014). Porter et al., (2005) compared memory and intellectual functioning in 8- to 14-year-olds that had

43

experienced multiple episodes of sexual abuse and matched controls. When SES and IQ were controlled-for, abused children performed within the average range of memory and no difference was found in memory performance. Using the Deese – Roediger –McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), Howe, Cicchetti, Toth and Cerrito (2004) assessed memory and false memory in abused (physical abuse, sexual abuse, emotional abuse and neglect) and non-abused low SES status 5- to 12-year-olds. The DRM paradigm involves presenting lists of words that all have a common theme e.g. 'sleep', but never actually presenting the word 'sleep' itself, instead items such as 'bed', 'dream', and 'blanket', may be presented. The children listened to recorded word lists and to recall the presented words and then to listen to lists of words and say whether or not they had been previously presented. Neither susceptibility to false memory recall (recall of the 'critical' non-presented word e.g. 'sleep' was taken as evidence of a false memory) nor recognition of false items, differed as a function of maltreatment status. Recall of true items did not differ as a function of group either, instead showing improvement with increasing age.

Episodic memory is a type of memory specific to recalling past experiences, as opposed to memory for facts and knowledge or habits and skills (Tulving, 1972; 1985), and is the type of memory that a forensic interviewer would be attempting to elicit when interviewing a child about suspicions of abuse. Neuroimaging evidence has shown that there are a number of cerebral structures involved in the encoding and retrieval of episodic information, such as the hippocampus (Desgranges, Baron, &

44

Eustache, 1998). The hippocampus is also an area of the brain that is more susceptible to the long-term effects of stress than some other brain areas. When an individual experiences a threat in their environment, the neuroendocrine system responds by facilitating the release of corticosteroid hormones via the pituitary gland, such as glucocorticoid into the bloodstream in high amounts in order to elicit a behavioural response. However, if the exposure to stressors is long-term, the normal adaptive function of these reactions to stress becomes maladaptive with changes in the brain occurring as a means of helping the individual to adapt to high levels of lifelong stress. Due to the high number of glucocorticoid receptors located in the hippocampus (Teicher et al. 2003) chronic stress has been shown to alter the function of the hippocampal cells (Joëls, 2008). Hippocampal changes have been frequently associated with trauma exposure, such as reduced size in the left hippocampus of adults who have experienced trauma (Woon, Sood & Hedges; 2010; Woon & Hedges, 2008).

While differences in hippocampal volume are found in adults who suffered maltreatment as children, childhood exposure to trauma is not associated with hippocampal volume reduction in children (Woon & Hedges, 2008). These findings could however support the idea that brain regions have their own unique windows of sensitive periods where stressors and nurturing experiences can have a lasting effect. Therefore, although the developing hippocampus is vulnerable to exposure to early stress, the stress induces gradual changes on the hippocampus, meaning the detrimental effect on hippocampal volume is not apparent for some

time. Indeed, Andersen et al. (2008) found reduced hippocampal volume in adults who had experienced childhood sexual abuse at a particular stage in childhood between 3 and 16 years of age. Those who had suffered abuse at 3-5 years old and 11-13 years had significantly smaller hippocampal volumes compared to those who had experienced abused between the ages of 6-8, 9-10, and 14-16 years of age.

A limited number of studies have compared abused and non-abused children's recall of episodic memories in an effort to test whether the lack of differences between abused and non-abused controls' memory remains so for personally experienced events (Chae et al., 2011; Eisen, Goodman, Qin, David & Crayton; 2007; Eisen et al., 2002; Goodman, Bottoms, Rudy, David, & Schwartz-Kenney, 2001). Goodman et al. (2001) compared abused (sexual and physical) and non-abused 3- to 10-year-olds' recall of a dress-up play session with a research assistant following a 2-week-delay. Abuse status did not predict the amount of correct information freely recalled (instead age was the predicting factor with older children recalling more correct information than younger children) or the amount of correctly answered misleading questions. In response to specific questions (wh- and yes/no) non-abused children gave more correct responses than abused children, however, the abused children differed from the non-abused children on IQ and behavioural disturbance measures and these factors may have been the cause of this effect rather than abuse status. Chae et al. (2011) compared 3-to 16-year-old abused (sexual, physical, neglect) and non-abused controls' recall of a different play event following a 3-day delay.

46

No significant abuse-related differences were found in children's memory performance as measured by the amount of correct information in response to free recall and cued recall question or the errors in response to specific and misleading questions. Again, older children provided more correct information in response to free recall questions. Children with higher cognitive functioning and less trauma-related psychopathy (assessed by tests of intellectual ability, language comprehension and short-term memory) provided more correct information in response to free recall questions and fewer errors in response to specific and misleading questions.

In an effort to apply abused children's memory performance findings in forensic contexts, researchers have examined their memories of stressful experienced events that involved physical contact. Eisen et al., (2002) asked specific and misleading questions to 3-to 17-year-old control and abused (sexual and physical) and neglected children regarding an anogenital examination they received. Measures of psychopathology, short-term memory, and intellectual ability were also assessed. Abuse status was not significantly related to the children's memory performance, instead age (older children evinced fewer memory errors and greater suggestibility resistance than younger children), psychopathology, short-term memory, and intellectual ability all predicted facets of children's memory performance. Eisen et al. (2007) compared memory performance of 3- to 16-year-olds that had been either sexually abused, physically abused, sexually and physically abused, or neglected with those of non-abused controls for a medical examination and

47

venepuncture. Maltreated children's memories did not differ from control children's in the amount of correct information provided to free recall and open-ended questions or the proportion of errors to specific questions, instead abuse status was significant only for errors to misleading questions. When maltreatment status was taken into account however sexually abused and physically abused children were significantly less suggestible than neglected children. Older children provided more correct information and fewer errors overall than younger children did and children that self-reported higher levels of trauma symptoms produced the more errors to specific and misleading questions.

Studies that have directly compared abused and non-maltreated children's eyewitness memory, for both neutral and stressful events, have all pointed to the same conclusion; experiencing childhood maltreatment does not predict diminished episodic memory ability. There were no differences in the amount of correctly recalled information during free-recall or the number of errors made to specific questions by abused and non-abused children and only one study found a difference in the number of incorrect responses to misleading questions. Maltreated children appear no better or worse than their non-maltreated peers at recalling neutral or stressful events. Instead children's recall was consistently predicted by age, psychopathology and cognitive functioning. Older children, children with lower levels of psychopathology and children with higher levels of cognitive functioning, recalled more during free-recall and answered a greater number of specific and misleading questions correctly than did younger children, children with higher levels of psychopathology

48

and children with lower levels of cognitive functioning. When different

subtypes of abused children have been compared, sexually and

physically abused victims performed better than neglected children,

providing more correct information in response to open-ended questions,

making fewer omission errors, and showing significantly lower levels of

suggestibility (Eisen et al., 2007; Gaudin, 1999). It has been theorised

that neglected children may perform more poorly on a variety of tasks,

including those assessing memory, due to their particular abuse status.

Neglected children have suffered from a lack of their basic needs being

met and inattention of their parents, which may in turn have jeopardised

their overall cognitive functioning (Gaudin, 1999).

 1.4.1 Summary. Research shows that maltreated children evince

poorer academic performance than non-maltreated children (Ayoub et al.,

2006; Hostinar, Stellern, Schaefer, Carlson & Gunnar, 2012; Kinard,

1999, 2001; Mills et al., 2011). Whether this is evidence of poorer global

cognitive skills (Armsworth & Holaday,1993; Friedrich et al., 1983; Tarter

et al., 1984), language delays (Eigsti & Cicchetti, 2004) or lack of familial

support of their education (Kinard, 1991) remains to be clarified. It is

surprising, considering the long-term effects traumatic and stressful

experiences can have on brain structures (Andersen et al., 2008; Woon

et al., 2010), that such changes have not been found in the abused

children (Woon & Hedges, 2008), instead appearing gradually over time.

Often adults who have experienced childhood abuse present with

psychopathology and/or alcohol/drug abuse, which may explain damage

to these brain structures in adulthood. Indeed abused children's memory

49

performance parallels non-abused children's, both groups recall similar amounts of correct information and resist misleading questions about events they have personally experienced (Chae et al., 2011; Eisen et al., 2002; 2007; Goodman et al., 2001). Instead the most consistent predictors of memory performance are age and presence of psychopathy. Differences between abused subgroups have been found in memory recall, neglected children exhibit poorer performance than sexually and physically abused children (Eisen et al., 2007).

Future research is needed to address the discrepancy between abused and non-abused children's academic performance. Extra learning support programmes for the children are needed as well as any family support that could help engage the children's parents' interest in their children's education e.g. educating mothers (Kinard, 1999) and addressing absenteeism at an early stage. Providing education to children and adolescents about the effects of substance abuse to brain regions and memory structures should be provided and alternative coping strategies should be explore with abused populations before maladaptive coping strategies such as using substances have become a cycle in their lives. More research on how different aspects of psychopathology specifically affect eyewitness memory and interviewing strategies is also needed in order to reduce these potential problems.

1.5 *Implications for forensic interviewing*

Child abuse can have serious consequences on some children's social, emotional and cognitive development, which could, in turn, affect

abused children's ability to recall and recount a past event to an interviewer. Abused children are more likely to have insecure attachments and social difficulties with both caregivers and peers, and come to expect their needs will go unmet and expect less help from parental figures (Edwards et al., 2005; Egeland & Sroufe, 1981; Finzi et al., 2000; Macfie et al., 1999; Shipman et al., 2000, 2005).Maltreated children may automatically expect less support from an interviewer and have a pervasive anxious or avoidant style of attachment, making it difficult to establish a secure relationship that fosters disclosure in an interview setting. There is field evidence to show that from the outset of a forensic interview non-disclosers are less cooperative, and their expectation of less support is then confirmed as interviewers then offer fewer supportive comments (Hershkowitz, Orbach, Lamb, Sternberg & Horowitz, 2006). Fostering a supportive environment for maltreated children is vital as they may already be mistrustful and interviewers must resist any inclinations to withdraw support. Using interview strategies that emphasise emotional and social support techniques from the outset, such as the Revised NICHD (National Institute of Child Health and Human Development) Investigative Interview Protocol (Hershkowitz, Lamb, Katz & Malloy, 2013), can help reluctant children to form a connection with the interviewer and enhance their motivation to disclose (see chapter 4 for a more detailed description of the Revised NICHD Protocol).

The cognitive difficulties that maltreated children may present, such as refraining from using conversational interactions as a positive social tool, engaging in shorter interactions and failure to describe

elements to do with the self, could also impact abused children's performance in forensic interviews. As maltreated children may not be used to lengthy social interactions or to describing their own experiences in detail, building verbal rapport and engaging them in narrative elaboration training, where the child describes a neutral real life experienced event, is useful in providing an example to maltreated children of how to structure and report their experiences. The NICHD (National Institute of Child Health and Human Development) Investigative Interview Protocol (Lamb, et al., 2007a) was developed specifically to be used with children and was designed in mind of their shorter attention spans, limited linguistic skills, partial mastery of concepts and poorer memory retrieval skills. It is also suitable for use with children who suffer from disabilities (Hershkowitz, Lamb & Horowitz, 2007) and therefore is suitable for use with children with abuse related language deficits. The effect of experiencing child abuse does not appear to affect eyewitness memory, but age does, with younger children recalling less information during free recall than older children (Lamb et al., 2003). Younger children may benefit more from 'cued invitations' (e.g. "You said he touched your bottom. Tell me more about that.") that narrow the focus of the information requested from children using cues they have previously mentioned, than from the less specific 'general invitations' (e.g. "And then what happened?") (Lamb et al., 2003; Lamb, Hershkowitz, Orbach & Esplin, 2008; Orbach & Lamb, 2000).

Individual characteristics must also be taken into account when planning an interview. For example, children's particular abuse statuses

may require interviewers to be more cautious about particular aspects of the interview with different abuse groups. As neglected children are more susceptible to suggestive influence than other abuse statuses (Eisen et al., 2007; Gaudin, 1999), interviewers must take care not to inadvertently lead a neglected interviewee. On the other hand, social dynamics may be more important when interviewing physically abused children, who may be highly vigilant to threat and interpret innocuous communications to be threatening, and sexually abused children who report reduced trust in people. Psychopathy and cognitive functioning also impact how accurate children are when answering questions about events they have experienced; therefore it may be useful to conduct evaluations of children's clinical symptoms and intellectual functioning prior to conducting a forensic interview.

1.6 Overall chapter conclusion

While research shows that experiencing child abuse often results in social, emotional, and cognitive detriments, not all children that are abused will experience such adverse outcomes. Further, not all children that are alleged abuse victims have actually been abused. Many children participate in investigative interviews because there is a suspicion of abuse or an allegation has been made by a third party, while some of those children really have experienced abuse, some will be interviewed due to incorrect suspicions and some make false allegations, both mistakenly and knowingly. Therefore, in order to understand how to best train interviewers that question children about suspected abuse, the development and dynamics of all aspects of how children remember both

53

traumatic and non-traumatic experiences as well as other cognitive

factors that may impact children's ability to recount their experiences

must be understood.

Chapter 2: The development of episodic memory

The primary purpose of an investigative interview is to elicit children's accounts of events that they have personally experienced (Scottish Executive, 2003; 2011; Lamb et al., 2008). Sometimes these events were in the very distant past and often experienced on more than one occasion (Andrews & Lamb, 2014; Gaudagno, Powell & Wright, 2006). Such memories are known as episodic memories (Tulving, 1972, 1985, 2002, 2005) and this chapter will begin by discussing the various ways episodic memory has been defined in the literature. Lack of agreement on a definition has led to difficulties in determining when this memory system becomes functional in children, therefore, the differing criteria used to assess episodic memory and how this affects the age at which episodic memory is thought to be present in children is also discussed. In order to understand how to best elicit accounts of specific instances of suspected abuse it is vital that interviewers' training educates them about the development and dynamics of how children remember their experiences at different points during their childhood. This chapter will review children's memory capabilities and limitations for events that have happened in their lives from the time of their earliest memories in infancy, throughout early and middle childhood.

2.1 Defining episodic memory

Memory is a higher cognitive process involving the encoding, storage, consolidation and retrieval of information (McDermott & Roediger, 2016; Melton, 1963; Shing, Werkle-Bergner, Brehmer, Müller,

55

Li & Lindenberger, 2010). However, there is great variability in the nature

of information that can potentially be processed, the principles of

operation and the brain structures involved in these processes

(McDonough, Mandler, McKee, & Squire, 1995; Scoville & Milner, 1957;

Tulving, 2002). This has led many researchers to propose that memory is

not a monolithic entity, but in fact a system comprised of multiple

qualitatively distinct systems which are capable both of interacting and

also of functioning independently (Cohen & Squire, 1980; Kandel &

Squire, 2001; Moscovitch, 1985; Nelson & Fivush, 2004; Schacter, 1987;

Squire, 1994, 2004; Tulving, 1985; Tulving, & Markowitsch, 1998). One of

the most commonly accepted distinctions in multiple memory theory is

that of procedural memory or "knowing how" and declarative memory or

"knowing that" (Cohen & Squire, 1980). Procedural versus declarative

memory is also referred to as non-declarative and declarative memory

(Squire & Zola, 1996; Zola-Morgan & Squire, 1990) or implicit and explicit

memory respectively (Graf & Schacter, 1985).

Procedural memory is memory for non-conscious learning abilities,

such as habits and priming, procedural learning and perceptual

representations. This type of memory requires no conscious awareness

to transfer information into memory and no conscious effort to recall or

utilise the knowledge in the present (Cohen & Squire, 1980; Graf &

Schacter, 1985). Declarative memory, on the other hand, is memory that

encompasses memory for both data based materials and facts and

knowledge, as well as memories of personally experienced events.

Declarative memory requires thought or directed conscious attention to

bring stored information back into consciousness awareness and to be able to utilise the information stored (Cohen & Squire, 1980; Graf & Schacter, 1985). For example, Cohen and Squire (1980) found that amnesiacs could learn new skills, such as mirror reading, which involves the learning of a pattern-analysing skill, at a rate equivalent to matched controls and that they could retain these skills over time. These same patients however, had no conscious recollection of ever having previously performed the task and did not recognise words that had been repeatedly presented during testing sessions as being repeated. In addition to the well-replicated finding of preserved procedural memory and damaged explicit memory in amnesiac patients, dissociations in adults and children with normally functioning memory between tasks thought to measure procedural and declarative memory have been found (e.g., Jacoby, 1983) providing further support for these two systems.

In 1972, Endel Tulving proposed that declarative memory can be further categorised into subsystems: semantic memory and episodic memory. Semantic memory pertains to facts and knowledge about the world while episodic memory represents memory for our experiences and specific autobiographical events, including the temporal and spatial aspects, the 'where' and 'when', of events that have been personally experienced. For example, knowing that World War II ended in 1945 is an example of a memory that one might recall from semantic memory. This information can be utilised without the individual having any memory of where and when they acquired this piece of information. On the other hand, recalling arriving late to a history lesson on the day that World War

57

II was being studied is an example of an episodic memory because it is a personally experienced recollection from one's life known to have been experienced in the past in a specific time and space. Tulving later added additional criteria to the definition of episodic memory, firstly, autonoetic or 'self-knowing' consciousness (a sense of recollecting or re-experiencing of one's own previous experiences, a conscious awareness that this event happened to "me" in the past, that does not accompany retrieval of other kinds of memories; Tulving, 1985). Then later, chronesthesia, a consciousness that enables the individual to mentally time travel, not only backwards, but also forwards (Tulving, 2002; see also Suddendorf & Corballis, 2007 for a review).

Tulving's (1985) original definition of episodic memory made reference to the necessity that in order for a memory to be truly episodic that it must contain both temporal (when) and spatial (where) components about experienced events in addition to the semantic content (what). This component of Tulving's criteria has become known as www (what-where-when) memory and some researchers (e.g. Clayton & Dickinson, 1998; Fugazza, Pogány & Miklósi, 2016) have identified the presence of 'episodic-like' memory based on such criteria. This approach accepts behavioural evidence alone, with no reference to the phenomenological experience of recollection, and is known as the 'minimalist approach' (Burns, Russell & Russell, 2015; Russell & Hanna, 2012) to episodic memory. In direct contrast to the minimalist approach is the 'conceptual approach' (Perner, Kloo & Gornik, 2006; Tulving, 2005). This approach moves away from the contents of the memory, instead defining it by the

process which it was recollected, therefore, in order for memory to be determined as episodic, it is essential that the rememberer understands that they are recollecting a past experience.

A third approach, known as 'Kantian minimalism', recognises qualities from both approaches (Russell & Hanna, 2012). It is minimalist and non-conceptual but also captures a form of phenomenological experience not reliant on a consciousness criterion. The authors claim that all experiences 'present objects in spatial positions and in temporal sequences relative to where the organism was located and to when the objects appeared to the subject' (Russell, Cheke, Clayton & Meltzoff, 2011, p. 369). Therefore, the defining features of Kantian minimalism compared to the other minimalist approach are that 1) the 'temporal' aspect of an episodic memory does not have to mean time from the present but can be measured by the temporal order of elements within the episode. 2) The spatial component of an episodic memory can relate to the depth and distance of objects within the episode (spatial relationships) rather than being landmark-based (Russell et al., 2011), and 3) phenomenological experience is a part of episodic memory.

This lack of consensus on what makes a memory truly episodic and the difficulty in identifying the presence of features such as autonoetic consciousness and mental time in the absence of language make it difficult to identify when episodic memory is functional and its developmental trajectory. Using the minimalist approach, even non-human animals can be attributed episodic memory (Clayton & Dickinson, 1998; Fugazza, Pogány & Miklósi, 2016). According to the conceptual

approach, conceptualising a memory as a re-experience is what makes a memory truly episodic, therefore, non-human animals and infants are not capable of episodic remembering because these concepts evolve from processes such as theory of mind which are not developed until around 3-or 4-years-old in humans (Perner & Ruffman, 1995). Recent research has aimed to demonstrate that by three years of age children have some episodic memory abilities (Atance & Sommerville, 2014; Scarf, Gross, Colombo & Hayne, 2013; Scarf, Smith & Stuart, 2014; Suddendorf, Nielsen & Von Gehlen, 2011) a point of view shared by the Kantian minimalists (e.g. Russell & Hanna, 2012). According to Tulving (2002) episodic memory is a "recently evolved, late-developing, and early deteriorating" (p. 5) memory system. Tulving and others (e.g. Nelson, 1993; Nelson & Fivush, 2004; Perner, Kloo & Gornik, 2006; Perner & Ruffman, 1995; Shing et al., 2010) propose that children are not capable of forming and retaining memories of an episodic nature until at least 4-or 5-years-old.

The following section will discuss children's memory for events they have experienced and whether these are episodic based on the approaches discussed above. For this literature search, the author's familiarity with the literature, and literature cited by Patricia J Bauer, Robyn Fivush and Harlene Hayne were used as starting points. I then performed searches using Web of Science to find more recent and additional studies, based on the keywords "children" or "infant" used in combination with "memory development" and "episodic memory" .

2.2 Early episodic memory - Infancy

Having not yet acquired language, it a formidable task to pinpoint whether infants and young children possess a functioning episodic memory system, and if so how early it becomes available during childhood. One of the defining features of episodic memory operation is that it is conscious (Bauer, 2004; Tulving, 1985; Newcombe Lloyd & Ratliff, 2007). Infants cannot confirm whether or not conscious processes are involved in their memory recall, as this would require asking for their insight into their cognitive processes (Newcombe, Balcomb, Ferrara, Hansen & Koski, 2014; Rovee-Collier, 1997; Russell et al., 2011). Therefore, researchers have developed non-verbal tests of episodic memory, the deferred imitation task is appropriate for use with infants, making it the earliest measure of pre-verbal episodic memory (Bauer & Hertsgaard, 1993; Collie & Hayne, 1999; Meltzoff, 1985; 1988). Deferred imitation tasks involve subjects being shown an action or sequence of actions modelled using props, then after a delay being given the opportunity to behaviourally re-enact the action or sequence previously witnessed (see Bauer & Mandler, 1989; Bauer & Shore, 1987; Meltzoff, 1988, for examples). The following section will discuss the development of episodic memory through infancy, reviewing the research on deferred imitation, with particular reference to studies that that have attempted to isolate and test specific elements thought to be representative of episodic memory in their tasks.

Deferred imitation tests have demonstrated that by 6-months-old infants can imitate actions they witnessed being performed 24-hours

61

previously (Barr, Dowden &Hayne, 1996; Collie & Hayne, 1999), with retention increasing to one month for sequences demonstrated to 9-month-olds, and 3 months retention evident for sequences shown to 10 month-olds (Carver & Bauer, 2001, experiments 2 and 3). Such changes indicate significant development in infant memory retention in the latter half of the first year of life, that is, they can recall their experiences over longer delays. Further, while the 9-month-olds could recall the sequences after 1 month, they could not recall them after 3 months, whereas the 10-month-olds could recall sequences after both one and three-month-delays (experiment 3) and even over 6-month-delays (experiment 4, Carver & Bauer, 2001). The second year of life is a time of further improvement, where infants can cope with even longer delays. Here, six months retention has been shown at 13-14-months-of-age (Bauer, Wenner, Dropik, Wewerka & Howe, 2000; Meltzoff, 1995) and 12-months retention at 20-months-of-age (Bauer et al., 2000). In addition to increased retention, older children require less exposure to stimuli than younger children to evidence long-term recall after a delay. For example, the 6-month-olds in Barr et al. (1996) required six demonstrations in order to imitate the sequence 24-hours later, whereas groups of 12-, 18-, and 24-month-olds exhibited imitation 24-hours later after only three demonstrations. By 13 months infants recalled sequences they had been exposed to once as well as sequences they had been exposed to twice (Carver & Bauer, 2001).

The latter half of the first year of life has been implicated as a period of important development in memory encoding (Bauer et al.,

2006). Bauer et al. (2006) compared infants' encoding processes as they observed different action sequences electrophysiologically (measuring their brain activity through recording of event related potentials) and their long-term recall of the sequences using measures of behaviour (deferred imitation), when they were 9- and then 10-months-old. The pattern of encoding employed by the infants when they were 9-months-old was the opposite of the encoding pattern they used once they had reached 10 months of age, further, after 1-month delays, events encoded at 9 months of age were not recalled, whereas events encoded at 10 months were. Although infants encoded the event regardless of age, improvements in the encoding strategies employed in the older infants resulted in more robust memory traces that were more resistant to forgetting.

These behavioural findings of memory improvement compliment the neurological research (Bauer et al., 2006; Carver & Bauer, 2001; Ghetti & Bunge, 2012; Gómez & Edgin, 2016; Newcombe et al., 2007; Usher & Neisser, 1993; Utsunomiya, Takano, Okazaki, & Mitsudome, 1999). The structures that support organisation and consolidation of explicit memory, such as the hippocampus (Bauer et al., 2006; Newcombe et al., 2007), are not fully developed at birth (Usher & Neisser, 1993), instead undergoing dramatic maturational changes through the first two years of life (Utsunomiya et al., 1999). Other structures that support explicit memory in humans, such as the medial temporal and cortical areas, which serve as long-term repositories for explicit memories, and the connections within and between them (Bauer et al., 2006; Carver & Bauer, 2001) only reach functional (though not full)

63

maturity late in the first year of life, with development continuing over the course of the second year (Bauer, 2006; Carver & Bauer, 2001). Such changes would be expected to increase reliability and robustness of recall and temporally ordered recall around this period of life (Bauer, 2004).

As these processes involved in forming and consolidating memories undergo protracted development, they are less efficient in younger compared to older children. This means that during consolidation loss of information is more likely, in turn making traces more vulnerable to forgetting (Bauer, 2015a). Pathman and Bauer (2012) found initial post-encoding processes predicted 16-, 20-, and 24-month-old infants' long-term memory recall of sequences using an elicited imitation paradigm over different time-points. Infants' recall of sequences was tested immediately, after a 1-week and after a 1-month delay. When success of encoding was free to vary (infants did not have to reproduce target sequences in order to be included) the authors could include encoding (immediate performance) as well as post-encoding (1-week performance) as predictor variables of long-term recall in regression analyses. Encoding was a significant predictor of the variance in infants' long-term recall only when it was sole variable, when post-encoding performance was added to the analysis encoding became non-significant.

In a second experiment they measured infants' immediate recall and long-term recall (1-month-delay) of sequences using different post-encoding delays (15 minutes, 48 hours and 2 weeks). Multiple regression analyses revealed the initial period of consolidation (the 15 minute and 48 hours post-event tests) significantly predicted variance in long-term recall.

64

Immediate encoding and the 2-week assessment were rendered non-significant, despite the latter testing point being the closest in time to the final assessment. This indicates that post-encoding processes soon after an event explain more of the variance in infants' long-term memory than measures of initial encoding processes. Bauer (2015a, 2015b) posits that a complimentary processes account of memory development can explain differential long-term recall in older and younger children. This account suggests that with increasing age improvements in the quality of memory traces combine with a decrease in the vulnerability of memory traces, which in turn lead to enhanced memory performance throughout infancy and childhood.

2.2.1 Deferred imitation tasks that measured specific components of episodic memory.
Some researchers remain unconvinced that deferred imitation tasks truly measure episodic memory, instead attributing such memories to be of a semantic nature, concerned with general knowledge rather than episodic (e.g. Mandler; 2002; Newcombe et al., 2007). In an effort to address these concerns, researchers have isolated specific components of episodic memory and integrated them into deferred imitation tasks (e.g. Bauer & Leventon, 2013; Bauer & Lukowski, 2010; Lukowski, Garcia & Bauer, 2011). Bauer and Leventon (2013) tested 13-, 16- and 20-month-olds' memories of 3-step sequences that they had only been exposed to on one occasion, following a delay of either one or three months. Episodic memories are highly specific memories of experienced events in one's life located in a particular time and place (Tulving, 1985), therefore recall of a one-time

exposure should indicate that the memory is of an episodic and not of a semantic nature as it must be based on a one-time experience and cannot have been amassed over multiple learning opportunities. After the one-month-delay, both 16- and 20-month-olds could re-enact the actions of the sequence they had seen performed only once and the 20-month-olds could also reproduce the sequences in correct temporal order (Bauer & Leventon, 2013). After a delay of 3 months neither age group could recall the sequences experienced only once, instead multiple exposures were needed for memory to have been maintained over this longer delay.

In a second experiment examining 13-month-olds' long-term recall of two-step sequences, no recall was exhibited after a one-month-delay of sequences witnessed only once and after a 3-month-delay long-term recall was not demonstrated even for sequences they had had multiple exposures to. The authors believe that forgetting was a reasonable explanation for lack of recall of a three-step-sequence in 13-, and 16-month-olds' after one or three months respectively because they produced fewer actions after three months than they did after one month. However, 20-month-olds failure to recall the sequences after 3 months did not appear to be due to forgetting. Instead the 20-month-olds' failure to produce the target actions was likely to be due to their preoccupation with producing spontaneous new actions that had been shown as control sequences at the one month testing session.

Memory for the specific features of an event are another component indicative of episodic memory (Tulving, 1983). Bauer and Lukowski (2010) tested 16- and 20-month-olds' long-term recall and

66

recognition memory for three-step sequences and the props used to enact them one month after observing and imitating the sequences. The infants completed a delayed forced choice recognition test where they chose the prop they had used before from the actual prop and a similar one and memory for the sequences was tested using delayed imitation. Both ages evidenced memory for the actions and sequences with the older infants outperforming the younger ones. While there were no differences between the 16- or 20-month-olds' memories for the specific props, memory for the props predicted unique variance in recall performance of the event one month later in the 20- but not 16-month-olds. Therefore, both age groups encoded the specific features of the event (the props) but the older children were able to utilise this information to aid recall of the event itself while the younger children were not. This experiment lends support to 1) the idea that long-term recall is facilitated by memory for the specific features of an event, and 2) the increasing body of evidence that suggests that episodic memory is gradually emerging by the end of the second year of life as this skill does not appear to be available until the latter part of the second year of life.

The spatial 'where' and temporal 'when' aspects of a memory are specific components of episodic memory (Tulving, 1985), Lukowski et al. (2011) found evidence for such location memory in infants in the latter half of the second year of life. Thirteen-, 16-, and 20-month-old infants participated in 3 exposure sessions of a one-step elicited imitation task in a room that contained a wall sticker of a giraffe that was hidden by a curtain. The curtain was slid back and the infant encouraged to 'say hello

to the giraffe'. At delayed recall sessions of either 1-, 3-, 6-, or 12-months the infants completed the elicited imitation task in the same room as the testing sessions had occurred and were verbally prompted to 'say hello to the giraffe' again, correct responses were counted as looking or pointing to the curtain. All infants immediately orientated to the giraffe, however, only the 16- and 20-month-olds showed significantly better performance at the delayed recall than the baseline, with 20-month-olds outperforming 16-month-olds. This difference was not due to superior encoding as infants of all ages performed the target action 100% of the time in the immediate recall test.

Recall of temporal order is thought to measure episodic memory as temporal ordering cannot be recalled unless genuinely encoded and retrieved from memory as no perceptual support remains for order once the action has been modelled (Bauer et al., 2006). Recall of temporal order has been shown to increase over the first year of life. Carver and Bauer (2001, experiment one) found that aged 10 months, only some children recalled the temporal order of events, when tested again at 13 months all of the children recalled temporal order. Throughout the second year of life and by 20 months, children could recall temporal order even after a single exposure (Bauer & Leventon, 2013). Research has also shown performance on an imitation task at 20-months-old predicts performance in later episodic memory at 6 ½ years old (Riggings, Cheatham, Stark & Bauer, 2013).

2.2.2 **Summary**. As early as six months old, infants show evidence of recalling events they have experienced over short (24 hour)

68

delays (Barr et al., 1996; Collie & Hayne, 1999), the length of time infants can retain their experience increases throughout the first two years of life (Carver & Bauer, 2001; Bauer et al., 2000; Meltzoff, 1995) and fewer exposures are required to facilitate recall (Barr et al., 1996; Carver & Bauer, 2001). Neurological research has shown that improvements in encoding strategies (Bauer et al., 2006) and consolidation processes (Pathman & Bauer, 2012) contribute to memory becoming more robust and to less information being lost to forgetting (Bauer, 2015a; 2015b). While infants do have explicit memory during the first year of life, it is likely that episodic memories are being created only during the second year of life, due to inability to fulfil criteria thought to measure episodic memory such as temporal ordering (Carver & Bauer, 2001). Of course, episodic in this context is based on fulfilling the minamilist criteria (Clayton & Dickinson, 1998), the remembering of 'what' (props) (Bauer & Lukowski, 2010), 'where' (location) (Lukowski et al., 2011) and 'when' (temporal ordering) (Carver & Bauer, 2001), without making any reference to whether or not infants have awareness of their own processes of recall. During the second year of life episodic memory has emerged and shows continual improvement throughout this year. For example, temporal order is consistently recalled (Bauer & Leventon, 2013) and while 13-month-olds fail to reliably show location memory (Lukowski et al., 2011) and event recall after a single exposure (Bauer & Leventon, 2013), both 16- and 20-month-olds do. Older infants (20-month-olds) were also able to use specific information encoded as part of an event to enhance memory of the event itself in a manner that younger

infants (16-month-olds) were not (Bauer & Lukowski, 2010)

Improvements in the quality of memory representations and decreases in

vulnerability to forgetting, lead to memories being recalled over longer

delays with increasing age throughout infancy, which improves further

throughout childhood (Bauer, 2015a, 2015b).

2.3 *Early childhood*

Once children have developed sufficient language skills to be able

to talk about their previous experiences event recall can be tested

verbally. The following section will focus on both the field research that

has assessed young children's memories of their real-life experiences

and the experimental research that has attempted to assess episodic

memory using more conceptual criteria. The strengths and weaknesses

of children's memory as well as the likelihood that they can display

conceptual aspects thought to represent episodic memory throughout

early childhood are discussed. The strengths and the weaknesses

associated with both approaches in this area are also outlined.

Fivush, Gray and Fromhoff (1987) asked two-and-a-half- and

three-year-olds to recall memorable events that they had experienced

only once or twice (e.g. a trip to Disneyland). The children recalled as

much about distantly experienced events (occurring more than 3 months

beforehand) as they did about recently experienced events (occurring up

to 3 months ago) and all children recalled at least one event that occurred

more than 6 months earlier. These results suggest that by 2-years-old

children are forming memories of personally experienced events that are

available for recall 6 months later. Hamond and Fivush (1991) showed

that children as young as 3-years-old demonstrated retention of a novel personally experienced event even after delays of 18-months. Two groups of children that had visited Disneyland were interviewed about their experience. The youngest group visited Disneyland at 3-years-old on average and the older group at 4-years-old; half of the children in each group were interviewed after a 6-month-delay and the other half after an 18-month-delay. Children recalled a mean of 40 propositions and inaccuracies were "virtually non-existent" (p.437). There was no effect of age or retention interval; even the youngest children gave highly accurate and detailed accounts of personally experienced events even after an 18 month delay, performing as well as older peers who experienced only a 6-month delay.

Fivush, Haden and Adam (1995) interviewed children on four occasions, asking each time about 3 experienced events (such as a trip to a science centre) with the first interview taking place when they were 3 ½ years old. At the subsequent three interviews when the children were roughly 4-, 5- and 6-years-of age, two of the events they recalled were new events they had not previously discussed and one event had been discussed at the preceding interview. At the first interview, when the children were aged 3 ½ years old, they provided lengthy and detailed accounts of events they had experienced previously, for some this was experienced before they had reached 3-years-old. At the second recall opportunity of a previously discussed event, 73% of 4-years-olds recalled an event they had recounted 6 months ago when they were 3 ½ years old, 78% recalled an event at 5-years-old that they had recalled one year

71

previously, and 90% recalled an event at 6-years-old that they had discussed a year ago. Therefore, a total of 80% of the events they had discussed at preceding interviews were still available for recall over delays of either six months or one year. With increasing age both the length of their narratives and the delay after which the event could be recalled increased.

An important caution regarding field studies of children's memory is it cannot be ascertained how many times the events under investigation have been recalled out with testing sessions. Parents were asked how often the novel events had been discussed at home and how often photographs of them they had been looked at by the child, parents had difficulty reporting such information. Hamond and Fivush (1991) obtained estimates of number of times the events had been discussed and photographs had been looked at and divided children into either a "high rehearser" or "low rehearser" group. The "high rehearser" group included children that had looked at pictures of the trip and discussed it six or more times (or "often", "a lot" or "frequently" if parents could not estimate a numerical value). The "low rehearser" group consisted of children that either discussed the trip or viewed pictures of the trip fewer than six times or were rated by parents as "occasionally" or "rarely" discussing it. The results showed that children who visited Disneyworld 18 months ago in the low rehearsal category recounted an average of 34.91 prepositions, while those in the high rehearsal group recalled a mean of 49.92 Therefore, children that discussed their trip more often recalled more information about it at the testing session.

72

Furthermore, as we do not know the content of these discussions, it is possible that in addition to having an opportunity to rehearse their own memories, that adults provided children with information about the trip that they actually did not remember, that become incorporated into memory. Further, parents are not privy to any discussions about the trip that they were not present for, a trip to Disneyland is likely to have been a topic of discussion with other family members and care-givers and friends, therefore their estimates could be undermining how many times these events have been discussed. In the Fivush et al. (1987) study, children actually only recalled 56% of events they were asked about, these events were possibly not as notable as the trip to Disneyland, and were therefore not discussed as often in the aftermath, which may be why they were not as readily recalled as the trip to Disneyland. In addition, as these were field studies, accuracy cannot be reliably measured.

2.3.1 *Tasks that measure specific components of episodic memory*. While the aforementioned studies categorised children's prepositions into categories, they did not have the specific aim of identifying the presence of episodic memory. While the categories encompassed some of these criteria for episodic memory, e.g. Fivush et al (1987) included a location category, and demonstrated children of this age fulfilled some the minimalist criteria of episodic memory, they did not look at whether all three aspects were reported. Further, no reference is made to any conceptual processes involved in the children's recall. Experimental studies have attempted to assess a feature that is involved in episodic memory, but not semantic memory, contextual binding (what-

where-when) of associations that represent experiences (Newcombe et al., 2014) in young children. Testing paradigms developed to measure contextual binding are thought to measure 'episodic-like memory', as they require all three features of 'www' to be remembered but do not measure any conceptual criteria. Further, they can be tested non-linguistically and following only one experience, which is thought to be a measure of episodic recall (Bauer & Leventon, 2013).

Hayne and Imuta (2011) assessed the www binding aspect of episodic memory by playing a game of hide and seek with 3- and 4-year-olds. Three toys (what) were hidden in three different rooms (where) in a specific location e.g. under the bed (where) in a specific order (when). After a 5 minute distractor involving reading stories to prevent the child rehearsing the event the experimenter asked the child what-where-when questions such as "What room did we go into first?" about hiding the toys to obtain a measure of verbal recall. The children were then provided them the opportunity to non-verbally demonstrate their memory by finding the toys in the order that they had been hidden. Four-year-olds were superior in verbal recall on all categories and children of all ages recalled more about the 'where' aspects of the room than the 'what' and 'when'. Non-verbally the 4-year-olds only outperformed the 3-year-olds on the test of 'when' (the order in which the toys were hidden) there were no age-related differences between the where (room or location). This shows that when allowed to recall behaviourally 3-year-old children perform as well as 4-year-old children on some measures of www memory. The 'when' component may have been particularly difficult because the toy

hiding was all conducted during the same session, making differential fading cues (using the strength of the memory trace to judge how recently it was experienced) unavailable. The authors conclude that a rudimentary form of episodic memory is in use by the age of three years old and that the way children are tested (verbal v non-verbal) can have an effect on whether or not episodic memory is attributed.

Following Hayne and Imuta's successful demonstration of episodic memory ability (evidenced by contextual binding) in 3- and 4-year-olds, Newcombe et al., (2014) conducted a similar experiment requiring binding of novel environments and unique cues. Children were shown two unique rooms, both contained four identical containers (in different arrangements), one of which contained a toy. The container that held the toy and the toy itself differed between rooms, e.g. bubbles were hidden in the basket in the rainbow room and crayons in the box in cloud castle, and children entered the different rooms with a different research assistant. The children were provided with two opportunities to retrieve the toys; searches were either scored as correct, semantic search errors (the child searched in the container that held the toy in the other room) or random search errors (the child searched in a container that did not hold a toy in either room). In experiment one the comparison groups were15- to 20-month-olds and 21- to 26-month-olds, experiment 2 compared 34- to 40-month old children with 64-to 72-month-old children and in experiment 3 there were 3 groups of children; 34-to 40-month-olds, 42- to 48-month-olds and 50- to 56-month-olds.

In all three experiments, all children performed better than chance at finding the toys and made more semantic than random errors (apart from the 5-6-year olds who barely made any errors at all). The authors assessed episodic memory as present when children made more correct choices than semantic errors, only one group of children (the 21-26-month olds) failed to do so, even the younger 15- to 20months olds fulfilled this criterion. The average number of correct responses and semantic errors were similar in both of these groups though there was greater variability in performance within the older age group which may account for this finding. Therefore, episodic memory was present in even the youngest group. Errors were negatively correlated with age and the only difference between groups' accuracy was between the oldest children (64-72-month-olds) and the four youngest age groups. This suggests that contextual binding is present as early as 15 months, increasing in skill gradually across pre-school period and maturing at about five years of age. As semantic errors were more common than random errors in these experiments, this suggests that explicit memory is available in young infants but the binding of specific elements of experiences that denotes episodic memory, rather than the general elements of situations, improves over the preschool years.

In addition to examining www contextual binding, Russell et al., (2011) looked for evidence of conceptual processing ability in children, specifically episodic foresight (demonstrating Tulving's chronesthesia element of episodic memory) in 3-, 4-, and 5-year-olds. Children were shown two mini coolers, a red 'hot' box with a sticker of a sun on it which

warmed its contents and a white 'cool' box which kept its contents cold and had a picture of a snowman. Training included discussions and presentations of the effects of the boxes on a chocolate bar and a plain biscuit. After short delays (3-5 minutes) the chocolate remained intact and after a longer delay of 30-45 minutes it melted, whereas the biscuit remained intact after long delay. All children tested showed a preference for chocolate over biscuits and children were told they were not allowed to eat the melted chocolate as it would make a mess. The chocolate was put in the hot box and the biscuit in the cold box; children were taken to play and on return could pick the food out of one box to have.

Children were deemed as having bound the www elements of the episodic like task if they demonstrated episodic foresight and chose the preferred item- chocolate- only when they experienced the short delay, if they experienced the longer delay then the biscuit, the only edible option, would be chosen. None of the age groups performed better than chance although 4- and 5-year-olds showed a trend towards better than chance performance. However, rather than being reflective of a lack of episodic memory, children's floor performance may have been because it was difficult for the children to inhibit their choosing of the chocolate as it was the preferred item.

When provided with only the hot box and asked which food they would put inside before they left for a short or long delay, 3-year-olds did not perform better than chance but 4- and 5-year-olds did. Enhanced performance on this task may be due to only one item being available as children not having to execute inhibitory control. Alternatively it could

77

suggest that some elements of episodic memory may be available earlier than others. Caution should be exercised when interpreting results from this study as it is likely that the task demands were too high due to the use of highly desirable items (chocolate) and a lack of executive control in this age group.

In 2005 Tulving developed a criterion for non-verbal testing of his autonoetic component of episodic memory. The "spoon test", derived from an Estonian children's story, tells of a little girl who dreams one night that she is at a birthday party where her favourite chocolate pudding is served. However, guests were expected to bring their own spoon and as she did not bring one she does not get any pudding. The following evening the girl takes a spoon to bed with her. The girl in the story demonstrates behaviourally by taking the spoon to bed that she can mentally time travel; she anticipates a future return to the party and takes her spoon to bed to be prepared for eating pudding. To test autonoetic episodic memory behaviourally an experiment must fulfil Tulving's (2005) four criteria, 1) temporal separation must occur between episode and test to ensure long-term memory, 2) the episode and preparatory act must occur in different spatial contexts to avoid cueing, 3) a single trial must be used and 4) the result must not be driven by any physiological states e.g. hunger.

Using Tulving's criteria, researchers have developed experimental procedures to identify the presence of autonoetic episodic memory according to the conceptualist approach, in 3- and 4-year-olds (e.g. Atance & Sommerville, 2014; Scarf, Gross, Colombo & Hayne, 2013;

78

Suddendorf, Nielsen & Von Gehlen, 2011) Scarf et al., (2013) took

children individually to a large outdoor sandbox they were told contained

buried treasure. The child and experimenter find a treasure chest in the

sand, but it is locked and neither have a key. Twenty-four hours later the

children return and are offered a choice of three items that they are

allowed to keep: a ball, a wind-up toy and a key. After selecting their

object they are taken back outside to the sandbox. Control groups of 3-

and 4-year-olds did not visit the sandbox but were offered one of the

three items as a gift. A greater number of 4-year-olds (75%) chose the

key than 3-year-olds (33%). The 3-year-olds did not select the key at a

rate above chance while the control group selected the key a rate of

chance. When children that chose the key were asked why they had

chosen as such only 2 children (1 3-year-old and 14-year-old) did not

make reference to the locked treasure chest they had found in the

sandbox 24 hours ago.

Additional 3-year-olds were tested using the same paradigm but

changing the delay from 24 hours to either no delay, a 15-minute delay or

a 30-minute-delay, in an effort to determine the reason behind why the 3-

year-olds did not appear to remember and/or link the key to the treasure

box,. In the absence of a delay, 75% of the 3-year-olds chose the key.

Following a 15 minute delay this rate dropped to 50% and a further drop

to 33% was observed after a 30 minute delay; the same rate of success

children had in the previous experiment following a 24 hour delay.

Compared to the control group, a greater numbers of 3-year-olds in this

experimental condition chose the key after no delay and after the 15

minute delay, but the groups did not differ after a 30 minute delay. As the length of the delay increased the number of children passing the test decreased, indicating children encoded the event but forgot it rapidly. Finally, the experiment was repeated with 4-year-olds using a one week delay. All of the 4-year-olds chose the key, with 75% of children referring to the future utility of the key. These and others' results (Atance & Sommerville, 2014; Suddendorf, Nielsen & Von Gehlen, 2011) indicate that episodic memory, according to Tulving's strict conceptual criteria, is in fact present in children as young as 3 years of age. However, rapid forgetting at this age mean that the episodic memories formed by 3-year-olds are not retained over substantial delays making them unavailable for later retrieval.

This is not an entirely new theory; Brainerd's (1985) review of experimental memory studies concluded that while both storage and retrieval processes improve throughout early childhood to adolescence, the largest improvement in storage processes happens in the pre-school years. During the pre-school years, memory consolidation processes are not fully developed thus memories are not consolidated as efficiently as they are in older children. This is why episodic memories, although formed in both 3- and 4-year-olds, are only retained by the 4-year-olds who have more advanced post-encoding processes, resulting in the memories being better consolidated and thus available for recall over longer delays. Conversely, younger children form such memories but have less mature post-encoding processes, leaving their memories more vulnerable to forgetting. This review evaluated experimental research on

80

children's retention and recognition performance on word lists, therefore the conclusions may not be generalisable to episodic memory performance. However, recent research, including neurological studies, have now provided evidence to further substantiate this conclusion (Bauer et al., 2012; Bauer, Larkina & Doydum, 2012). For example, Bauer et al., (2012), found 3- and 4-year-olds' recall of multi-step event sequences after 1-week predicted significant variance in long-term recall after one month, suggesting that it is post-encoding processes that predict whether an event is available for long-term recall (similar findings are reported among both younger and older infants in the above section; see Pathman & Bauer, 2012).

2.3.2. Summary. As early as two-years-old children are forming verbally accessible memories of personally experienced events that they are able to recall six months later (Fivush et al., 1987). The length of time such memories can be retained increases throughout early childhood, with 3-year-olds demonstrating retention of a novel personally experienced event even after delays of 18-months (Hamond & Fivush, 1991). Further, increasing age increases the amount of information children can recall about things they have previously experienced (Fivush et al., 1995). Commenter's have argued whether or not this is demonstrative of true episodic memory, leading to studies of specific features indicative of episodic memory, such as contextual binding and episodic foresight, which have produced mixed findings. Some have claimed episodic memory has been demonstrated non-verbally in children as young as 15-months-old (Newcombe et al., 2014), while others have not deemed it

available until 5-years of age (Russell et al., 2011). It is likely that the reason for such disparate findings is the use of differing testing paradigms, therefore, the way that children are tested is likely to have an effect on whether or not they are credited as having functional episodic memory. Adhering to a criteria developed by Tulving (the 'spoon test') to assess the conceptual aspects of children's memories, thus rendering it truly episodic, research has conclude that episodic memories are created in 3-year-olds (Atance & Sommerville, 2014; Scarf et al., 2013; Scarf, et al., 2014; Suddendorf et al., 2011). However, due to immature post-encoding consolidation processes that are still undergoing developmental improvements in early childhood, they are rapidly forgotten by 3-year-olds rendering them unavailable for long-term recall (Bauer et al., 2012).

2.4 Middle childhood

By four years of age children can readily talk about events they have experienced more than a year previously (see Bauer, 2007; Peterson 2002; Peterson, 2012, for reviews). In fact, memories created when children are in their third year of life have been shown to be available for recall over longer delays of 5 and even 6 years (Fivush & Schwarzmueller, 1998; Van Abbema & Bauer, 2005). Fivush and Schwarzmueller (1998) interviewed children when they were 3-years-old, and again just before they turned 4-, 5-, and 6-years-old about events occurring in the months before those interviews. They were then asked about these events a second time at an interview occurring when they were 8 years old. Following this 5-year delay 78% of events were recalled

82

from the most distant time-point when the children had been on average 34 months old.

Increasing age and retention interval/delay has been shown to linearly decrease the number of distant events available for recall. For example, Van Abbema and Bauer (2005) interviewed 3-year-old children about unique events they had experienced, the children were re-interviewed either when they were 7-years old after a 4-year delay, when they were 8-years old after a 5-year delay, or they were 9-years old following a 6-year delay. The 7-year olds recalled 60% of events, 8-year-olds recalled 36% of events and the 9-year-olds recalled 34% of the events from when they were three. Similarly, Cleveland and Reese (2008) interviewed children about events they had experienced over the course of 4 years, from age 1 ½ until they were 5 ½ years old. They also found that children recalled fewer distant (experienced between approximately 1 ½- to 2-years-of-age) events than recent ones (experienced between 5- and 5 ½- years of age). The children also recalled less information about these distant events and were less accurate when describing them (50% compared to 75% accuracy for recall of events that occurred before and after 2 years of age respectively).

With increasing age the content of children's recall of events from earlier in their life changes in a number of ways. Although the total amount of information recalled about the events at the second interview did not differ from the amount of information that had been recalled originally recalled, the information in the second interview was primarily

83

new, yet accurate, information about the events that they had not reported during the first interview (Cleveland & Reese, 2008; Fivush & Schwarzmueller,1998; Van Abbema & Bauer, 2005). As children got older their recollections also contained a broader range of episodic components (e.g. more temporal information and explanations of why things happened) about the events than their original accounts of these same events at the first session (Van Abbema & Bauer, 2005).

While these studies have demonstrated impressive memory retention and recall abilities by children for personally experienced events, these findings may be unique to the contexts that they were studied in and are therefore not generalizable to all situations where children are asked to recall life events e.g. a forensic interview. The events children were asked to recall in these studies were of a highly memorable, emotionally positive nature and likely to have been lengthy experiences e.g. trips to Disneyland. The types of experiences children may be asked about in a forensic interview may not be highly memorable, lengthy or connected to emotions of a positive nature, thus they may not be directly comparable. The early recall opportunities children had to discuss their experiences may also have bolstered their memory for the original event at the later interview (an early interview has been shown to enhance later recall e.g. La Rooy, Pipe & Murray, 2005), and it is possible the children also recalled these events outwith the interview opportunities in the studies. Finally, no true measures of accuracy were obtained, while the accuracy of children's information was judged as correct by their mothers, it is possible they may have also been mistaken.

2.4.1 Beyond middle childhood. As discussed throughout this chapter, children clearly can remember events from very early in their lives, adults on the other hand are able to recall few memories from the period of life when they were 3-to-4-years-old (Davis, Gross & Hayne, 2008; Dudycha & Dudycha, 1933a, 1933b, 1941; Henri & Henri, 1898; Kihlstrom & Harackiewicz, 1982; Howes et al, 1993; Mullen, 1994; MacDonald et al, 2000; West & Bauer, 1999; Sheingold & Tenney; 1982), a phenomenon known as childhood amnesia (Freud, 1916). Research investigating the onset of childhood amnesia has shown that childhood amnesia begins during childhood itself; as children get older the earliest memory they can recall is from later in life (Jack, MacDonald, Reese & Hayne, 2009; Peterson, Grant & Boland, 2005; Peterson, Noel, Kippenhuck, Harmundal & Vincent, 2009; Peterson, Wang & Hou, 2009; Peterson, Warren & Short, 2011; Tustin & Hayne, 2010). For example, Tustin and Hayne (2010) asked 5-year-olds, 8- to 9-year-olds, 12-13 year olds and adults to report their earliest memories. The children reported a greater proportion of earlier memories than the adolescents, who in turn reported a greater proportion of earlier memories than the adults.

Similarly, 6-to 9-year-olds recalled earlier first memories than 10-to 13-year-olds and 14-to 16-year-olds (Peterson, Grant & Boland, 2005), however, when re-interviewed two years later the children reported their earliest memories being from later in life (more than 7 months later) than they had originally (Peterson, Warren & Short, 2011). This demonstrates the boundary of childhood amnesia moves with increasing age, early memories that could be recalled at one time become unavailable for

85

recall with as children get older. The reason for this has yet to be determined, failures in memory consolidation (Bauer & Larkina, 2014a) and a lack of effective retrieval cues for memories early memories, have been proffered as potential explanations for the inability to recall such early memories (Hayne, 2004; see also Hayne & Jack, 2011). By the age of 8 to 9 years, the pattern of recall apparent in adults is emerging in children (Bauer & Larkina, 2014b)

    2.4.2 Summary. Research has shown that during middle childhood, the amount of information children can recall about events they have experienced after lengthy delays (e.g. 4 to 6 years), is the same as the amount of information they recalled about such events when asked about them at the time when they were in early childhood (Cleveland & Reese, 2008; Fivush & Schwarzmueller; 1998; Van Abbema & Bauer, 2005). However, what they recall at different time points changes, at a second recall opportunity during middle childhood the majority of the information they recall is new information that they did not report when they were in early childhood (Cleveland & Reese, 2008; Fivush & Schwarzmueller; 1998). The new information they report is accurate (Cleveland & Reese, 2008), more coherent and detailed, and contains more temporal information (Van Abbema & Bauer, 2005) than their earlier recollections. While events from early childhood survive into middle childhood, it is likely that upon entering adolescence and adulthood that most of these events will be forgotten, due to childhood amnesia (Freud, 1916) which likely begins in middle childhood (Peterson, Morris, Baker-Ward & Flynn, 2014). As chronological age increases more

86

early events are forgotten, with children evidencing younger first memories, a greater number of early memories than adolescents and adults (Jack, MacDonald, Reese & Hayne, 2009; Peterson, Grant & Boland, 2005; Peterson, Noel, Kippenhuck, Harmundal & Vincent, 2009; Peterson, Wang & Hou, 2009; Peterson, Warren & Short, 2011; Tustin & Hayne, 2010) and an inability to recall early events that were previously available for recall (Peterson, Warren & Short, 2011).

2.5   Overall chapter conclusion

Evidence from both behavioural and neurological research suggests that episodic memory is emerging and improving during the second year of life (Bauer et al., 2006; Carver & Bauer, 2001; Newcombe et al., 2007; Utsunomiya et al., 1999). During this period infants' recall of event sequences becomes more robust, becoming available for longer retention intervals and requiring fewer exposures to be recalled in the long-term (Barr et al., 1996; Bauer et al., 2000; Carver & Bauer, 2001; Meltzoff, 1995). Further, aspects of their recall begin to feature elements thought to measure episodic memory, such as temporal recall (Bauer & Leventon, 2013), location memory (Lukowski et al., 2011), specific features (Bauer & Lukowski, 2010) and recall of events experienced only once (Bauer & Leventon, 2013). Improvements in memory trace strength as well as decreases in vulnerability to forgetting occur over this period, as well as more effective post-encoding processes (Bauer 2015a; 2015b; Bauer et al., 2006; Pathman & Bauer, 2012).

During early childhood toddlers begin to recall their experiences verbally, dramatic increases in retention are further evident during this

period of life, now recall can be exhibited over durations of years (Fivush et al., 1987; 1995; Hamond & Fivush, 1991). When higher order elements of episodic memory such as contextual binding and mental time travel are examined, the research on these elements has produced inconsistent results (Atance & Sommerville, 2014; Hayne & Imuta, 2011; Newcombe et al, 2014; Russell et al, 2011; Scarf et al, 2013; Suddendorf, Nielsen & Von Gehlen, 2011). Some binding tasks show that even 5-year-olds do not display future episodic foresight (Russell et al, 2011). Other paradigms thought to measure autonoetic time-travel have shown that even 3-year olds can form memories of this nature, however, post-encoding consolidation processes are not yet mature enough to stabilise the trace for it to be recalled over longer delays, whereas four-year-olds both encode and stabilise the trace making it available for retention over longer durations than their younger counterparts (Hayne & Imuta, 2011; Scarf et al, 2013). Thus testing paradigms used will undoubtedly have an effect on how early episodic memory is attributed to young children.

During middle childhood, children can recall events over impressive delays (e.g. 6 years, Van Abbema & Bauer, 2005), and surprisingly they recall as much information at these time-points as they did years ago when the events were fresh in their minds (Fivush & Schwarzmueller; 1998; Van Abbema & Bauer, 2005). The content of their recall changes over time with the majority of information recalled later being new, accurate (Cleveland & Reese, 2008; Fivush & Schwarzmueller; 1998), more coherent, more detailed and containing more temporal references (Van Abbema & Bauer, 2005) than their earlier

accounts. However, it is likely these events, even though they were recalled after such impressive delays, will be forgotten once children move to adolescence and to adulthood because with increasing age the age of our youngest memory increases and the number of events remembered from early in life decreases (Jack et al., 2009; Peterson et al, 2005; 2009a; 2009b; 2011; Tustin & Hayne, 2010). Therefore, the forgetting of events, possibly due to lack of retrieval cues (Brainerd, 1985; Hayne, 2004; Hayne & Jack, 2011) and immature consolidation processes (Bauer & Larkina, 2014) occurs within this age bracket.

An important feature of episodic memory is that once retrieval is cued, the probability of future retrieval is more successful (Tulving, 1983). However, the reconstructive nature of memory (Bartlett, 1920; Conway & Playdell-Pearce, 2000; Loftus, 2005; Pezdek & Lam, 2005; Pollio & Foote, 1971; Roediger & McDermott, 1995; Rubin & Umanath, 2015) means that post-event recall can have serious effects on both the strength of the memory trace (the likelihood of subsequent recall) and its content (Hirst & Manier, 2008). That is, the reconstructive nature of episodic memory coupled with post-event factors can impact how accurately an episodic memory is remembered and recalled. This will be the focus of the next chapter.

Chapter 3: Additional factors that predict episodic memory recall

The previous chapter explored the key developmental milestones in memory development, this chapter will explore additional factors and developmental considerations that can affect children's ability to recall past experiences. Many factors affect what will be remembered about an experience. These factors can relate to the experience itself, for example, how much time has passed since the experience (Ornstein, Gordon & Larus, 1992) and whether any similar experiences have occurred between encoding the original experience and recall (Murachver, Pipe, Gordon, Owens & Fivush, 1996). In addition, other cognitive skills such as language (Simcock & Hayne, 2002), memory for source (Foley & Johnson, 1985; Foley, Johnson & Raye, 1983) and vulnerability to suggestive influence (Leichtman & Ceci, 1995) can also affect what is recalled. As the primary task of a forensic interviewer is to elicit the recall of personally experienced memories in a way that maximises both the amount and the accuracy of information recalled, interviewers must be aware of what aspects of our experiences are likely to be remembered and the ways that retrieval can be enhanced or distorted.

For the literature search, the author's familiarity with the literature, and literature cited by Maggie Bruck, Stephen J. Ceci, Harlene Hayne, Carole Peterson and Kim P. Roberts was used as a starting point. I then performed searches using Web of Science to find more recent and additional studies, based on the keywords "children" in combination with

"development" and "language" (Chapter 3.1), "suggestibility" (Chapter 3.2), "source monitoring" (Chapter 3.3) and "delay" (Chapter 3.4).

3.1 Language

The development of language supports the development of verbally accessible memory (Simcock & Hayne, 2002; 2003). This claim has been supported by two lines of research. One showing that children fail to express preverbal memories verbally, despite demonstrating non-verbal memory of events and having since acquired the relevant vocabulary to do so (Cordón, Pipe, Sayfan, Melinder, & Goodman, 2004; Simcock & Hayne, 2002, 2003). The other, focussing on age of adults' earliest memories, which in fact occur on average between the ages of 2 ½- and 4 ½-years-old; the period of life where biggest increase in words occurs (Hayne & Jack, 2011). Together these findings suggest that lack of language may be what prevents early experiences from becoming part of long-term autobiographical memory. If this is the case then language is not an add-on that can be used to recall pre-verbal experiences, instead it is the prerequisite for verbal memory and a factor in ending the dense period of infantile amnesia.

Using a novel game 'magic shrinking machine' that involved pulling a lever, placing a toy into the machine, turning a handle, and retrieving an identical, but smaller, toy from a door on the machine, Simcock and Hayne (2002) demonstrated children failed to verbally recall memories they could demonstrate non-verbally. Children aged between 2- and 3–years-old learned to operate the machine independently, over two sessions separated by a 24-hour delay. After either a 6- or 12-month-

91

delay, during which time parents confirmed that the children's language skills had improved to the point that they could easily produce the target words necessary to describe the event, children's memory of the event was tested. Children in all age groups had no difficulty demonstrating non-verbally that they remembered the event (choosing photographs of the items used during the game and behavioural re-enactment) over both retrieval intervals. However, performance was poor when they were required to verbally recall the event. In fact, none of the children ever used any word in their verbal recall that had not been part of their productive vocabulary at the time of encoding the event.

Following an injury and treatment at the hospital, Peterson and Rideout (1998) interviewed 12-to 34-month-old infants a few days later about their experiences. At this early interview only the oldest children (aged 26-to 34-months-old) could verbally answer questions about their accident and treatment. In follow-up interviews after 6-, 12-, and either an 18- or 24-month delay, these children were able to provide verbal narrative accounts that contained information about how the injury occurred and/or the hospital treatment they had received. On the other hand, the youngest children (12-to 18-month-old pre-verbal toddlers at the time of the accident) could not provide verbal recall (free-recall/wh-questions) about the incident. However, when asked 'yes/no' recognition questions there were no age effects on the proportion or accuracy of information provided by the children. Taken together, this shows infants do encode and retain some memory of their pre-verbal experiences, but only those can provide some verbal information at the time of the initial

92

experience can verbally describe them after a delay. Therefore, it is likely that prior to the acquisition of language, infants and young children must encode information in a nonverbal way (Bauer, Kroupina, Schwade, Dropik & Wewerka, 1998; Hayne & Rovee-Collier, 1995) but these non-verbal memory representations of events cannot be translated into language. However, it is important to note that recent research has found 3-year-olds demonstrated verbal recall of aspects of an event that they could not articulate at encoding when strong contextual support was available (props used during the event were present during recall), but whether this is true memory or online reasoning needs further investigation (Dahl, Kingo, & Krøjgaard, 2015).

Verbal ability at the time of encoding not only affects verbal recall, it also affects non-verbal recall (Bauer & Wewerka, 1995; Simcock & Hayne, 2003). Bauer and Wewerka (1995) found that 12-, 16- and 20-month-olds all demonstrated non-verbal memory for test sequences after delays of 1-, 3-, 6-, 9-, or 12-months. The children's productive language skill at time of encoding the event predicted their verbal memory of the event (mnemonic verbalisations e.g. requesting the "windmill" before it was presented) during recall even after removing the variance associated with age and delay. Verbal ability at time of encoding was also related to later non-verbal production of target actions, though not dependent on it. Simcock and Hayne (2003) used the magic shrinking machine game again, with children aged 2-to 4-years-old, and tested their memory for the event 24-hours later using verbal recall and non-verbally (choosing photographs of the target items used in the game and re-enacting

93

working the machine). Older children reported more about the event verbally, as language skill increased so did verbal recall, but there were no effects of age on behavioural re-enactment. Correlations were also found for verbal skill and non-verbal tests of memory (behavioural re-enactment and photograph recognition).

These studies demonstrate that greater verbal ability at the time of encoding an event leads to later increases in both verbal and non-verbal recall of an event. Bauer and Wewerka (1995) suggest that greater language proficiency at the time of encoding allows children to use others' (e.g. experimenters') language, as well as their own to produce a verbal "script" of the event. This allows for verbal cues to facilitate retrieval. For example, Hayne and Herbert (2004) found that 18-month-old infants that experienced a verbal commentary of the actions and event goals during a demonstration of sequences evidenced superior retention following a 1-month-delay compared to infants that did not. In a follow-up experiment infants were exposed to either verbal cues only at either encoding but not retrieval, or at retrieval but not encoding, however, the infants who received cues at retrieval outperformed those that received the cues during encoding. They suggest that because the same language is used at encoding and retrieval, language as a retrieval cue promotes superior performance.

Conversely, field studies where children engaged in conversations with adults whilst participating in events have shown that aspects of the events not discussed during their occurrence were not recalled later

(Haden, Ornstein, Eckerman, & Didow, 2001; Tessler & Nelson, 1994). Haden et al., (2001) organised three novel events that children participated in with their mothers that involved a wide variety of props. At 2 ½-years-old the children participated in a camping trip event, at 3-years-old they participated in a bird-watching event and finally aged 3 ½-years-old they participated in an event involving setting up and opening an ice-cream shop. They recalled all events with a research assistant after both a 1-day and a 3-week-delay. Videotapes of the child-mother interactions were coded and items that had been attended to during the event were coded as present features, attention could be coded as either verbal (e.g. talking about the items) or non-verbal (e.g. by pointing to or picking up items) by either the mother or child or both 'jointly'. In the memory interviews, items that had been discussed during the sessions by both the mother and the child were significantly more likely to be recalled in children's recall to open questions than were items only the mother verbally alluded to or items that had not been discussed at all.

Similar findings were reported by Tessler and Nelson (1994, study 1), who filmed mothers' interactions with their children (aged 3-to 3 ½-years-old) on a visit to the American Museum of Natural History, one group was asked to interact as normal and the other group of mothers were asked to refrain from elaborate engagement and discussion with their child and told to respond to their questions and comments only. One week later the children were interviewed about their trip to the museum and no children recalled items that they had seen but that had not been discussed during the visit. This finding was replicated in their second

study, 4-to 4 ½ year old children were recorded with their mothers during an event in which they took photographs together, one week later an interviewer came to the child's home and asked them about their picture taking experience. No recall was demonstrated by children of instances that had not been discussed during encoding, however, with this group of older children the discussion did not have to be joint; if either mother or child verbalised something then it could be recalled.

Tessler and Nelson (1994) suggest that parents and other experienced narrators highlight the salient aspects of the present which focuses children's attention to these aspects, providing a model for children to learn what aspects of events to focus on. Discussion keeps attention on these salient points which may mean they are encoded more elaborately and more likely to be stored with multiple verbal 'tags' that helps children use their own and others' language to aid recall of these experiences. The importance of children learning narrative structure from parents or other adults has been shown repeatedly, whereby the way parents discuss the past with their children shapes the way their children come to talk about their past (e.g., Bergen, Salmon, Dadds & Allen, 2009; Farrant & Reese, 2000; Fivush, 2014; Fivush & Fromhoff, 1988; McCabe & Peterson, 1991; Morris, Baker-Ward & Bauer, 2009; Peterson Morris, Baker-Ward & Flynn, 2013; Reese, 2002; Reese, Haden, & Fivush, 1993). Children have been shown to internalise their mothers' narrative style, children's mothers' narrative style was assessed as wither narrative (elaborative and dynamic) or paradigmatic (descriptive) and children reflected this narrative style in the way they describe their experiences

both when interacting with their mother and with a researcher (even when the researcher used the opposite narrative style to the child's mother) (Tessler & Nelson, 1994). There are further benefits for children whose mothers use an elaborative narrative style when conversing with them, they can recall earlier memories than children whose mothers had used a less elaborative style (Jack, Macdonald, Reese & Hayne, 2009; Reese, Jack & White, 2010) and provide more coherent narrative accounts than children of parents who used a pragmatic style (Fivush & Fromhoff, 1988).

Language also allows us to organise our experiences into a coherent form, aiding retention of entire experiences rather than fragments. As causal understanding of the narrative structure of an event improves, memory for the event has been shown to also improve (Fivush, Haden & Adam, 1995; Fivush & Hamond, 1992; Morris, Baker-Ward & Bauer, 2010; Trabasso & Shu, 1993). The superior performance of older children at recalling episodic memories can be thus attributed in part to increased understanding of the relationships between the individual components of events. Pillemer, Picariello and Pruett (1994), interviewed 3 ½- and 4 ½- year-olds about their experience of an emergency evacuation at pre-school, both 2 weeks later and after seven years. Only the older children, whose first interview narratives showed evidence of temporal and causal understanding of the evacuation at the time, were able to produce an intact narrative account of the event after a seven-year delay. This suggests that comprehension of an experience aids its

retention; the children who did not show comprehension of the event when they experienced it were not able to recall it at a later date.

3.1.1 *Summary*. Instead of merely allowing children to express their memories, language development supports the development of verbally accessible memory (Cordón et al., 2004; Simcock & Hayne, 2002; 2003). While children have the ability to express their pre-verbal memories behaviourally (Simcock & Hayne, 2002, 2003) or in recognition memory tasks (Peterson, & Rideout, 1998; Simcock & Hayne, 2002) they remain unable to provide narratives of their pre-verbal experiences, even when they have developed the relevant vocabulary to do so (Simcock & Hayne, 2002). Non-verbal encoding during this period of life (Bauer et al., 1998; Hayne & Rovee-Collier, 1995) does not appear to be able to be translated into verbal recall. Language skill at the time of encoding an event has been shown to correlate with verbal memory (Bauer & Wewerka, 1995; Simcock & Hayne, 2003), and non-verbal memory of the event (Bauer & Wewerka, 1995; Simcock & Hayne, 2003). It is likely that using language during encoding allows children to use language cues during retrieval which in turn aids recall (Hayne & Herbert, 2004).

During the development of narrative ability the role of conversational partner is an important one because children internalise and mimic the narrative style of their mothers when they begin to discuss the past (Bergen, Salmon, Dadds & Allen, 2009; Farrant & Reese, 2000; Fivush, 2014; Fivush & Fromhoff, 1988; McCabe & Peterson, 1991; Morris, Baker-Ward & Bauer, 2009; Peterson Morris, Baker-Ward &

Flynn, 2013; Reese, 2002; Reese et al., 1993), even when later

conversational partners use the opposite narrative style (Tessler &

Nelson, 1994). Further, children whose mothers' use more elaborative

narrative styles produce more coherent narratives (Fivush & Fromhoff,

1988) and earlier first memories (Jack et al., 2009; Reese et al., 2010)

than children whose mothers used less elaborative narrative styles. More

coherent event narratives aid memory retention through increasing

elaboration and understanding of an event and the relations of its

components to each other, which in turn aids retention of the event itself

(Pillemer et al., 1994).

3.1.2 *Forensic interviewing implications*. Due to the

importance of language as a retrieval cue, it is unsurprising that when

children are asked complex questions that are developmentally too

advanced (by the way they are worded or the content they contain), this

can undermine children's recollections. When children are asked

questions that are too complicated for them to understand, they do not

have the opportunity to express the information they may possess. For

example, Saywitz and Camparo (1998) describe a court case in which a

pre-schooler denied seeing a weapon at a crime scene, but answered

'yes' when asked if he saw a gun. In addition to failing to obtain

knowledge children have, asking questions beyond children's

developmental level can further undermine their competency because

they tend to answer questions even if they do not understand what they

are being asked (Hughes & Grieve,1980; Pratt,1990).This may be in part

due to the fact that children have difficulty when it comes to

99

comprehension monitoring; that is, they may not yet have the skills necessary to identify when they do not understand the question they are being asked, particularly when they are being asked complex questions (Perry, McAuliff, Tam, Claycomb, Dostal & Flanagan, 1995).

Researchers have attempted to explore the effects of teaching children to engage in comprehension monitoring (CMT) by showing them how to identify confusing questions. Peters and Nunez (1999) found that when CMT was coupled with TDT (task demand training) which emphasises that adults make mistakes and children should tell the interviewer when they don't know, children requested rephrasing more than children trained in the TDT alone group. While this would likely be too lengthy a process to undertake with children prior to a forensic interview, ground rules, such as telling the child to let the interviewer know if they do not understand a question, should be communicated in an effort to encourage children to request clarification when they do not understand questions (NICHD protocol, Lamb, et al., 2007a; Scottish Executive, 2003, 2011). In addition, when interviewing children, age-appropriate language should always be used in order to avoid inhibiting or confusing children's recall and children should be allowed to tell interviewers what happened using their own words (NICHD protocol, Lamb, et al., 2007a; Scottish Executive, 2003, 2011).

3.2 Suggestibility

The vulnerability or proneness to 'yield' to suggestive influence is known as suggestibility (Gudjonsson, 1984). Suggestive influence has been shown to have an effect on the content and accuracy of both children and adults' memories, though there is evidence that children are more suggestible than adults and younger children more suggestible than older children (Bruck & Ceci, 1999; Ceci & Bruck, 1993; Ceci, Ross & Toglia, 1987; Cohen & Harnick, 1980; Goodman & Aman, 1990; Goodman, Jones & McLeod, 2017; Goodman & Melinder,2007; King & Yuille, 1987; Paz-Alonso & Goodman, 2016; Volpini, Melis, Petralia & Rosenberg, 2016). Therefore, it is important from a legal standpoint for interviewers and decision makers to be aware of what factors can elicit inaccurate testimony from children. Suggestibility is negatively correlated with greater cognitive capacities, such as source monitoring ability (Giles, Gopnik, & Heyman, 2002), pre-existing knowledge about experienced events (Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1997), cognitive inhibitory control (Melinder, Endestad & Magnussen, 2006), theory of mind (Bruck & Melnyk, 2004; Welch-Ross, 1999; Welch-Ross, Diecidue & Miller, 1997), verbal ability (Curci, Bianco & Gudjonsson, 2017) and intelligence (Ceci & Bruck, 1993; Chae et al., 2011). Cognitive factors, that are still under developed in children and improve with age, may in part explain age differences in suggestibility, with pre-schoolers seeming to be the most vulnerable to misleading information (Bjorklund et al., 2000; Ceci & Bruck, 1993).

101

Social factors have also been implicated in suggestibility, an individual may acquiesce or 'go along with' suggested information for social reasons, such as feeling under pressure or wanting to appear helpful, however, the memory for the original event is not altered (see McCloskey & Zaragoza, 1985, for a discussion). It is important to note that under certain conditions no such age differences and even "reverse developmental effects" where conversely older children and adults, have been more susceptible to certain types of suggestions than younger children have been found (Brainerd, Reyna & Ceci, 2008; Kim, Kwon & Ceci, 2017; Otgaar, Howe, Peters, Sauerland & Raymaekers, 2013). A review of all the suggestibility literature is beyond the scope of this thesis, therefore, this section will focus on research that has investigated suggestibility in children in a manner that more closely resembles contexts that have elicited false reports of abuse.

Research has shown that the way children are questioned has an effect on the accuracy of their testimony; leading questions are more likely to elicit inaccuracies, produce contradictions and allow interviewees to go along with misleading information (Brady, Poole, Warren, & Jones, 1999; Dent, 1986; Dent & Stephenson, 1979; Oates & Shrimpton, 1991; Waterman, Blades & Spencer, 2004). Leading questions by their very nature introduce information and can be used to 'lead' an interviewee in a particular direction. Loftus and Palmer (1974) demonstrated that manipulating the wording of questions had an effect on the answers given by adult participants who viewed an identical film of an automobile accident. Suggestive questions imply a particular response is expected

from an interviewee, and may use some degree of pressure or feedback to get the interviewee to 'go along with' what the interview is saying (Baxter, Charles, Martin & McGroarty, 2013; Bruck, Ceci, Francoeur & Barr, 1995b; Leichtman& Ceci, 1995; Thompson, Clarke-Stewart & Lepore, 1997). Interviews that have relied heavily on leading and suggestive questioning and suggestive techniques have been shown in the real world to elicit false accounts of child abuse (Duke, Uhl, Price & Wood, 2016; Kitzinger, 2000; Schreiber et al., 2006).

Some experimental studies have questioned children about their interactions with adults that have involved innocuous physical touching, in an effort to determine whether children readily adopt suggestions of touching that could be construed as abusive. Following play sessions where they interacted alone with a male confederate, Goodman and Aman (1990) interviewed 3- and 5-year-old children after a one-week delay and Rudy and Goodman (1991) interviewed 4-year-old and 7-year-old children following a 10- to 12-day delay. The children were asked free-recall, and specific and misleading questions (a subset of which were abuse related e.g. "Did he touch your private parts?" or "He kissed you, didn't he?"). Age-related differences in suggestibility were apparent, with older children giving more correct answers and fewer commission errors to specific questions and fewer errors to misleading questions than the younger comparison age group. Children of all ages answered the misleading 'abuse' questions correctly the majority of the time (3-year-olds; M =.76; 5-year-olds) and commission errors consisted of children nodding their head without verbally elaborating (Goodman & Aman,

103

1990), with no commission errors made by Rudy and Goodman's participants. When children were allowed to freely recall the events no children ever falsely recalled any sexualised acts.

Either one week or one month after they had received a doctor's examination, Saywitz et al., (1991) interviewed 5- and 7-year-old girls about the appointment. Half the children in each age group received a routine examination of the exterior of the vagina and anus (genital condition) while the others received an examination for scoliosis instead (non-genital condition). There were no age differences in response to specific questions and commission errors were extremely rare. There were effects of age and delay on suggestibility, older children and children interviewed after one week answered a higher proportion of misleading questions correctly than younger children and children interviewed after one month, however, the overall rate of commission errors made to misleading 'abuse' questions was extremely low (M = .01). When asked for their free recall, no children in the non-genital condition claimed to have been touched on either the vagina or anus, instead, of the children that truly had been touched 81% failed to report vaginal and 89% failed to report anal touching.

These studies showed few children accepted false suggestions of things that did not really happen (Goodman & Aman 1990; Rudy & Goodman 1991; Saywitz, Goodman, Nicholas & Moan, 1991). However, different interviewing styles e.g. persistent suggestive interviewing that utilises lots of leading and suggestive questions, have produced

104

dramatically different results. For example, Leichtman and Ceci (1995) interviewed 3- to 6-year olds once a week for 4-weeks about a 2-minute-long visit of a confederate known as 'Sam Stone' to their nursery, that had consisted of Sam greeting the children, walking around the classroom and then waving goodbye. Children were assigned to one of four interview conditions; some had been told stories about Sam being clumsy and breaking toys before his visit, some were shown a torn book and a soiled teddy bear after Sam's visit, some received both the pre- and post-interview suggestions and some received neither. The children that had not been told anything about Sam previous to his visit, were either interviewed in a neutral manner (the control group) or suggestively following the event. The children that had been given the pre-event information about Sam being clumsy were interviewed in a neutral manner and the children that received the stereotype and suggestion were interviewed suggestively.

Ten weeks after the original event the children were re-interviewed, when asked if they saw Sam Stone do something to the teddy and book the control group were the least suggestible group with an overall accuracy rate of 96%. Accuracy rates decreased linearly as the number of suggestive influences increased; the stereotype only group were accurate 83% of the time, the suggestive interview group were accurate 72% of the time, and the children that received both the stereotype and suggestive interview evidenced the lowest accuracy rate (64%). As was the case in the aforementioned studies, no children in the control group or the stereotype only group claimed any false events had

105

happened during free-recall or even during direct questioning. Children in both the suggestive interviewing group and the suggestive interviewing plus stereotype group claimed in free-recall, as well as specific questioning, that they had witnessed Sam doing something to the book and/or the teddy.

Interviewing style has even been shown to affect the accuracy of children's recollections of things they have truly witnessed. For example, Thompson et al., (1997) set up an interaction where 5- and 6-year old children were left alone in a room with a confederate 'Chester the janitor ' who either cleaned the room and some toys, cleaned the room then played with the toys, or just played with the toys. Following a one-hour delay, children were interviewed twice in either a neutral, an incriminating ("You know he sometimes stops working and plays with the toys") or an exculpating ("These toys always get dirty, he is supposed to clean them all over") manner, using leading questions in the same direction as the interview condition. Parents then interviewed their child in a neutral manner about what they had witnessed. After a one week later delay all children were interviewed in a neutral manner, children's free recall accounts of what they had seen were consistent with the style they had originally been interviewed in. Children that had always been interviewed neutrally gave accurate accounts of what they had seen, children in the incriminating condition described the janitor's actions in a manner more like playing and children in the exculpating interviews described the janitor's actions like cleaning. Children that saw the janitor play most frequently made errors biased toward cleaning when they were in the

exculpating condition, conversely, children that saw the janitor clean more frequently made errors suggesting playing when they were in the incriminating condition. Notably, the exculpating condition (where the interviewers suggested the janitor had cleaned) had a stronger effect on children's reports than the incriminating interviews, this is important because it demonstrates that children were less easily swayed to falsely accuse Chester of wrongdoing than they were to falsely deny his misbehaviour.

Misleading post-event information has even been shown to affect children's suggestibility for physical events involving their own bodies when they are interviewed in a suggestive manner on repeated occasions. After receiving a vaccine, Bruck et al., (1995b) exposed five-year-old children to four 45-minute-long suggestive interview sessions, where they received repeated feedback and misinformation about the event. The suggestive interviewing phase was conducted over a 2-week period on average 11 months after the inoculation. Children were given "positive ' (they acted brave and had not cried) or "neutral" (no feedback) feedback about how they acted during the inoculation and were either given misleading or no information about who had administered the shot and who had read them a story. At the fourth session children that had been given positive feedback reported less hurt and less crying than children in the neutral group despite there having been no difference in ratings at week one. Children who were given misleading information were more likely to make false allegations than children who were not given this information (e.g. that an RA had given them a shot when it was

their paediatrician and that the paediatrician had read a story, when it was the RA).

Contrary to the earlier discussed studies these results would suggest children are extremely vulnerable to misleading and suggestive questioning. However, the study design undoubtedly affects children's degree of suggestibility and there are a number of factors that likely explain why children displayed elevated rates of suggestibility in the latter few studies. Firstly, lengthy delays were imposed between the event and test sessions in some of the studies e.g. approximately a year between event and test in the Bruck et al., (1995b) study. Memory traces degrade with time, so longer delays between the event and testing session make suggestions easier to accept because the real memory trace is weaker (Saywitz et al., 1991). Secondly, in between event and testing sessions children engaged in repeated suggestibility and misinformation exposure sessions (e.g. Bruck et al., 1995b; Leichtman & Ceci; 1995). In some cases children were not prepared initially to falsely assent, however, the sheer number of times they were exposed to untrue aspects of the event may have come to make them doubt their own representations. For example, even though children eventually yielded to suggestion in the Bruck et al., (1995b) study, an earlier phase of the study where children were given immediate positive "that didn't hurt", negative "that hurt" or neutral "the shot is over" feedback about their own reaction to the vaccine did not produce group differences in rating of pain or how much they had cried following the vaccination after a one week delay. Even pre-school children resisted attempts to sway their reports when they were exposed

108

to suggestive influence on only one occasion. Thirdly, the interviews were of a highly suggestive nature (Leichtman & Ceci; 1995; Thompson et al., 1997), a nature that critics claim is unrealistic to the level of suggestion than would be the case in standard forensic interviews (Lyon, 1999). Finally, the main finding of the Bruck et al., (1995b) study, that children supplied with misinformation were more likely to make false claims than children not exposed to misinformation, was based on aspects of events that have been shown to be not well remembered by children. Children confused actions performed by their paediatrician and by an RA that had been performed in their paediatrician's practice. At the first session of phase 2 children were asked to pick the photograph of their paediatrician and of the RA study from a group of distractors. All children selected the correct photo of their paediatrician but only 35% of the children selected the correct picture of the RA. The doctor was highly familiar to the children, whereas they did not remember the RA even before they were given any misinformation. Children more often chose the paediatrician as the actor that had carried out the 6 target actions than the RA, this may reflect general knowledge that when they go to the doctor office that it is the paediatrician there and may have chosen him more often than the RA as in absence of a true memory as it is the more likely option. Suggestions may be more readily adopted when they are consistent with the child's pre-existing schemas and expectations about an event than when they are not (Bartlett, 1932; Thompson et al., 1997).

*3.2.1 **Summary**.* The degree to which children are vulnerable to suggestive influence depends on a number of factors. How children were questioned affected how likely suggested information was to appear in children's reports, when children were allowed to provide free-recall of their experiences before any misleading information was provided by interviewers, free-recall never elicited any reports of false touching (Goodman & Aman 1990; Saywitz et al., 1991). However, specific and misleading questions (Goodman & Aman 1990; Rudy & Goodman, 1991; Saywitz et al., 1991) and highly suggestive and biased interviewing elicited false accusations from children (Bruck et al, 1995b; Leichtman& Ceci; 1995; Thompson et al., 1997). Age has also been shown to affect children's ability to resist suggestion, with older children in general showing a tendency to answer more specific (Goodman & Aman 1990; Rudy & Goodman, 1991) and misleading (Goodman & Aman 1990; Saywitz et al., 1991) questions correctly and make fewer commission errors than younger children. However, when misleading abuse related questions were kept to a minimum and children were only interviewed once, commission error rates were impressively low for 5-year-olds (Goodman & Aman, 1990; Saywitz et al., 1991) and non-existent for 4-year-olds (Rudy & Goodman, 1991) and 7 year olds (Rudy & Goodman, 1991; Saywitz et al., 1991). Finally, the number of suggestive techniques used and the degree to which they are suggestive also affect children's ability to resist suggestion. When children were provided with multiple and repeated pre-interview suggestions and stereotypes, they provided false allegations even during free-narrative (Leichman & Ceci; 1995) and

110

when interviewed repeatedly in a biased manner produced reports that consistently converged further in the direction of the interviewers' biases (Bruck et al., 1995b; Leichtman& Ceci; 1995; Thompson et al., 1997).

   3.2.2   *Forensic interviewing implications*. Memory is a reconstructive process (Bartlett, 1920; Conway & Playdell-Pearce, 2000; Loftus, 2005; Pezdek & Lam, 2005; Pollio & Foote, 1971; Roediger & McDermott, 1995; Rubin & Umanath, 2015) and recalling an experience shapes subsequent remembering (Hirst & Manier, 2008). This means that whenever a memory is recalled it is reconstructed, making it particularly vulnerable to suggestive influences encountered during recall, as these can affect the original memory trace through the incorporation of errors and biases (Bartlett, 1932). Free recall usually elicits the most accurate responses when children (and adults) are asked to recall their experiences (Lamb et al., 1996a; Sternberg et al., 1996; Lamb et al., 2007b). Therefore, children should be encouraged to provide narrative accounts in their own words before any specific or leading questions are asked to avoid errors being incorporated into their memory (NICHD protocol, Lamb, et al., 2007a; Scottish Executive 2003, 2011). Interviewers must take care when asking specific and leading questions, even if they are not obviously suggestive, because introducing specific information into the interview, particularly when is more believable than implausible (e.g. that a child was spanked by a doctor), is more likely to become incorporated into children's accounts as it fit with general expectations and schemas children already have of events (Bartlett, 1932; Thompson et al., 1997). Interviewers should make it a priority to

111

find out about any prior disclosures and discussions children have had about the alleged abuse with other people before the forensic interview (NICHD protocol, Lamb, et al., 2007a) as discussions with others can lead to post-event information being incorporated into the child's memory by which point even using a free-recall account can still have misleading and suggested information present (Warren & Lane, 1995). This is also the case when children have been exposed to suggestive interviewing, any future free recall account tainted after suggestive interviewing (Leichman & Ceci, 1995; Thompson et al., 1997).

3.3   *Source monitoring*

Another cognitive ability that affects memory and has important consequences in forensic interviewing is source monitoring. Source monitoring is the ability to judge the origin, or source, of information such as our memories, beliefs and knowledge (Braha-Zeitoun et al., 2015; Johnson, Hashtroudi & Lindsay, 1993). Judgements of the origin of our memories may have to be made between external and internal sources, also known as 'reality monitoring' (Johnson & Raye, 1981). This involves judging whether a memory is self-generated or perceptual, for example, 'did I say something or was it somebody else?' (Foley, Raye & Johnson, 1983). Discriminations may also have to be made between two different external sources that a memory could have originated from (e.g. was it Steven or Laura who told me that Mark is moving house?) or between two internal sources (e.g. what one thought versus what one said; Johnson et al., 1993). Source monitoring is important in the forensic context because particularisation (specific contextual details such as time

and place something happened) is expected when attempting to prosecute a crime in order to allow the defendant to provide an alibi and to provide leads that the police can use to gather evidence to support an allegation (Guadagno, Powell & Wright, 2006). Also, children often experience repeated instances of abuse (Connolly & Read, 2006) and being able to provide accounts of specific instances instead of a generic 'script' of combined experiences enhances children's credibility (Brubacher, Powell & Roberts, 2014).

Significant developments in source monitoring ability happen between the ages of 3- and 8-years-of age (Roberts, 2002; Sluzenski, Newcombe, & Ottinger, 2004). The ongoing development of source monitoring ability throughout childhood does not mean that children exhibit difficulty judging between all sources of their memories; instead they have difficulty making particular source judgments yet have demonstrated performance comparable to adult levels in other source monitoring tasks (Foley & Johnson, 1985; Foley et al., 1983; Foley, Santini & Sopasakis, 1989; Lindsay, Johnson & Kwon, 1991). For example, Foley et al., (1983) found although 6-year olds had trouble distinguishing between words they had said aloud from words they had only imagined themselves saying, their performance was equivalent to adults' when distinguishing between words they said aloud and words they had heard another person say aloud. Similarly, Foley and Johnson (1985) found that 6 year olds were as able as adults to discriminate the source of their memories for actions they performed compared to actions they had watched another person perform. However, 6- and 9-year-olds

performed more poorly than adults when distinguishing the source of their memories for things they had really done from things they only imagined themselves doing. The children were more likely to say they had actually performed the action that they had only imagined themselves performing rather than vice versa.

Further illustrating this finding, Lindsay et al., (1991), had subjects perform actions (e.g. touching their nose) or observe a video of an actor performing the actions, and then either imagine themselves performing actions or imagine the actor performing those actions. They found 7-10 year-olds had difficulty distinguishing the source of their memories when attempting to discriminate between actions that had really been performed and actions that had been imagined only when the actor was the same, that is, when they had both seen and imagined the actor performing actions or when they had performed actions and imagined themselves performing actions. The children performed at an equivalent level to adults when judging between actions performed by the other person and actions they imagined themselves performing and judging the source of memory for actions they themselves performed and actions they imagined the other person performing. These studies have shown that children can discriminate between the correct source of their memories for things they have done or said compared with memories of what other people have done or said (Foley & Johnson, 1985; Foley et al., 1983; Lindsay et al., 1991). It is thought that the self is well differentiated from others by 6 years of age (Foley et al., 1983), therefore, cues that discriminate the self versus others may aid children in

differentiating between memories of things they have done or said compared to things another person has done or said. Indeed, see recent research by Hala, Brown, McKay and Juan (2013) that found even 2 ½-year-olds could accurately judge the source of memories for actions performed by themselves or an experimenter.

However, the ability to distinguish between memories for thoughts versus actions have not yet developed (Foley et al., 1983). Incorrectly identifying imagined events as being experienced happens because the mental representations formed when imagining an event contain qualitatively similar properties to the representations formed when actually viewing or participating in the event (Johnson et al., 1993; Johnson & Raye, 1981). Both the reality monitoring framework (Johnson & Raye, 1981) and source monitoring theory (Johnson et al., 1993) stipulate that judgments about the source of a memory are made during memory retrieval, some of these judgements are quick and automatic while some employ more effortful decision making procedures. The quality of memories for experienced events are often rich in temporal, spatial and affective characteristics and may be easily and quickly attributed as having been experienced. An imagined event on the other hand lacks vivid details but includes many cognitive operations (e.g. imagery), coupled with general knowledge these features will be used to judge the source of such memories, though it may require more conscious and painstaking judgement and decision making processes. However, sometimes the content of an imagined memory can be rich and the mental operations involved in processing the imagined event can be

115

the same as those of a real event, reducing the amount of discriminating cues available at retrieval, making it possible the imagined event is quickly judged as being an experienced event.

Confusing the source of memories for actions performed with actions one has imagined performing has lead children to believe they have actually performed these actions (Foley & Johnson, 1985). There is evidence to show that children have even come to believe they have experienced entire events when really they have only been instructed to repeatedly imagine them (Ceci, Crotteau, Smith & Loftus, 1994a; Ceci, Loftus, Leichtman & Bruck, 1994b). Ceci et al. (1994a) instructed 3-to 6-years-old to 'think real hard' about a list of events, some of which had really been experienced by the children and some had not (e.g. getting their hand caught in a mousetrap and going to hospital to have it removed) and this procedure was repeated 7-10 times. Despite interviewers warning children that some of the events had not really happened to them, 34% of the children claimed the false events had really happened to them, even providing narratives about the fictitious experience at the final session. Ceci et al., (1994b) repeated the procedure, this time telling children the false events really had happened to them and providing instruction to visualise those events. Initially 29% of children claimed the false events to be true, however, at the final session, following repeated visualisations this figure rose to 43%.

Adults are not immune to the effects of confusing multiple sources, in fact, they exhibit some of the same difficulties that children do, for

116

example, asking adults to imagine an event increases the chance of them believing that they actually experienced that event (Garry & Polaschek, 2000; Garry, Sharman, Wade, Hunt & Smith, 2001). Garry, Manning, Loftus and Sherman (1996) presented a checklist of childhood events, such as 'having a lifeguard pull you out of the water' and asked participants to rate how confident they were that each event had happened to them in the past. At a follow-up session the participants were asked to imagine a randomly selected subset of the events from the checklist. Following the imagination period, participant's confidence that each item had happened to them was measured again, under the guise that the participant's original confidence ratings had been misplaced. Participants became more confident that events that had not really happened to them had in fact occurred following the imagining exercise and this effect was only found for events that participants had not been initially confident occurred. Therefore, simply imagining events that had not really happened to them inflated participants' confidence so greatly that they came to believe the events had actually occurred.

When the cognitive operations used to create representations of events are similar, people may mistakenly judge imagined events as being truly experienced events (Ceci et al., 1994a; Ceci et al., 1994b; Garry et al., 1996). It is therefore unsurprising that when the perceptual qualities of truly experienced events are similar this affects one's ability to discriminate between these events (Brubacher, Glisic, Roberts & Powell, 2011a; Brubacher, Powell & Roberts, 2014; Connolly & Lindsay, 2001; Danby, Brubacher, Sharman, Powell & Roberts, 2017; Murachver at al.,

1996). When children are requested to recall details that occurred in a specific occurrence of a repeated experience, fixed items (details that are the same every time) are well remembered while variable items (details of repeated items that change in different occurrences, e.g., ordering a cheeseburger on one trip to McDonald's and Chicken McNuggets on another) are problematic and commonly misattributed to incorrect time (Brubacher et al., 2011a; Connolly & Lindsay, 2001; Powell, Roberts, Ceci & Hembrooke, 1999). For example, Murachver et al., (1996) found that 5- and 6-year-old children that experienced 3 exposures to a play event 'visiting the pirate' recalled more of the core actions that occurred in invariant activity sequences (sequences that were the same every time) than actions that were variable within the sequence (e.g. during each pirate visit an instrument was played, however the instrument was different at every session e.g. drum, cymbals, bell).

Greater recall of fixed occurrences compared to variable ones reflects young children's reliance on "scripts", incorporating and reporting what usually occurs into their memory representations and narratives, using the general "you" and present tense ('You do x'); instead of focussing on details of specific experiences (Hudson, Fivush & Kuebli, 1992). Scripts allow children to form predictions of the future based representations of past repeated experiences, even very young children have general well-organised scripts and the organisation of these scripts is qualitatively similar to those of older children and adults, though usually less complete (Hudson et al., 1992). According to script theories such as the schema confirmation-deployment model (Farrar & Goodman, 1990),

118

children should be able to recall new items (unique to a specific occurrence) because they are inconsistent with an already formed script of a repeated experience and so are episodically encoded, leading to enhanced memory for these details. Brubacher et al., (2011a) questioned 4-5 and 7-8-year old children about a play event they had experienced four times, some of the actions/items they experienced were 'variable' (e.g. the story read changed every time), some were 'fixed' across 3 occurrences (hi) but changed on one occasion (lo) (e.g. children used hand sanitiser in three occurrences and a fan in one to 'get refreshed'), and some were 'new', occurring only during one event. All children demonstrated greater source memory for 'hi' details that occurred across 3 fixed occasions than for all other low frequency details (variable, 'lo' and 'new' details). However, children's source attribution of the 'new details' were more accurate than 'lo' and 'variable' details, particularly for the older children who correctly attributed new details to the correct source 73% of the time.

Powell, Roberts, Ceci and Hembrooke, (1999) compared 3-to 8-year-old children's memories for a play event they had either experienced only once or 6 times. After participating in the event(s) they then took part in a biasing interview where the interviewer asked if certain instantiations had occurred during the activities, half of which were false suggestions. At a final recall interview, children that experienced the event repeatedly were more accurate at recalling truly experienced fixed instantiations and were less likely to report the false suggestion instantiations than children that only experienced the event once. However, when asked about items

119

that had varied across occurrences (the single experience children were questioned about a suggested instantiation as they had no other comparison experiences), the children that had experienced the repeated event were more prone to making source monitoring errors of an internal nature (mistakenly attributed variable items to the incorrect instance of the event). Therefore, repeated experiences of an event has both bolstering and detrimental effects on memory (also see Price, Connolly & Gordon, 2016), fixed items/occurrences are well remembered but the source for items/occurrences that varied between events are confused. Importantly, even though children confused the source of the memory they did not accept or generate incorrect information in their reports.

3.3.1 Summary. Significant developments in source monitoring ability happen between the ages of 3- and 8-years-of age (Roberts, 2002). While developmental trends in the ability to accurately discriminate between sources are evident, the development of source-monitoring skills is not uniformly linear; children gain competence at some types of source before other types (Foley & Johnson, 1985; Foley et al., 1983; Foley et al., 1989; Lindsay et al., 1991). Children are able to accurately discriminate memory sources for things they had done (actions and speech) when compared with things another person had said or done (Foley & Johnson, 1985; Foley et al., 1983). However, children have trouble discriminating the source of their memories when they had imagined themselves saying words and when they had really said them (Foley et al., 1983) and when they had performed actions and imagined themselves performing the action (Foley & Johnson, 1985) and were

120

more likely to say they had actually performed the action they had only imagined performing than vice versa (Foley & Johnson, 1985). Imagining or visualising events has lead children to believe that they had in fact experienced the non-event (Ceci et al., 1994a; Ceci et al., 1994b). The reality monitoring framework (Johnson & Raye, 1981) and source monitoring theory (Johnson et al., 1993) state that this confusion can happen through different processes. Firstly, imagining an event may produce a representation of that event that has similar vivid qualities to that of a real event, and quick automatic source judgments made during retrieval judge such memories as really having been experienced. On the other hand, the cognitive operations involved in imagining the event coupled with general knowledge or the general likelihood of the memory being real, may cause it to be judged as so despite more conscious judgment processes applied during decision making. When similar experiences occur the fixed details that occur at every instance are better remembered than details that are variable and change at different instances (Brubacher et al., 2011a; Connolly & Lindsay, 2001; Murachver et al., 1996; Powell et al., 1999). However new details, that only occur during one instance of a repeated event are particularly well remembered, especially by older children (Brubacher et al., 2011a) as they do not fit the current schema of the event and are therefore encoded episodically (Farrar & Goodman, 1990).

3.3.2 Forensic interviewing implications.

Source monitoring ability is still developing throughout childhood (Roberts, 2002), therefore, children may make more source errors than adults (Foley & Johnson,

121

1985; Foley et al., 1983; Foley et al., 1989; Lindsay et al., 1991). However, forensically it is important to note that source errors do not automatically equal inaccurate memory (Powell et al., 1999), in fact, the reverse has been demonstrated (Ratner, Foley & Gimpert, 2000). Four-year-olds traced and imagined tracing pictures from either a goal related story book with text or the same pictures in isolation with no story. Children in the picture only condition better discriminated whether they had traced or imagined tracing the pictures, however, the children that heard the story whilst completing the tracing exhibited greater memory of the pictures (Ratner et al., 2000). The presence of a goal aided memory but hindered source monitoring (see also Foley, Passalacqua, & Ratner, 1993; Foley & Ratner, 1998; Foley, Ratner, & House, 2002). Further, children that confused the source of their memories of multiple incidents rarely reported entirely false information (Powell et al.,1999). Techniques interviewers can use to minimise source errors in children's reports are asking open questions as children make fewer source errors when they are allowed to freely recall events (Brubacher et al., 2014; Roberts & Blades, 1998, 1999) and in the case of children that have experienced repeated instances of abuse asking about unique aspects of repeated experiences can elicit new and accurate details, particularly from older children (Brubacher et al., 2011a). Finally, suggestive interview techniques that present alternative representations of what the child has experienced can cause source errors to be incorporated into children's testimony during an interview (Roberts, 2002) and should always be avoided. This can include, but is not limited to, techniques that involve

122

imagining, visualising and the presentation of post-event information to children.

3.4 Delay

Children often delay disclosure of abuse (Finkelhor, 1984; Goodman-Brown, Edelstein, Goodman, Jones & Gordon, 2003; Kellogg, 2017; London, Bruck, Wright & Ceci, 2008; Malloy, Brubacher & Lamb, 2011;McElvaney, 2015; Pipe, Lamb, Orbach, Sternberg, Stewart & Esplin, 2007), consequently, by the time children are interviewed (Goodman et al., 1992) and testify in court (Bruck et al., 1995b; Cashmore, 1994; Oates, Lynch, Stern, O'Toole & Cooney, 1995) often weeks, months, or even years have passed since the abuse occurred. Therefore, it is important for interviewers to understand how delay affects children's memory for their experiences. Memory decay, or forgetting, occurs rapidly after experiencing an event (Hardt, Nader & Nadel, 2013; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991), this can potentially have a serious impact on the amount and accuracy of the information recalled. The following section will discuss research that has investigated the effect of varying delays on children's recall of their experiences. In an effort to render the findings applicable to investigations of child abuse, the events reviewed in this section focus on recall of naturally occurring potentially stressful/traumatic events or events that involved the children's bodies being touched.

Children's routine medical examinations have provided an avenue for researchers to examine the effects of delay on memories for both

123

bodily touching and stressful experiences because these events involve procedures such as drawing blood. Ornstein et al., (1992) examined 3- and 6-year-olds recall of a check-up with their paediatrician that included such a procedure. Children were interviewed immediately post-examination and again after either a 1-week-delay or a 3-week-delay. Children of both ages remembered most of the features of the check-up at the immediate memory test, though older children recalled a greater number (92%) of the overall features than younger children (82%). The performance of the 3-year-olds significantly decreased over delay intervals of 1- and 3-weeks (73% and 67% correct respectively), whereas that of the 6-year-olds remained constant over this period. Both age groups produced 'new' information at their second interview that they had not recalled at their first interview with no difference between age group or delay condition. Due to the standardised interview format some children were asked questions about standard features of a check-up that they had not actually experienced, therefore, some questions were misleading. After 1-week the 3-year-olds evidenced a significant increase in the percentage of assents to misleading questions than they had made originally, whereas for the 6-year-olds the increase was only significant after 3 weeks.

Using this same procedure again with slightly older children (4- to 7-year-olds) Ornstein et al. (2006) elicited recall immediately post-examination and again after a 6-month-delay (half of the children were also given an extra interview 3 months after the examination). Age affected the number of features of the event recalled immediately, with

124

older children (6- and 7-year-olds) recalling more than younger children (4- and 5-year-olds). However, delay affected recall scores similarly for all age groups, with recall scores declining between the immediate and 3-month interviews, but no further decline from 3 to 6 months. Overall the results of these studies show delay decreased the amount of information recalled and increased the number of errors made, however, children were still able to recall a great deal of what they had experienced accurately following a delay. Ornstein et al., (1992) reported an age effect, with younger children's recall decreasing to a greater degree following the delay than older children's, while Ornstein et al., (2006) reported similar recall decreases in older and younger children. Possibly the different length of the delay used in the studies contributed to these contradictory findings, shorter delays (1-3 weeks) were enough for younger children to show forgetting while it was not until the passage of 3-6 months that older children began to forget to the same degree than the younger children did.

Peterson and her colleagues studied children's memory for naturally occurring, personally salient and upsetting (injuries they had sustained and the subsequent hospital treatment they received) experiences over various delays. Peterson and Bell (1996) elicited recall from children aged between 2- and 13-years-old, regarding their injury and hospital treatment, after delays of a few days and again after 6 months. Children recalled more information and were more accurate at the first interview than 6-months later, though the majority of information provided by children was accurate. Older children (9- to 13-years-old)

recalled more information than younger children and made so few errors that their data could not be included in analyses. Children made more errors when talking about the hospital event rather than the injury itself, while the details of the actual injury were well remembered. Peterson (1996) interviewed 2- to 9-year-olds about their memories of sustaining injuries and the hospital treatment after a short delay (2-5 days) and a long delay (6 months). The initial interviews were more accurate than after the 6-month delay but the later inaccuracies were related to the hospital treatment and all children were 100% accurate when reporting information about the events surrounding the sustaining of the injury after both the short and long delay. Therefore, salient features of personally relevant events may be well remembered, even by very young children (2- to 3-years-old). However, errors did decrease with age and 2-year-olds made more errors than older children.

Two years later, all the children were re-interviewed and some received an extra interview 1 year after the target event (Peterson, 1999). All children recalled just as much information two years later as they did initially, but were less accurate. Accuracy rates for children that had participated in just three interviews fell from 98% to 80%, while accuracy rates for the children who received the extra interview fell slightly less from 95% to 83%. Peterson and Whalen (2001) re-interviewed the children again, 5 years after the original incident. While recall and accuracy decreased over the lengthy delay, this was not uniform. Instead only significant for the hospital treatment; the injury was remembered as completely and accurately as it had been after a few days. Older children

126

(5- to 6-years-old at the time of injury) evinced more complete and accurate recall than the youngest children, however, the youngest children (2-years-old at time of the event) did not differ in accuracy from children that had been 3- and 4-years-old at the time of recall. It is important to note that a major limitation of the Peterson studies is that accuracy was measured using a parent or other adult witness' account of the accident. Children were scored as making an error if their account contradicted details in the adult's account, and it was taken for granted that the adult's account was correct when it is likely that at least on some occasions the child may be correct and the adult bystander mistaken.

Three-to 4-year-old children that experienced Hurricane Andrew were interviewed after a few months and again after a 6-year delay (Fivush, McDermott Sales, Goldberg, Bahrick, & Parker, 2004). Children recalled a high number (M = 57.19 prepositions) about their experience after a few months, with the amount children recalled dramatically increasing after 6 years (M = 116.64). No measure of accuracy could be applied, therefore consistency was measured, propositions made in interviews were compared and categorised as either appearing only in the first interview, only in the second interview, or in both interviews. The number of propositions recalled in both interviews was low, ranging from 4.35 to 13.37 propositions on average depending on stress group membership. The findings of this study are extremely surprising from a memory perspective, while new details are often recalled at subsequent interviews the children in this study recalled twice as much about the event as they had originally. For example, the earlier mentioned Ornstein

127

et al. (1992) reported an average of 1.6 'new' items after 1- and 3-week delays. It is likely that the hurricane was a popular topic of conversation for quite some time afterwards and discussing the event resulted in reconstruction and incorporation of information from other sources (conversation partners, the news) into children's memories. Improved language skills at the second recall opportunity may also have facilitated a much more elaborate account as the children were older (Bjorklund & Douglas, 1997).

3.4.1 *Summary*. The results of these studies suggest that personally relevant, stressful events are well remembered by even very young children. In general, long delays (more than a few weeks) have been shown to affect the amount and accuracy of children's recall, with less information reported and decreased accuracy following delays (Ornstein et al., 1992; 2006; Peterson & Bell, 1996; Peterson & Whalen, 2001). However, some studies have demonstrated following lengthy delays children's reports can be just as complete (Peterson, 1999) and as accurate (Peterson, 1996; Peterson & Whalen, 2001) as they were initially, with some studies even reporting increases in amount recalled following a delay (Fivush, et al., 2004). Recent research has attempted to investigate why some experiences that are stressful and personally salient are well remembered (e.g. the injuries) yet others (doctor appointments, the hospital treatment) are not (for a review, see Cordon, Pipe, Sayfan, Melinder, & Goodman, 2004). In fact, the cause may relate not only to the content of the memory but to the coherence of the memory (how well it was understood and how well organised the narrative

surrounding the event was during encoding) affect how well it will be remembered.

Morris, Baker-Ward and Bauer (2010) found that children's memories that were reported with higher levels of narrative coherence (a clear focus) were more likely to be retained, while the percentage of content reported about the events did not inform estimates of the memories' survivability. Peterson, Morris, Baker-Ward and Flynn (2014) demonstrated using logistic multi-level modelling that children's memories containing emotional characteristics and thematic coherence (how well it was understood at the time) were more likely to be retained two years later than memories that did not contain emotional content or narrative coherence. Indeed, it is likely that the injuries were highly emotional memories which may be why they were so well remembered. Further, the events that suffered less degradation (the hurricane and injuries) are likely to have been topics of discussion with friends and family, leading to those narratives being reconstructed into a more coherent fashion, compared to the routine doctor check-ups which were not as well remembered.

Finally, it is important to note that in order to assess the effect of delay, children in these studies were interviewed as soon as possible (immediately or a few days after the to-be-recalled event) in order to obtain an initial measure of memory. Memory decay, or forgetting, occurs rapidly after experiencing an event (Hardt, Nader & Nadel, 2013; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991), after this initial period of vulnerability, memory traces remain relatively stable over time (Bauer &

129

Larkina, 2014) Interviewing soon after an event has been shown to preserve memory of an event (Dent & Stephenson, 1979; Brainerd & Ornstein, 1991; Poole & White, 1995; Warren & Lane, 1995). Therefore, we do not know how well children recall events after delays when they have not had this initial interview.

3.4.2   Forensic interviewing implications. Delay between the event/s and interview can be problematic as forgetting occurs rapidly following an event (Hardt, Nader & Nadel, 2013; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991). This can reduce the accuracy, as well as the amount, of information elicited from children. On the other hand, the salient nature of child abuse may mean it is well-remembered. For example, Lamb Sternberg and Esplin (2000) found delay was not associated with the length or level of detail of 4- to 12-year-old children's responses in a forensic interview, even though the delay since the experience and interview ranged from days to 14 months. Interviewers should focus their efforts on how they question children, even after 6-month-delays children made no errors in the free-recall (Peterson, 1996) while yes/no questions produced a disproportionately large amount of errors (Peterson & Bell, 1996).

3.5   Overall chapter conclusion

The amount and accuracy of information recalled about an event depends on many factors. Memory recall improves as linguistic ability develops, language allows verbal encoding (Simcock & Hayne, 2002), children better understand their experiences (Pillemer et al., 1994) and

130

language allows children to verbally express their memories. Further, language also provides additional retrieval cues that can be used to aid recall (Hayne & Herbert, 2004). Source monitoring ability is an important development that prevents suggested or imagined events becoming incorrectly attributed as being true memories that have happened to the self (Ceci et al., 1994a; Ceci et al, 1994b). This development also enables children to discriminate between episodes of repeated experiences (Brubacher et al., 2011a). Post-event information and inappropriate interviewing techniques (reliance on leading and suggestive questions and conveying interviewer biases), especially when experienced repeatedly, can alter children's memory traces, causing inaccurate information to become incorporated into an existing memory trace (Bruck et al., 1995b; Leichtman & Ceci; 1995; Thompson et al., 1997). Delay impacts memory traces, forgetting occurs soon after an experience (Hardt, Nader & Nadel, 2013; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991), therefore delay can negatively impact recall of an experience. However, highly salient events, such as abuse, may be well remembered (Peterson, 1996; Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001).

In conclusion, there are a number of complex interactions that can affect the quantity and quality of children's recall during a forensic interview. Inappropriate interviewing strategies may mean interviewers struggle to elicit memories from children and this may in turn lead them to think children are unreliable sources of information. However, a general consensus on how best to question children in order to elicit accurate

131

testimony has been developed following decades of research and this

has led to the development of interview protocols that acknowledge the

above limitations. This will be the focus of chapter 4.

Chapter 4: Conducting developmentally appropriate interviews with children

Due to the secretive nature of child abuse and the fact that victims often delay disclosing (Finkelhor, 1984; Goodman-Brown, Edelstein, Goodman, Jones & Gordon, 2003; Kellogg, 2017; London, Bruck, Wright & Ceci, 2008; Malloy, Brubacher & Lamb, 2011; McElvaney, 2015; Pipe, Lamb, Orbach, Sternberg, Stewart & Esplin, 2007) often no medical evidence of abuse can be found (Bays & Chadwick, 1993), hence, the account given by children may be the only evidence available to aid the prosecution build a case against the perpetrator (Goodman, Batterman-Faunce, Schaaf & Kenney, 2002; Lamb, 1994; Samra & Yuille, 1996). This chapter examines research surrounding interview protocols designed to guide interviewers through the process of conducting developmentally appropriate forensic interviews with children. As discussed in Chapter 2, children are capable of remembering and accurately recalling personally experienced events and Chapter 3 presented evidence that showed the way children are questioned is as important as their underlying cognitive skills. Probing memories using complex, leading and suggestive questions elicited greater numbers of errors (Hughes & Grieve, 1980; Leichman & Ceci; 1995; Pratt, 1990; Saywitz & Camparo, 1998; Thompson et al., 1997), while using open prompts to elicit free recall produced few errors (Goodman & Aman 1990; Roberts & Blades, 1998; 1999; Saywitz et al., 1991).

Based on such psychological research, developmentally appropriate interview protocols have been designed for use during interviews with children (see for example: ABE (Achieving Best Evidence) (Ministry of Justice, 2011), DNE (Developmental Narrative Elaboration) (Saywitz & Camparo, 2014), MOGP (Memorandum of Good Practice) (Home Office, 1992), NICHD (National Institute of Child Health and Human Development) Protocol (Lamb et al., 2007a, Lamb et al., 2008), Scottish Executive Guidelines (Scottish Executive, 2003, 2011) and the Step-Wise interview (Yuille, Hunter, Joffe, & Zaparniuk, 1993)/ Step-Wise Guidelines (Yuille, Cooper & Hervé, 2009)). Because these guidelines have been influenced by psychological research it is no surprise that they converge on a similar overall structure; a rapport-building phase to engage children and explain the interview procedure, a practice interview to familiarise them with responding to open prompts, the eliciting of a narrative account of what happened in the children's own words, focussed questions delayed until the end, and a closure phase where neutral topics are discussed before the interviews is closed. This chapter will review the psychological research that has influenced this structure of these interview protocols as well as the outcomes when interviewers use them.

For the literature search, the author's familiarity with the literature, and literature cited by Irit Hershkowitz, Michael Lamb, and Yael Orbach were used as a starting point. I then performed searches using Web of Science to find more recent and additional studies, the keywords "investigative interview" and "children" were combined to find highly-cited papers in the field.

134

4.1. Interviewing phases

4.1.1. Introductions.
The rapport phase serves a variety of important functions. It is there to familiarise children with the unfamiliar interviewer in an effort to put them at ease to facilitate the discussion of personal information with a stranger (Brown et al., 2013; Hershkowitz, 2009, 2011; Sternberg, Lamb, Davies, & Westcott, 2001a; Roberts, Lamb, & Sternberg, 2004; Siegman & Reynolds, 1984; Wood, McClure, & Birch, 1996) and to explain the purpose, roles and expectations of the unusual interview situation in an effort to transfer control to children and promote accurate responding (Brubacher, Poole & Dickinson, 2015; Lamb & Brown, 2006; Orbach et al., 2000; Saywitz, Lyon & Goodman, 2011; Saywitz, Snyder, & Nathanson, 1999). Then a practice interview, where children are asked to describe personally experienced events using open prompts that mimics the format of the substantive phase, allows children to practice responding to open prompts (Roberts, Brubacher, Powell & Price, 2011) and provides interviewers with an opportunity to gauge the children's language and memory skills (Poole & Lamb, 1998) and willingness to talk. This is an opportunity for interviewers to adapt their questioning to meet the children's needs if necessary and they may wish to spend longer in this phase to reduce anxiety and build trust with more reluctant children before moving to talking about substantive issues (Hershkowitz, 2011; Hershkowitz, Lamb & Katz, 2014).

To begin with interviewers introduce themselves and explain the purpose of the interview and the roles and responsibilities of all persons

135

present (e.g. the interviewer is 'here to listen' and the interviewee's role is to 'do the talking'). Field research has shown that interviewers' efforts to explain ground rules and the purpose of the interview improved the quality of children's reports (Teoh & Lamb, 2010). The greater the proportion of instructive prompts that informed children about the purpose and procedure of the interview in the pre-substantive phase, the greater the proportion of informative responses provided by 5- 7-year-olds in the substantive phase. Therefore, instructions and preparations during an interview are of paramount importance in preparing children to be the best informants they can be and to overcome contextual, motivational and social factors that may prevent them from reporting everything they know.

Next, interviewers should establish rapport and explain there are a different set of conversational rules during the interview, however, despite such widespread recommendations that rapport should always be built and ground rules always be communicated in forensic interviews with children, surprisingly little research has pointed to how this task is best accomplished and findings on the effectiveness of the ground rules is inconclusive and often yields contradictory results (Brubacher et al., 2015; Dickson, Brubacher & Poole, 2015; Ellis, Powell, Thomson & Jones, 2003; Geddie, Beer, Bartosik & Wuensch, 2001; Lamb et al., 2008; Roberts, Lamb & Sternberg, 2004; Saywitz et al., 2011; Van Roestel, 2016). For example, there is no consensus on where the ground rules should be placed, while most guidelines advise they are communicated immediately after the introductions (e.g. the Scottish

Executive 2003 guidelines; the NICHD Protocol, Lamb et al., 2007a), others support an approach where the ground rules are dispersed throughout the interview and selected for use as and when needed as judged by the interviewer to suit the particular situation (The Revised NICHD Protocol, Hershkowitz et al., 2014; APSAC, Lamb, Hershkowitz & Lyon, 2013; The Developmental Narrative Elaboration Interview, Saywitz & Camparo, 2014).

One explanation for some ground rules being ineffective is that they have been incorporated into guidelines not because they are supported by research, but because of other factors. For example, the common competency measure of assessing children's understanding of truth and lies. While a legal requirement, the assumption that just because children can identify lies and claim hypothetically that they would tell the truth, the competence test has no bearing on children's actual behaviour. Even when children could discriminate between truth and lies and claimed they'd tell the truth if they were in a hypothetical child's place, they lied to interviewers. When children were explicitly asked to promise to tell the truth, fewer of them lied to interviewers than children given no such truth instruction (Talwar, Lee, Bala & Lindsay, 2002). Therefore, a 'truth oath' promotes honest responding (Lyon & Dorado, 2008) while the test of discrimination, a principle that was included in many guidelines, does not.

Field research has guided the creation of some rules, for example, Roberts and Lamb (1999) found that children failed to correct interviewers when they distorted details (e.g. through paraphrasing) two thirds of the

time. Children are not used to correcting adults and need to be told it is ok for them to do so, therefore, the instruction to 'Correct me if I make a mistake' is conveyed to transfer this power to children. Warnings that some questions may be 'repeated questions' also has a basis in research, studies have shown that children often change their answers when questions are repeated (Fivush & Schwarzmueller, 1995; Fivush & Shukat, 1995; Lyon, 2002; Poole & White, 1993). A warning that questions may be asked more than once, but this is not because the child's first answer was incorrect, are to provide a rationale for repeated questions and to discourage children from changing their answers because they have interpreted the re-presentation as a sign their initial answer was unsatisfactory.

Other common ground rules include interviewers' communicating they are naive to what has happened and the child is the one with the 'knowledge' because children are used to being questioned by adults that already know the answers to their questions so it must be made explicit that in this situation it is the child that is knowledgeable. Further, children are encouraged in real life situations to guess when they don't know the answer to a question and therefore the 'don't guess' instruction and providing alternative ways to indicate lack of knowledge such as saying 'I don't know', 'I don't remember', or 'I don't understand' where appropriate are also commonly communicated.

Research has shown some positive results of using ground rules, for example, Mulder and Vrij (1996) found telling children the interviewer

138

cannot help them answer questions (interviewer is 'naïve') and that saying "I don't know" is an acceptable answer elicited fewer incorrect answers from children that were not given the rules or only given one of the rules. Saywitz et al., (1999), found that telling children to indicate their lack of comprehension when asking them difficult to comprehend questions facilitated accurate recall. Krackow and Lynn (2010) combined ERT (Event Report Training) and instructions that interviewers are naive and to correct the interviewers and found that ERT decreased suggestibility to abuse-related questions in pre-schoolers during a memory interview. However, other studies have shown no benefit of instructing children to say 'don't know' (Ellis, Powell, Thomson & Jones, 2003; Peterson & Grant, 2001), to correct the interviewer and not to guess (Ellis et al., 2003) or delivering warnings that repeated question may be asked (Geddie, Beer, Bartosik & Wuensch., 2001; Memon & Vartoukian, 1996).

An alternative explanation for children's lack of utilisation of interview instructions could reflect the fact that these studies did not provide children the opportunity to practice using the ground rules. It is recommended that interviewers do not just ask whether the child understands the rules but check by giving examples as it is expected this will make the ground rules more effective than just stating the rules alone (Brubacher et al., 2015; Danby, Brubacher, Sharman & Powell, 2015; Hamilton, Brubacher & Powell, 2016; Lyon, 2010; Saywitz, Camparo, & Romanoff, 2010). For example, after instructing children to say 'I don't know' interviewers can say "If I say, 'What day is my birthday?' you

139

should say….". The only study to date that has compared children's performance when told to use ground rules against children that were told and had the opportunity to practice applying the rules was conducted by Danby et al., (2015) who found that practice aided children in using the 'don't know' rule but not the 'don't understand' or 'correct me' rules. This rule has been previously shown to be more easily used by children compared to other more complex rules (Dickinson, Brubacher & Poole, 2015). Therefore, as concluded by Brubacher et al., (2015) much more research is needed in order to understand what skills are required in children to use these instructions to best effect, such as theory of mind and metacognition. As these cognitive skills may not yet be developed in some age groups, not all children will benefit from the standardised delivery and different age groups may require different instructions in order to reap any benefit from such discussions.

Despite recommendations to convey ground rules, field interviews have shown that interviewers fail to communicate basic rapport building components such as introducing themselves and giving the child permission to say "I don't know/understand" and to correct the interviewer (Sternberg, Lamb, Esplin & Baradaran, 1999; Warren, Woodall, Hunt & Perry, 1996). Research on use of ground rules in field interviews has also produced contradictory results (Earheart, La Rooy, Brubacher & Lamb, 2014; Hamilton, Brubacher & Powell, 2016)

Earheart, La Rooy, Brubacher and Lamb (2014) looked at use of the 'don't know' rule. In their sample less than half (49%) of the

interviewers communicated the 'don't know' rule and none gave children a practice in using the rule. Interviews that contained the 'don't know' ground rule were then matched with interviews where the rule had not been communicated. There were no differences in the likelihood of children saying don't know whether they had or had not received the 'don't know' instruction. However, interviewers rejected around 30% of children's indications that they didn't know by following up such responses with continued questioning about the same topic or increasing pressure to answer the question. This resulted in children who indicated a lack of knowledge about the topic to then provide an answer 81% of the time. Therefore, even when interviewers communicated the rule their responses to children's utilisation of 'don't know' responding may have sent a contradictory message that they did not accept don't know as an answer by continuing to pursue the answer to the question which may have ceased children's don't know responding. Hamilton, Brubacher and Powell (2016) on the other hand found that delivery and practice of ground rules at the beginning of the interview was positively associated with the spontaneous usage of rules in children's narratives of abuse. When specifically examining the "don't know" rule, however, only practice had an effect of children's usage of the rule.

Next, interviewers build rapport by asking children about things they like to do using open prompts. There is no consensus on the length of time that rapport should last but research has indicated that shorter rapport building sessions (less than 8 minutes) are more effective in aiding children's responding later in the interview (Davies, Westcott &

141

Horan, 2000). Hershkowitz (2009) also found in interviews with short rapport sessions using open prompts children responded with a greater number of details on average to the open prompts asked later in the substantive phase. The more rapport building prompts (e.g. talking about family and school) and words interviewers used by interviewers with young children (5-7-year-olds) in the rapport phase of the interview, the fewer informative responses elicited from children in the substantive phase. However, this was not the case for instructive prompts (ground rules), which were positively correlated with informative responses. Too much rapport building is counterproductive, possibly because children's attention and cognitive resources are limited, the lengthy rapport may be rendering them fatigued (Hershkowitz, 2011; Lamb & Teoh, 2010), while shorter rapport sessions are less demanding on children's limited cognitive resources.

4.1.2. Narrative Elaboration Training. Once the ground rules and rapport have been established, a practice interview, otherwise known as narrative elaboration training (Saywitz & Geiselman, 1998; Saywitz & Goodman, 1996; Saywitz & Snyder, 1993, 1996; Saywitz et al., 1999), is recommended. The topic for the practice interview should be a neutral or positive experience that will hold the child's interest such as a national holiday (Roberts et al., 2011). The interviewers should ask children to describe event(s) from the beginning to the end using open questions that request episodic information about individual event(s), thereby providing opportunities to 'train' children to respond with lengthy narratives about specific events before the substantive phase of the interview begins

142

(Lamb et al., 2008; Orbach et al., 2000; Saywitz & Snyder, 1993;1996) and giving interviewers practice questioning the child using open prompts (Roberts et al., 2011). Even during the substantive phase children provide longer responses to open prompts than they do to directive and option-posing questions, and prompts inquiring about events elicited longer responses than questions about hobbies/likes (Price, Ahern & Lamb, 2016).

As discussed in chapter 1, some maltreated children have been shown to refrain from using conversational interactions as a social tool and may not be used to lengthy social interactions. Thus, engaging children in narrative elaboration training will be useful in providing an example of how to structure and report their experiences and the level of detail that will be required from them. In cases where children have been repeatedly abused, practice in responding to episodic questions about a single instance of an event produces responses of a more episodic (what happened during a specific instance) rather than generic (what usually happens) nature during later questioning (Brubacher, Roberts & Powell, 2011b).

Field research has shown that when interviewers conduct an open-ended practice interview the amount of information obtained from children in the substantive phase is greater than the amount obtained from children who did not receive this opportunity (Price, Roberts & Collins, 2013; Sternberg et al., 1997). Children provided almost five times more details in response to substantive open-ended prompts when interviewers

conducted a practice interview than when no practice narrative was conducted. This relationship was further enhanced when the practice narratives were conducted as recommended in an open-ended manner (Price et al. (2013). Sternberg et al. (1997) found that when interviewers build rapport using open prompts, children provide 2 ½ times as many details and words in response to the first substantive utterance and continue to respond more informatively to open-ended utterances during the remainder of the interview, relative to children whose rapport-building session involved direct questions, even though the rapport phases are of the same duration and cover the same topics. Therefore, any concerns interviewers may have about fatiguing children do not seem founded as children provide longer and more detailed responses after practice (Brubacher, Roberts & Powell, 2012; Price et al., 2013; Sternberg et al., 1997). Further, just a few minutes of rapport is sufficient, which is consistent with the research discussed earlier and so should not extend the overall length of the interview to an extent that will fatigue children (Brubacher, Powell & Roberts, 2014).

4.1.3. Free-recall narrative account. Next, interviewers should obtain accounts of each of the alleged incident(s), in the children's own words using open prompts e.g. "Tell me what happened". Field studies have shown that when interviewers use open questions in forensic interviews with children they elicit longer and more detailed responses from children than directive, option-posing and suggestive questions (Aldridge et al., 2004; Brown et al., 2009; Feltis, Powell, Snow & Hughes Scholes, 2010; Lamb, 1994; Lamb et al., 1995; 1996a; 2003; 2007a;

144

2007b; 2009; Orbach & Lamb, 2000, 2001; Orbach et al., 2000; Powell & Snow, 2007; Sternberg et al., 1996; 2001a; 2001b). Open prompts are questions, statements or imperatives which are free from interviewer input or constraints, that is they do not introduce information that has not already been mentioned by the child and they do not dictate or limit what information the child is to disclose. Open questions probe recall memory, which requires respondents to conduct a memory search in an attempt to provide as much relevant information as they "remember," whereas focused questions involve recognition processes which do not require this search (Lamb et al., 1995; Sternberg et al., 1997).

Importantly, studies that have had the means to assess the accuracy of details provided in field interviews with children have shown that details elicited by open questions are highly accurate (Lamb & Fauchier, 2001; Lamb et al., 2007b; Orbach & Lamb, 1999; 2001; Orbach, Lamb, La Rooy & Pipe, 2012). For example, Lamb et al. (2007b) compared accounts provided by victims of child sexual abuse with the confessional accounts given by the perpetrators of the abuse in order to assess the convergence between the details provided. Each detail reported by the alleged victim was classified as confirmed, contradicted, ambiguous or ignored by the suspect. While most of the details were ignored (66%) about 30% of the details elicited using open prompts were confirmed by the suspects. Details elicited by open prompts were more likely to be confirmed than details elicited by focussed prompts, making this type of prompt superior to focused prompts with respect to the proportion of reported details that were confirmed.

145

An important caveat for forensic interviewers regarding open questions is that young children typically provide little information in free recall or in response to open-ended questions (see Pipe, Thierry, & Lamb, 2006, for a review) while directive questions are highly effective in producing more details from young victims (Andrews, Lamb, & Lyon, 2015; Korkman, Santtila, Drzewiecki & Sandnabba, 2008; Lamb Sternberg, Orbach, Esplin, Stewart & Mitchell, 2003; Yi, Lamb, & Jo, 2014). This is likely because directive questions contain a 'cue' taken from children's previous responses that helps children focus their answers. One way interviewers can maintain open prompting with young children is by using cued invitations. Gagnon and Cyr (2017) found children aged 3- to 6-years-old gave more detailed answers to open-ended question using cues (cued invitations or directive open-ended) compared to questions that did not contain cues (general invitations) when interviewed about sexual abuse. The youngest children (aged three and four) gave the same number of details following directive questions and invitations, while those aged five and six provided more details following invitations compared to all other types of questions, including directive questions. The ability to respond to open prompts increases with age (Brown & Lamb, 2015). This also further underscores the paramount importance of narrative elaboration training because young children need practice responding to these unusual prompts in order to provide information.

If the child has been abused on more than one occasion, interviewers must be aware that memories for repeated experiences often form 'scripts' (what usually happens), therefore, techniques that elicit individual accounts of separate incidents rather than 'gist' accounts must be used. Separation of incidents should be achieved by using episodic prompts (Schneider, Price, Roberts & Hedrick, 2011) and unique labels associated with each different incident (e.g. 'the first time,' 'the last time,' or 'the time you said x happened' (Powell & McMeeken, 1998). Further, these labels should be generated by the children themselves to prevent interviewers' cues biasing their memory searches (Brubacher, Malloy, Lamb & Roberts, 2013; Brubacher et al., 2014). See Brubacher, Powell and Roberts (2014) for a recent review on eliciting best-practice accounts of repeated experiences from children.

4.1.4. Questioning. Focussed questions or recognition prompts, such as option-posing 'yes/no' and 'forced-choice' questions, that require the child to affirm, negate or select an investigator-given option are problematic. They rely on recognition memory processes, which do not require interviewees to search memory, but instead allow them to acquiesce and to guess the answers to questions (Waterman, Blades & Spencer, 2000). By their very definition, while they introduce information, they also elicit erroneous information and allow interviewees to go along with misleading information (Bruck & Ceci, 1995, 1996; Ceci & Bruck, 1993, 1995; Brady et al.,1999; Dale, Loftus & Rathbun, 1978; Dent, 1986; Dent & Stephenson, 1979; Garven, Wood, Malpass & Shaw, 2000; Goodman & Aman, 1990; Goodman, Bottoms, Schwartz-Kenney & Rudy,

147

1991; Hutcheson, Baxter, Telfer & Warden, 1995; Lamb & Fauchier, 2001; Oates & Shrimpton, 1991; Orbach & Lamb, 2001; Ornstein et al., 1992; Peterson & Bell, 1996; Peterson, Dowden & Tobin, 1999; Peterson & Grant; 2001; Rocha, Marche & Briere, 2013; Waterman et al., 2004). Therefore, it is recommended that focussed questions are delayed until later in the interview after free recall is exhausted and be used to inquire about aspects of the incident(s) that children have failed to mention.

Field studies that have had the means to assess accuracy have shown that focussed questions are more likely to produce contradictions than open questions (Lamb & Fauchier, 2001; Orbach & Lamb, 2001). Lamb and Fauchier (2001) compared the details elicited in interviews with children with details the same children provided in later interviews, in order to assess what types of questions elicited information most likely to be contradicted later. Every detail that contradicted a previously elicited one was in response to a focussed question, with 86% of the contradictory pairs of details containing at least one question that was option-posing or suggestive and suggestive questions eliciting significantly more contradictory details than expected by chance. Similarly, Orbach and Lamb (2001) looked at within-interview contradictions made by a 5-year-old who had been asked an excessive number of repeated, leading and suggestive questions and found that a high proportion (41%) of the option-posing and suggestive questions were involved in self-contradictions and that 94% of details that contradicted earlier details were elicited using option-posing or

suggestive utterances. Both studies found that details elicited using open prompts never contradicted earlier reported details.

Once an account of all alleged incidences has been elicited, interviewers are recommended to elicit information about any disclosures that have been made prior to the interview. Eliciting information regarding prior disclosures is important for many reasons, such as but not limited to, identifying additional witnesses that may be able to verify the account or in cases of delayed disclosure to enhance the child's credibility if the case goes to court, as a jury may be interested as to why a child may have delayed disclosing (Schaeffer, Leventhal & Asnes, 2011).

4.1.5. Closure. The closure phase of the interview should always be conducted; even if no disclosure is made or the children do not appear distressed. Neutral topics should be discussed. The children should be thanked for their participation and made aware of what will happen next, provided with contact details and offered the opportunity to ask any questions they may have. However, field research has shown that interviewers do not uniformly conduct closure. For example, Westcott and Kynan (2006) found that in a sample of interviews conducted in the UK between 1994-1997, only around half of the interviewers offered children the opportunity to ask questions, to report anything else they wished to say and to thank them for participating and only 14% conducted a neutral

149

closure despite recommendations that a neutral closure should always be conducted.

4.1.6. *Supplementary techniques.* Despite professional recommendations that anatomically-correct dolls should not be used as a diagnostic test of abuse (see, e.g., guidelines of the American Professional Society on the Abuse of Children, 1990), they are still used in many interviews for a variety of reasons. A full review of the literature on anatomically correct dolls is beyond the scope of this thesis (but see reviews by Ceci & Bruck, 1995 and Poole & Bruck, 2012). However, it is important to note the research has looked at their use in field interviews. Lamb, Hershkowitz, Sternberg, Boat, and Everson (1996b) were able to compare field reports of abuse elicited from two groups of children (aged 4-12 years), one group having been interviewed with anatomical dolls and the other without. Children interviewed without the dolls gave longer and more detailed responses. This could suggest that the use of anatomical dolls, instead of facilitating children's accounts as they are designed to do, in fact may inhibit them. The authors noted that they failed to find any clear benefits from using the dolls, and so suggested that caution should be exercised if any of these aids were used in forensic interviews.

Human figure diagrams (HFDs) are also used widely in forensic interviews with children. Laboratory studies have shown that body diagrams do indeed elicit accurate reports of touching but they also elicit inaccurate (potentially forensically relevant) allegations of touching (Brown, Pipe, Lewis, Lamb & Orbach, 2007; Bruck & Poole, 2016). The

150

diagrams have indeed been have shown to elicit many new details in forensic interviews, particularly with the youngest children (4- to 7-year-olds), however, the accuracy of these details cannot not be verified (Aldridge et al., 2004; Teoh, Yang, Lamb & Larsson, 2010). The paucity of research available on HFDs means that no firm conclusions regarding the utility of the diagrams in forensic interviews with children can presently be made and more research is needed to clarify whether they can in fact be used to aid children's recall of bodily touch.

4.2.   *Forensic field research*

With the general consensus on how to best interview children informing national child interviewing guidelines in many countries it was expected that field interviews with children would be conducted to a high standard in line with such guidance. It was not until the mid-1990s that researchers first began to conduct analyses on the quality of forensic interviews conducted by child protection workers with suspected victims of child abuse in the field (Lamb et al., 2008). One of the most formidable challenges in conducting research of this nature is obtaining interview transcripts. Due to the highly sensitive nature of the information contained in the interviews permission must be granted from ethical review boards and from the organisations that collected and store the data (e.g. family protection unit of the police jurisdiction that the interviews were conducted in). This has proved very difficult to achieve in the past (as evidenced by the small number of transcripts obtained in the first wave of field studies that are detailed later in this section). Further, the interview transcripts are usually either audio or visually recorded and are therefore stored on

tape or DVD. For the purposes of coding the interviews must be transcribed verbatim as the coders need to refer to earlier portions of the interview in order to determine where and how information was introduced which is too difficult to do in 'real-time' viewing recordings (Earheart, La Rooy & Lamb, 2016). This adds further difficulty to the research as tapes are required to stay on site where they are stored and therefore organisations must accommodate researchers a space on site where transcribing can be conducted.

When transcripts are obtained, in order to assess the quality of a forensic interview both the types of questions asked by the interviewer and the informativeness of children's responses must be closely examined. The entire substantive phase of the interview is coded, with each interviewer utterance being categorised into a question category based on both the structure and the content of the question. The following example of question categories was developed by Lamb et al. (1996a, discussed below) and has formed the basis of the general coding system used in forensic field studies to the present day.

1. Invitations (Open prompts). These are input-free utterances, including questions, statements or imperatives prompting free-recall responses from the child (e.g., 'Tell me everything that happened'), or utterances which use details disclosed by the child that refocus their attention using contextual cues to elicit free recall (these are known as cued invitations) e.g., 'You mentioned that he touched you. Tell me everything about the touching').

2. Directives. These refocus the child's attention on details or aspects of the alleged incident that the child has already mentioned, providing a category for requesting additional information using 'Wh-' questions (cued recall). For example, the interviewer might say 'what colour was the t-shirt'? when the child mentioned a t-shirt).

3. Option-Posing Utterances. These focus the child's attention more narrowly on details or aspects of the alleged incident that the child has not previously mentioned, asking the child to affirm, negate, or select an investigator-given option, thus using recognition memory processes, but do not imply that a particular response is expected. For example, the investigator might ask 'Did he touch you over or under your clothes'? when the child mentioned being touched.

4. Suggestive Utterances. These are stated in such a way that the interviewer strongly communicates what response is expected (e.g., "He forced you to do that, didn't he"?) OR they assume details that have not been revealed by the child (e.g. Child: 'We laid on the sofa'. Interviewer: 'He laid on you or you laid on him'?). An utterance may also be coded as suggestive when the interviewer presents the same option-posing question three times or more.

5. Facilitators. Utterances such as "uh-huh" or restatements of the child's last few words that are used as encouragement to keep the child talking.

6. Summaries. These are statements made by the interviewer which accurately restate what the child has just said without any explicit request for information or response.

153

7. Non-substantive Utterances. These are comments not related to the alleged incident such as "Do you not know the answer or is it hard to say?".

8. Introductory comments. These are utterances that related to the general interview procedure such as "I'm just going to pause there so Sophie can catch up with writing everything down".

It is also important to note that while some questions structurally may look like they belong to one category, the information the question is requesting must also be taken into account. For example, an interviewer may ask "Do you have a name for the place he touched you?" While it looks like an option-posing question due to the fact is can be answered with yes or no, the interviewer is actually requesting a name for the body part and would therefore be coded as a Directive. Previously elicited information must also be taken into account, for example, while "tell me about him punching you" looks like an open prompt, it would only be classed as one if the child had mentioned being punched. If the child had not mentioned being punched it would be a suggestive question. Often utterances contain two or more statements, in this case a trumping system is used and the question is allocated a main code based on the question that contained the most input (has the highest question number from the above list). As it is best practice to delay the use of focussed prompts until as late as possible in the interview, often the number and percentage of questions asked before both directive and option-posing

utterances is calculated, as well as the number and percentage of details elicited before the first use of directive and option-posing prompts.

Every response by a child is scored as either informative or uninformative. Informative responses include:

1. Responsive. Related to the content of the interviewer's previous utterance. Utterances can be assigned to this category even if they do not contain informative details, or when their meaning is unclear. *I: Where did it happen? C: It happened at my house.*

2. Responsive action. Child action response (gestures) not accompanied by a verbal response that relates to the interviewer's previous utterance e.g. pointing or nodding.

3. Unresponsive. Responses that do not relate to the content of the interviewer's previous utterance but do provide incident-related information. A child's response to an IC that contains informative details is also included. *I: Where did it happen? C: It happened when I was five years old.*

Uninformative responses include:

4.No answer.

5. Clarification requests.

6. Unclear responses.

7. Digressions. An off topic response where the answer is unrelated to the previous interviewer prompt or to incident under investigation.

155

Children's responses to each substantive interviewer utterance are also coded for both length and for level of detail, that is, the number of words and the number of incident related details their answers contain. A detail is the smallest unit for analysing information provided in the interviewee's account. A detail is any information pertaining to the incident that is conveyed by the interviewee. A detail consists of naming, identification or description of individual(s), object(s), event(s), place(s), action(s), emotion(s), thought(s), and sensation(s) that are part of the alleged incident, as well as any of their features (e.g.-appearance, location, time, duration, temporal order, sound, smell, and texture). Only *new* details per utterance are counted (repeated details are not counted again) and non-verbal cues (points, nods) are counted as details but not as words.

Conducting such a detailed psycholinguistic analysis is time consuming and labour intensive, however, the coding must also be subject to a further check of accuracy and consistency through comparison with a second coder. Two coders independently code a transcript and then a Cohen's kappa coefficient (Cohen, 1960) is calculated by taking the proportion of units of agreement minus the proportion of units which would be expected to be agreed upon by chance and dividing this number by 1 minus the proportion of units expected to be agreed upon by chance (the maximum which this difference could be). Raters train until reliability is achieved. A kappa value of 0.70 is generally considered to be satisfactory, however, the desired reliability level varies depending on the purpose for which kappa

is being calculated. Then both raters will independently code a random subset of the interviews in the sample, with one rater coding all transcripts, to ensure consistency in the method of coding. Cohen's kappa is the preferred method of calculating reliability in this field as it is a more conservative measure of agreement than calculating the percentage agreement alone as it takes into account the amount by which the observed agreement exceeds that expected by chance alone.

This first wave of field studies conducted such quality analyses on field interviews conducted in various different countries under each country's national guidelines, with all guidelines created in the spirit of the international consensus on how to appropriately question children. Lamb et al., (1996b) analysed 24 interviews conducted in the USA, finding that only 2% of utterances were open prompts, even though the open prompts elicited longer and more detailed responses than the focussed prompts when they were employed. Another study conducted in the USA by Sternberg et al. (1996) found that 4.7% of substantive interviewer questions were open prompts, despite the open prompts yielding more words and details from children than the focused questions. Similarly, in Sweden, Cederborg, Orbach, Sternberg and Lamb (2000) found open prompts accounted for only 6% of the total substantive questions asked and elicited only 8% of the overall details from children in forensic field interviews, with only 1 question on average being asked before interviewers moved to using focussed prompts. In Finland, Korkman, Santtila and Sandnabba (2006) obtained a small sample of 27 transcripts from legal professionals (lawyers, expert witnesses) involved in court

cases of alleged child sexual abuse and analysed them in a similar manner. They found open prompts accounted for only 2% of the total questions asked in the investigative interviews, the majority of details were elicited using directive questions and 66% of interviews used an option-posing or suggestive question as their first utterance. Similarly, in Norway, Myklebust and Alison (2000) sampled 11 investigative interviews with children and reported that "almost no open questions" (p.339) were used with interviewers instead relying mainly on closed questions.

Even locations that provided highly specialised training and specific interviewer guidelines did not produce high quality interviews (Lamb et al., 1996a; Sternberg et al., 2001a). Lamb et al. (1996a) obtained 22 transcripts conducted in Israel by specially trained youth investigators. They are the only people in this country allowed to interview children and are trained in a system that operates uniform policies and strict training procedures. However, these interviewers used only 2.2% open prompts even though the few open prompts they asked elicited longer and more detailed responses than any other prompt. Sternberg et al., (2001a) analysed the quality of interviews conducted with children in England and Wales between 1994 and 1997, following the introduction of the Memorandum of Good Practice (MOGP, 1992) interview guidelines. The authors predicted that the these interviews would be of a higher standard than previous samples of field interviews as the MOGP contains explicit and specific guidelines on how to interview children based on best practice recommendations and was introduced on a national level with every child interviewer trained in their use. However, this prediction was

incorrect, only 6% of the substantive questions asked were open and interviewers asked only 6.3 utterances on average before introducing their first option-posing question, by that time children had provided only 8% of the total details they went on to provide. Davies et al., (2000) also analysed the questioning in 36 investigative interviews in the UK conducted under MOGP guidelines between 1991 and 1997 and found that only 2% of questions were open prompts.

Similarly, In Norway in 1998 regulations on interviewing child witnesses were introduced (Norwegian Ministry of Justice and the Police, Section G-70/98) followed by the Norwegian Police University College developed manuals and training programs in the forensic interviewing of children in 1999. Thoresen, Lonnum, Melinder and Magnussen (2009) analysed investigative interviews with children conducted throughout Norway between 1990-2002, they divided the interviews into 3 time periods in an effort to measure any improvements made following the introduction of guidelines and training. Open prompts accounted for 1.4% of interviewers' questions in 1990-1994, 2.2% in 1995-1998 and 1.8% in 1999-2002. A follow-up study by Johnson et al., (2015) assessed the quality of interviews conducted in Norway in 2002-2012. The interviews were divided into 3 time periods; 2002-2005, 2006-2009 and 2010-2012. Despite the majority of interviewers having attended the newly introduced training programs, open prompts constituted 2% of all interviewer questions/statements and there were no significant differences in the frequency of open prompts used across the 3 time periods.

Surprisingly, despite undergoing interviewer training and the provision of national guidelines as a tool to aid interviewers in conducting high quality interviews with children, field research consistently shows that interviewers around the world fail to execute this task. Instead, they typically rely on focused questions, use few open prompts and structure their interviews poorly. The disappointing results of these field studies demonstrate that knowledge of appropriate interviewing and best practice guidelines alone are therefore not enough to prepare interviewers to conduct high quality interviews with children, even with the existing training. This conclusion lead the researcher to develop the highly specialised NICHD structured interview protocol (Lamb et al., 2007a) in an effort to provide operational guidelines that interviewers can refer to throughout the interview to help them put the guidance into practice and structure their interviews appropriately but are still flexible enough to not follow a standard script. Unlike many interviewer guidelines/protocols this child interviewing protocol has undergone rigorous field testing to assess whether or not it is effective in assisting interviewers to conduct high quality interviews that adhere to best practice.

4.3. Introduction of a protocol

The NICHD protocol aims to maximise the amount of information obtained from children using open questions (Lamb et al., 2007a). It provides a flexible structured script that demonstrates multiple ways open questions can be used to enhance children's testimony. Open prompts characterise the rapport building and a narrative elaboration training section, in the substantive phase, where riskier focussed questions are

often unavoidable. It is recommended that these types of prompt are 'paired with' (i.e. followed by) open prompts, that return the child to free recall responding. For example, when interviewers are investigating genital touching but children have only reported general touching, they may ask "Were you touched anywhere else on your body?" and if the child responds affirmatively they can then 'pair' the previous question with an open prompt such as "Tell me everything about that touching" (Lamb, Sternberg, & Esplin, 1994). Open prompts are equally effective at any stage of the interview (Orbach & Lamb, 2000).

Following the development of the NICHD protocol a series of field validation studies were conducted. Investigative interviewers in Canada, Sweden, the UK and the USA received NICHD Protocol training and then their subsequent protocol interviews were compared to interviews they had conducted before protocol training using the national guidelines in their respective countries. Training was extensive and included viewing videos of best practice interviews, conducting role-play interviews, and coding interviews. Written feedback, and follow up sessions were also implemented. Orbach et al. (2000) found that the Israeli investigative interviewers used a greater percentage of open prompts (30% compared to 6%) and fewer focused prompts once they had been trained to use the NICHD protocol than they had in their pre-protocol training interviews. Further, following protocol training interviewers asked a greater number of open-ended substantive questions before asking their first option-posing question than they had done before protocol training. Cyr and Lamb (2009) also found that following protocol training, the quality of

161

interviews conducted by experienced Canadian police officers and social workers improved. The interviewers used 3 times more open prompts and 2 times less directive, option-posing and suggestive questions after protocol training than they had in their pre-protocol interviews. They also asked a greater number of open prompts before asking their first option-posing prompt in their protocol interviews than they had in their pre-Protocol interviews.

In the USA Sternberg et al. (2001b) showed post-protocol training police officers' interviews contained three times more open prompts as well as significantly fewer option-posing and suggestive questions than their pre-protocol training interviews that had not been guided by any protocol. Here, more open prompts were offered before the first option-posing or suggestive utterance was used, both in absolute and proportional terms. Lamb et al. (2009) compared a sample of interviews provided by one police force in England. Half of the interviews were conducted by interviewers trained to conduct Memorandum of Good Practice interviews and the other half were conducted following NICHD protocol training. The Protocol interviews contained a greater proportion of open prompts and proportionately fewer directives, option-posing and suggestive prompts than the MOGP interviews and option-posing and suggestive questions were introduced later than in the Protocol interviews than the MOGP interviews. Some interviewers contributed interviews to both the MOGP and the Protocol interviews and their performance could be directly compared. These interviewers asked both absolutely and proportionately more open prompts and fewer suggestive and option-

162

posing questions in their interviews after protocol training than they had in their MOGP interviews.

In addition to improvements in questioning behaviour the quality of information elicited improved, the implementation of the NICHD protocol meant fewer of the details elicited from children came from focussed prompts and a greater number from open prompts than children in pre-protocol interviews (Lamb et al., 2009; Orbach et al., 2000). Information elicited from open prompts is more likely to be accurate than information elicited by focussed questions, therefore, protocol interviews are more likely to produce more accurate accounts from children. Sternberg et al. (2001b) reported children in the protocol group provided nearly two and a half times as many details in response to open prompts and Cyr & Lamb (2009) found that open prompts elicited 62% of all details elicited in Protocol interviews compared to only 14% in pre-protocol interviews. Lamb et al. (2009) suggest that protocol interviews produce this level of information from open prompts not only because interviewers ask more open prompts in the substantive phase, but also because children have been trained in the pre-substantive phase on how to respond to open prompts, making them more effective in the substantive phase. More recent research has examined the effect of training interviewers to use the Protocol in Canada (Cyr, Dion, McDuff & Trotier-Sylvain, 2012) and Korea (Yi, Jo & Lamb, 2016). Both studies found interviewers used more open prompts and elicited more details from children than in their pre-Protocol interviews.

163

When high quality interviews are conducted decision makers'
judgements are aided. Hershkowitz, Fisher, Lamb and Horowitz (2007)
demonstrated that when youth investigators assessed the credibility of
allegations made in interviews conducted using the NICHD protocol they
were more likely to be judged as credible or not credible as opposed to
'no judgement possible' than interviews not following the protocol. This
was due to the protocol interviews containing more CBCA (Criterion
based content analysis) criteria, that is protocol interviews contained
more components that are characteristic of truly experienced events.
Further, when interviews are conducted to a high standard using the
NICHD Protocol, charges are more likely to be filed against the accused
and courts are more likely to deliver guilty verdicts (Pipe et al, 2009; Pipe,
Orbach, Lamb, Abbott & Stewart; 2013). Therefore, implementing a
structured interview protocol and training interviewers in its use is a vital
step that should be implemented on a national level to prepare
interviewers to conduct high-quality interviews with children in order to
achieve justice for children.

A recent review by La Rooy, et al. (2015) discusses the various
countries around the world that have or are in the process of
implementing the NICHD protocol. The protocol has been officially
implemented in Israel and all child interviewers are expected to adhere to
it. Countries such as Canada, the USA, Finland, Japan, Norway, Scotland
and Korea have either implemented the protocol in some regions or have
national guidelines based on NICHD Protocol research. Recently the
protocol was also translated into Portuguese and research in its use in

164

Portugal is underway. It is vitally important to note that implementing a protocol alone, as is the case with best practice knowledge and guidelines, is not the only step that must be taken. Training in the protocol field validation studies was extensive while studies that have implemented a protocol but not extensively trained interviewers in its use have not had the same successes. For example, in Sternberg et al., (1999) almost 40% of the interviews they collected had to be excluded from the sample due to interviewers not following the interview protocol they had been given, therefore training is of paramount importance.

4.4.   Revised protocol

Further, a revised protocol has been developed for reluctant disclosers. Some groups of children are less likely to make allegations than others; younger children compared to older children, children suspected of being abused by parent/parental figures compared to children abused by non-parental figures and sexually abused boys compared to sexually abused girls, are less likely to disclose such abuses (Hershkowitz, Horowitz & Lamb, 2005). Further to individual and case characteristics, interview dynamics play a role in disclosure (Hershkowitz et al., 2006) reluctant children (those that do not make disclosures) receive less support (e.g. fewer supportive comments from interviewers) than children that are willing to discuss their abuse.

In an effort to increase disclosure rates from groups that are known to be reluctant disclosers, the revised protocol (RP) was devised to attend to the socio-economic needs of children. This is achieved by engaging

children in rapport building before relaying the ground rules in the introductory phase of the interview and by encouraging interviewers to use non-suggestive support in response to reluctant displays in both pre-substantive and substantive phases of the interview. In order for the supportive comments to remain non-suggestive, they are to revolve around the children's emotions, provide empathy about interview experience, and reflect appreciation of effort in general but not specific instances. A field test of the revised protocol vs standard protocol on 4- to 13-year old children that had experienced intra-familial abuse (that had external substantiation) found that allegation rates were significantly higher when the RP was used rather than the standard protocol (Hershkowitz et al., 2014).

4.5. Overall chapter conclusion

Decades of psychological research have informed existing guidelines on how to best interview children. Interviewers should use open prompts to elicit an account of the event(s) in children's own words with focussed questions being delayed as late as possible and paired with open prompts when possible (Lamb et al., 2007a, 2008) in order to elicit longer and more accurate responses. Giving children the opportunity to practice responding to open questions before discussing substantive issues makes open prompts more effective in the later substantive phase of the interview (Price et al., 2013; Sternberg et al., 1997).

Despite national level training and the implementation of best practice guidelines in line with this international consensus on how to best

interview children, field research has shown that interviewers still rely on focussed questioning and use few open prompts when interviewing children (Aldridge & Cameron, 1999; Cederborg, et al., 2000; Davies et al., 2000; Johnson et al., 2015; Korkman et al., 2006; Lamb et al., 1996a, 2007a, 2008; Myklebust & Alison. 2000; Sternberg et al., 2001a; Thoresen et al., 2009; Warren et al., 1999). However, when interviewers are trained to use and follow a structured interview protocol, such as the NICHD Protocol, the quality of their interviews improves, with protocol trained interviewers using a greater number of open prompts, fewer focussed prompts and delaying the use of focussed prompts until later in the interview (Cyr et al., 2012; Cyr & Lamb, 2009; Lamb et al., 2009; Orbach et al., 2000; Sternberg et al., 2001b; Yi et al., 2016).

These high quality protocol interviews are easier for fact finders to assess credibility of allegations elicited and are more likely to proceed to court and produce guilty verdicts (Hershkowitz, 2007; Pipe et al., 2009; Pipe et al., 2013). Knowledge of best practice guidelines and existing training courses alone have not been effective in preparing interviewers to conduct best practice interviews, instead, extensive training, on-going supervision and feedback and following a structured interview protocol are absolutely crucial components necessary to prepare interviewers to conduct high quality interviews with children. Therefore, it is a matter of urgency that interviewers' training courses are updated to reflect these needs and that a structured interview protocol is implemented and followed as standard. Importantly, training must accompany the

implementation of a protocol as research shows interviewers do not adhere to protocols without it (Sternberg et al., 1997).

Further, it is vital to implement programmes of research that examine the quality of field interviews on an ongoing basis all around the world. The field and protocol validation studies were conducted over a decade ago, new up-to-date research is essential in order to determine the current quality of investigative interviews that are being conducted. Baseline data for some countries, e.g. Scotland, has never been collected and is therefore a matter of national importance that quality measures of field interviews are collected, particularly with the recent update of their national guidelines, as we do not know whether or not they facilitate the intended improvement. The following chapter will look at how interviewers are typically trained to conduct forensic interviews and whether or not the individual aspects of training courses actually prepare interviewers to conduct best practice field interviews.

Chapter 5: Training Investigative Interviewers to conduct developmentally appropriate interviews with children

Based on the general consensus on how to best interview children, various guidelines have been developed to standardise the procedure. However, despite available resources, investigative interviewers' training courses, and the exercises they employ to train their interviewers, are not standardised (Powell, 2002; Powell & Wright, 2008). Currently training programmes are subject to time and financial constraints, with training, feedback, and ongoing practice and support often unavailable or inappropriate (Powell & Wright, 2008). This chapter will focus on how investigative interviewers are trained, with a particular emphasis on the practical training exercises that allow interviewers to apply the theoretical aspects they have learned and practice their interviewing skills for the first time. In Scotland (the context for the current PhD), 'mock' interviews conducted with adult actors that role-play an abused child are used in training. As such, the research on use of adult actors in training simulations will be reviewed, and the potential benefits and drawbacks of their use discussed. Recent research that has looked at different types of training exercises that enhance interviewers' skills, and recommendations on how to maintain gains made during training, are also discussed.

For the literature search, the author's familiarity with the literature, and literature cited by Michael Lamb, Martine Powell and Heather Price were used as a starting point. I then performed searches using Web of Science to find more recent and additional studies, the keywords

"investigative interview" in combination with "training" and "children" were used to find highly-cited papers in the field.

5.1. Investigative Interviewer Training Courses

Child protection workers receive initial training from their employing organisations to prepare them for the wide variety of tasks that they are required to do as part of their job. In order to qualify as forensic interviewers of children, they must undergo additional training to provide them with the knowledge and skills needed to conduct best practice interviews with children. Typically, investigative interviewer courses are short, intensive courses, completed in one stint, lasting for a few days to a week (Powell, Wright & Clark, 2010). While the content of individual courses varies, there are often national standards and curricula that courses are expected to meet (e.g. the National Curriculum, Scottish Executive, 2007). All courses aim to impart background knowledge about issues pertinent to interviewing children such as the dynamics of child abuse, child development, children's memory and suggestibility, question type, interview structure and the role of interviews in the wider context of the legal process. Interviewers may have the opportunity to put what they have learned into practice and conduct mock interviews, on which they will receive feedback from trainers (Aldridge & Cameron, 1999; Warren et al., 1999; Yi, Jo & Lamb, 2016).

Despite the emphasis on educating trainees about the desirability of asking open questions and explaining the effect of the different types of question on children's testimony, most training does not have the desired

170

effect on performance (Aldridge & Cameron, 1999; Craig, Scheibe, Kircher, Raskin, & Dodd, 1999; Davies & Wilson, 1997; Memon, Bull & Smith, 1995; Stevenson, Leung & Cheung, 1992; Warren et al., 1999; Westcott, Kynan & Few, 2006). Instead, as discussed in Chapter 4, field research has consistently shown that the quality of interviews with children is typically poor, with some studies reporting as little as 2.2% of all incident-related questions being open (Lamb et al., 1996a). Instead, interviewers typically rely on option-posing and suggestive questions, and structure their interviews poorly. Mastering the use of open prompts has proved a difficult task for interviewers, with myriad factors working against interviewers, including the natural tendency to ask focussed questions (Powell, 2000) even when receiving little information from witnesses (Powell, 2008). Interestingly, even when interviewers have elicited longer and more detailed responses from children in response to open prompts in the pre-substantive phase of their interviews, they still reverted back to using the more risky focussed questions in the substantive phase (Sternberg et al., 1997).

One of the most important components of investigative interviewers' training is practicing interviewing skills in a controlled environment where feedback can be provided, yet the mock interviews conducted employed on training courses provide inconsistent opportunities to practice the necessary skills. Trainees may be given the opportunity to practice their interviewing skills on fellow trainees, on children recalling a neutral event only (Warren et al., 1999) or on adult actors playing the role of abused children (Freeman & Morris, 1999;

171

Powell, Fisher & Hughes-Scholes, 2008; Yi, Jo & Lamb, 2016). Interviewing real children provides invaluable practice establishing rapport and talking to children, however, eliciting recall of a staged event does not present the same challenges as attempting to elicit disclosures of negative abusive events because, for example, the children didn't require as much prompting (Powell & Wright, 2008).

Powell, Cavezza, Hughes-Scholes and Stoove (2010) compared interviewers' performance in two mock interview contexts (a mock interview with an adult actor playing the role of an abused child and a mock interview with a real child recalling an innocuous event) with each other and with one of their field interviews. They found that performance was relatively stable; if problem behaviours were exhibited in one interview situation, interviewers were likely to exhibit the same problematic behaviours in other interview contexts. However, when it came to questioning behaviour, the adult actor interviews produced performance that was more similar to the field interviews than did interviews in which school children recalled innocuous events. Thus mock interview exercises with adult actors are more beneficial than ones with real children recalling innocuous events because they prompt similar behaviours and questioning strategies as field interviews, thereby providing an opportunity for inappropriate questioning strategies to be addressed during training.

The use of adult actors that role-play abused children has a number of hypothesised advantages. For example, trainees can practice

172

eliciting actual account of abusive events in a low risk environment where any mistakes will not have consequences for safeguarding children or for them professionally. This will enable them to develop confidence in eliciting accounts of abuse and to practice using open prompts and obtaining the level of detail they will need to elicit in a real interview. Specialised scenarios can be role-played and the level of difficulty varied. Trainers can evaluate trainees' competence using objective pre- and post-training measures and can repeat practice until desired competency is reached. The typical format involves trainees observing their peers conducting interviews, allowing additional learning by observation. Actors are readily available (though admittedly costly), their acting experience provides realism to the scenarios and their improvisation skills allow them to adapt to interviewers' questioning get the scenario going in the right direction regardless of any interviewer errors. Questioning an actor is preferable to practicing with fellow-trainees because familiarity between trainees can cause them to directly or inadvertently help fellow-colleagues by providing information quickly and easily. Finally, prompt feedback can be given from trainers while the interview is fresh in their mind and feedback can also be provided by the interviewee which is not the case in the field. However, these are the hypothesised advantages of role-play in general, the following section will review the research on the use of adult actors on professional training courses.

5.2. *Using actors in professional training simulations*

In some professions it is essential that practical skills must be mastered by trainees before they undertake work with their target client

173

base. Particularly when the consequences of making a mistake could have serious ramifications on the health and safety of clients and consequences for the practitioner and their employing organisation (Besharov, 1984). Practicing some of these skills requires practicing on human subjects, however, practicing on real client bases would be unethical, therefore, the use of adult actors that role-play real clients has been adopted. For example, the use of adult actors in training simulations with medical students such as doctors (Rosenbaum & Kreiter, 2002), nurses (Alexander & Dearsley, 2013) and clinical psychologists (Lane, 1989), has long been established as standard practice in undergraduate and graduate training courses. Other fields, such as counselling (Levitov, Fall, & Jennings, 1999; Osborn, Dean & Petruzzi, 2004), and more recently forensic interviewing, have also adopted the method of practicing interviewing skills on either adult actors or fellow-trainees (Lamb et al, 2000b; Yi, Jo & Lamb, 2016). Adult actors work in many different fields, therefore it is unsurprising that who can undertake the role of an appropriate actor, how they prepare/train for the role and whether they provide feedback to the trainees or not is highly variable.

In the medical literature adult actors (known as SPs (standardised patients)), are often required to present complex symptomology and are therefore trained by course instructors on how to believably present these symptoms. Alexander and Dearsley (2013) for example, trained their standardised patients to present medical conditions in interviews with trainee medical students. The actors were emailed a case history of their character, as well as details of how to role-play the symptoms they would

174

be expected to present. In addition, they met with the course facilitator for 45 minute sessions before simulations to discuss how to play the condition and patient, and for training on appropriate responding (actors were instructed to exhibit distress if the trainee failed to show empathy and conversely to open up when they empathised with them). They reported that 96% of the students found participation in simulation enhanced their confidence and that they preferred training with actors to fellow-peers due to the enhanced realism of the simulations.

Counselling courses have also emphasised the advantages of using actors as opposed to fellow-students due to ethical concerns raised about the disclosure of personal information during mock counselling practices (Levitov et al.,1999). In this study adult actors were selected carefully from a pool of undergraduate drama students using a rigorous screening process to choose the actors that were suitable for course participation. The role-play actors' then received training, included providing psychotherapy orientated reading material, character development, and weekly practice sessions with instructors to evaluate performance, discuss future directions and provide improvisation instruction. Their training also included advice on appropriate responding (actors were advised to respond to different counsellors personally and advised on how to play certain aspects of the role-play). Anderson, Gundersen, Banken, Halvorson and Schmutte (1989) also trained undergraduate (psychology) students to role-play clients for their counselling trainees. Two pages of information about their character was provided, followed by actors practicing roles with instructor prior to

interview, and receiving feedback to ensure credibility. They also employed a screening procedure and the authors noted that their students prefered training with actors because fellow-students may have agendas (helping them out). An essential aspect of learning skills is making errors and correcting them (McGeoch, 1942) therefore fellow-participants may be unknowingly preventing their peers from learning.

However, some actors receive no training to play the role of a respondent in professional training exercises. For example, in diagnostic practice interviews mental health students interviewed fellow-participants and psychiatric nurses played psychotic patients (Bogels, 1994). Presumably the course developers felt that psychiatric nurses had enough experience working with psychiatric patients to be able to portray them, and the students said these courses are instructive and effective. This was also the case with Osborn et al., (2004), who used predominantly trained counsellors as the actors on their graduate level advanced counselling courses which the authors say gives "legitimacy" to the process and that the students value their opinion as they are professionals in the field. Lane (1989) used undergraduate theatre majors to play clients exhibiting various clinical symptoms for trainee clinical psychologists' interview sessions. The actors read case studies of clients and with the help of instructor created characters. The interview component of the course consisted of 4 weeks, with the first 2 weeks involving students conducting a ten minute long interview with an actor, progressing to 20 minute interviews over weeks 3 and 4 that increased in difficulty. They conducted a final interview for assessment purposes and

176

students could request more specialised psychopathologies to be played in this one.

Within the field of child protection, Lexton, Smith, Olufemi and Poole (2005) developed two interagency child protection courses that include half-day sessions where hired actors role-play children, adults and various professionals and the trainees choose to participate in the scenario playing themselves and responding as they would do in real life. They also have an opportunity to ask the actors questions while they remain in character (hotseating) about their thoughts and behaviours. The authors emphasise the importance of adequate preparation of the actors; ensuring actors are familiar with the value base and principles of child-centred work, planning scenarios to tackle specific issues, and rehearsal to put these principles into practice in the scenario. Although no specific training was provided to the actors on how to role-play because the authors believe that professional actors already possess improvisation and role-play skills. Participants completed course evaluation forms at the end of the training day scoring a number of aspects from 1-6 and taking an average score from all the evaluations. The courses including actors score only slightly higher than other courses. However, they found that the additional written comments made by participants were very positive, highlighting multiple benefits of role-play with actors, including enhancing professionals' understanding of each other, improving their communication skills, and the usefulness of hearing the actors' outside perspectives as they were reflective of real world opinions that clients may hold.

177

While commentators who use role-play actors on their training courses emphasise the realism and professionalism of their courses, these studies have not evaluated objectively whether or not these simulations have improved trainees' skills. Extensive numbers of child protection training courses exist, however, the evaluative evidence for the effectiveness of training interventions is sparse (Carter, Bannon, Limbert & Barlow, 2006) and a recent review found only 2 studies (Freeman & Morris, 1999; Powell et al., 2008) of investigative interviewing training actually evaluated the effect of training with standardised actors on interviewers' questioning behaviour (Bogo, Shlonsky, Lee & Serbinski, 2014).

Freeman and Morris (1999) trained 3 adult actors to role-play abused children for investigative interviewer trainees. Prior to participating in interviews, the actors received 20 hours of instruction including learned scripted information about a hypothetical child, including abuse history, disclosure details and personal information. They were also educated about developmentally appropriate language and how to resist misleading suggestions. Twelve child protection workers completed a knowledge questionnaire of appropriate interview techniques and conducted an interview with one of actors playing a hypothetical child as their pre-training measure. The child protection workers were then given 6-hours of training (three 2-hour sessions) over 3 weeks regarding the development of memory, question type and interview phases. Within 2 weeks of completing the training, participants completed a second knowledge questionnaire and conducted another simulated interview

(immediate post-training measure). Finally, they conducted another interview at a follow-up 3 months later in an effort to assess whether any gains in training were maintained.

There was a significant increase on trainees' scores in the knowledge questionnaire immediately post-training and this improvement was maintained at the 3-month follow up. They also assessed whether the training had an effect on trainees performance in the simulated interviews and found that the proportion of open ended prompts used by the interviewers increased significantly, but modestly, from 12.2% pre-training to 17.5% post training. Again, this post training improvement was still evident at the 3-month follow up session with the proportion of open ended questions at the follow-up totalling 19.3%. However, no changes were found in the proportion of appropriate questioning or inappropriate behaviours at any time point post-training. While the study showed improvements in participants' knowledge on the topic and several interviewing skills, the training failed to impact several key interviewing capabilities, such as the types of questions asked or the length of the interview.

In an effort to investigate whether or not trained actors enhanced child protection workers use of open prompts, Powell et al., (2008) provided interviewers with a 2-hour training session in the use of open-ended questioning and two 2-hour interview practice session with a trained adult actor. A separate group of interviewers received the training session but conducted their practice interviews with a fellow-trainee. The

adult actors were psychology post-graduate students with knowledge of children's eyewitness testimony and forensic interviewing who completed 25 hours of training over 12 weeks that involved developing hypothetical child details and scripts, standardising role-play (developmentally appropriate language, responses consisting of no more than 4 details and use of pauses) and rehearsal and feedback until all 3 had similar performances. All interviewers conducted a mock interview with one of the actors both before and after training in order to obtain base-line and post-training measurements for comparison.

Immediately after training, interviewers asked a greater proportion of open-ended questions than pre-training, and the interviewers who practiced with the psychology post-graduate students used more open-ended prompts than those who trained with a fellow participant. Twelve weeks later the participants interviewed an actor again at a follow-up assessment, the number of open-ended prompts decreased over time for both training groups, however, those who practiced with the psychology post-graduate students still used more open-ended prompts than their counterparts. This is likely because the trained actors responded to open prompts with a greater proportion of event-related concepts during the practice sessions than fellow-participants, and were more likely to disclose abuse in the middle of the practice interview while fellow-participants more likely to disclose at the very beginning. The trained actors 'rewarded' interviewers with information when they used open-ended questions which in turn may have reinforced the interviewers' perceptions of the value of open-ended questions.

180

While there are many hypothesised advantages of using adult actors in interviewer training simulations, and the afore mentioned studies show they have enhanced interviewers' performance on some measures, most courses have failed to use objective methods to test what actors are actually doing. For example, Powell et al. (2008) reported the amount of event-related concepts actors responded with, but this still does not reveal exactly how the actors were responding to the different types of prompts because questions were categorised broadly as either open or specific. Therefore, we do not know whether the length of responses or the number of details provided by the actors was realistic in terms of how children actually respond to the different types of questions in forensic interviews.

Further, these two studies looked at the effectiveness of training with adult actors that had undergone extensive training prior to the role-plays. It is not always the case that the actors on forensic interviewing courses receive any such training to play the role of a child victim, and therefore we do not know whether or not they are reinforcing interviewers' use of appropriate questioning. Indeed, concerns have been raised about the use of adult respondents in mock interviews because their responses in simulated interviews often reflect superior memory and language skills than that typically portrayed by child witnesses in investigative interviews (Powell, 2002). Therefore, the findings of these studies may not be applicable to training programmes that use untrained adult actors.

Other issues have been raised with regards to the use of adult actors on training courses, one being that actors often provide feedback following simulations (Alexander & Dearsley; 2013; Lane, 1989; Levitov et al., 1999). Woodward (1998) points out that giving feedback itself is a learned skill, and accordingly provide training to their SPs on how to provide feedback in a constructive manner that is both sensitive and specific. Similarly, Osborn et al., (2004) provided specific feedback questions to their actors to designate what kind of feedback they should give. If actors have no knowledge of forensic interviewing they may provide feedback that encourages inappropriate behaviours using a 'real world' opinion as opposed to a professional one. Some commentators have reported that actors report finding it hard not to provide monologues in silences and that trainers and actors can become so involved in making an engaging theatre experience that it has little relation to the real world (Lexton et al., 2005). Finally, Liao, Kao, Liang and Hsieh (2015) point out that multiple performances a day can be mentally challenging for actors and there is a need to maintain accuracy and consistency as actors below par performances can invalidate an examination and therefore the same could happen in interviewing scenario.

5.3. *Maintaining training gains*

In addition to multiple practice opportunities, research recommends training programmes should include train interviewers in the use of a structured interview protocol (Cyr, Dion, McDuff & Trotier-Sylvain, 2012; Cyr & Lamb, 2009; Lamb et al., 2009; Orbach et al., 2000; Sternberg et al., 2001b; Yi, Jo & Lamb, 2016), provide expert ongoing

182

feedback (Cyr et al., 2012; Krause, Pompedda, Antfolk, Zappalá, & Santtila, 2017; Lamb, Sternberg, Orbach, Esplin & Mitchell, 2002a; Lamb et al., 2002b; Price & Roberts, 2011), and spaced learning opportunities (Poole & Lamb, 1998). Recently developed training programmes that include all of these components have been shown to increase interviewers' use of open prompting and reduced their use of option-posing questions (Cederborg, Alm, Lima da Silva Nises, & Lamb, 2013; Price & Roberts, 2011). Cederborg et al., for example, combined the NICHD protocol and the PEACE interviewing model, spaced the training out over 6 months, provided extensive supervision and feedback in a variety of forms (verbal and written) and showed trainees how to evaluate their own interviews.

As it has proven so difficult to change interviewers' questioning behaviour by teaching them about the theoretical issues behind question types alone, researchers have examined additional ways to improve their adherence to using open prompts. Tasks that give interviewers practice distinguishing and categorising questions using an objective coding scheme increase their use of open prompts in mock interviews (Yii, Powell, & Guadagno, 2014) and field interviews (Cederborg et al., 2013). Coding interviews may promote a deeper understanding of question type because, in order to categorise questions, coders must focus their attention on the structure of different prompt types. Learning to code interviews also allows investigators to evaluate their own work, which is one of the most useful elements of training highlighted by trainee interviewers (Powell & Wright, 2008).

Yii et al., (2014) found a positive correlation between interviewers' performance conducting mock interviews and performance on a written quiz that required participants to generate open prompts. They also found that interviewers who used more open prompts in their mock interview performed better in a computer simulated task by choosing a greater number of the best practice questions from computer generated options. However, consistent with previous findings that knowledge alone of best practice does not necessarily affect performance in the field (Warren, et al., 1999), performance on a knowledge quiz was not associated with performance in the mock interview. Instead, this suggests that training interviewers to identify and use open prompts may be more effective than knowledge-based training.

Online training activities have been successful in enhancing interview quality (Benson, & Powell, 2015; Brubacher, Powell, Skouteris, & Guadagno, 2015; Krause, Pompedda, Antfolk, Zappalá, & Santtila, 2017; Pompedda, Zappalà, & Santtila, 2015; Powell, Guadagno & Benson, 2014; Yii et al., 2014). They are a cost-effective way for interviewers to practice their skills, they also allow for spaced practice opportunities and are flexible with trainees completing the activities in their own time at their own pace. For example, Powell et al., (2014) found that interviewers' use of open prompts in mock interviews with trained adult actors significantly improved after they had completed approximately 36 hours of online activities, centred on eliciting best practice accounts of abuse. Tasks included learning a question-coding protocol and using it to code transcripts, conducting computer-simulated

interviews of virtual children, taking quizzes, assessing one's own and others' work and viewing examples of best practice. This improvement was still evident at a follow-up assessment 3-6 months later. These findings have even been replicated in field interviews with real children and the training gains were still evident 12 months later despite no further intervention or supervision (Benson, & Powell, 2015).

5.4.   *Overall chapter conclusion*

In sum, training programmes that focus mainly on knowledge of open prompts and best practice questioning alone do not effectively equip interviewers to conduct best practice interviews with children (Warren, et al., 1999; Yii et al., 2014). Training programmes that provide multiple opportunities to conduct practice interviews and receive feedback over spaced learning intervals, may be the key to improving interviewers' proficiency using open prompts (Cederborg et al., 2013; Cyr et al., 2012; Krause et al., 2017; Lamb et al., 2002a; Lamb et al., 2002b; Price & Roberts, 2011). Further, the type of training exercises employed should be carefully considered. Conducting mock interviews with trained adult actors can be very effective (Freeman & Morris, 1999; Powell et al., 2008), however, in order for actors to reinforce the use of best practice questioning, training should be provided for the adult actors. Actors should also be instructed in providing appropriate feedback if it is a requirement on the training course (Woodward, 1998). Exercises that involve generating and categorising prompts (coding interviews) promote deeper learning and understanding of question type (Cederborg et al., 2013; Yii et al., 2014). Further, these and other performance enhancing

185

tasks, such as practicing interview skills with an avatar, can be done cost-effectively and flexibly online (Benson, & Powell, 2015; Brubacher et al., 2015; Krause et al., 2017; Pompedda et al., 2015; Powell et al., 2014; Yii et al., 2014).The next chapter will examine the structure and content of initial child interviewers' training courses in Scotland in an effort to determine whether or not these courses adhere to the best practice recommendations made above.

Chapter 6: Investigative interviews and investigative interviewers' training in Scotland

This chapter describes the practices and procedures for interviewing child victims and witnesses, as well as those for training investigative interviewers of children, presently used in Scotland. The limited research that has assessed the quality of field interviews and investigative interviewers' training interviews in Scotland will be reviewed. Specifically, because adult actors role-play children in mock interviews on joint investigative interviewing training (JIIT) courses in Scotland, the use of adult actors to role-play children on JIIT courses is also discussed. Based on the forgoing review of memory development and forensic interviewing, the strengths and limitations of the current approaches in Scotland are discussed and conclusions and recommendations for future practice are made.

6.1.   First appearance of guidelines and training in Scotland

Despite the importance of conducting investigative interviews with witnesses and victims of crimes appropriately, there was a period of time in Scotland (as was the case in the rest of the UK and indeed many other countries) when there were no standardised evidence-based guidelines or formal training methods for investigative interviewers. Instead, interviewers' training procedures for police officers comprised largely of watching more experienced colleagues conduct interviews (Moston & Engleberg, 1993). It was not until 2003 that the Scottish Executive published their child interviewing guidelines ('Guidance on Interviewing

Child Witnesses in Scotland') and a further three years until the standardised child interviewer training 'JIIT' was developed collaboratively by ACPOS (Association of Chief Police Officers in Scotland), the Association of Directors of Social Work, Scottish Children's Reporters Administration, Crown Office and Procurator Fiscal Service and the Scottish Government (2006). These guidelines were based on the conclusions of international child interviewing research and included a shortened version of the NICHD Protocol (see Appendix A) for interviewers to use during forensic interviews. While interviewers are trained locally and the course structure varies between jurisdictions, each course is expected to be designed, delivered and run in accordance with the content and framework set out in the National Curriculum (Scottish Executive, 2007).The 2003 guidelines have since been updated and the current procedures for conducting investigative interviews with children are laid out in the Scottish Executive (2011) guidelines 'Guidance on Joint Investigative Interviewing of Child Witnesses in Scotland'. The updated guidelines no longer contain the NICHD Protocol sample.

6.2. *Joint Investigative Interviewing Training*

Joint Investigative Interviewing training is a week long course; this can be a continuous week or split into different sessions, so long as the five days are completed in order. First trainees learn about theoretical issues pertinent to children and conducting developmentally appropriate interviews, such as memory development, suggestibility, communicative issues, question type, interview structure and the Scottish Executive Guidelines, by attending presentations delivered by course instructors

188

and invited professionals. The second part of the course is focussed on practical training exercises and trainees have the opportunity to put theory into practice and conduct one or two practice interviews with adult actors. They receive feedback on their practice interview(s) from course instructors and also learn about the role of the interview in the wider context of criminal proceedings (some courses include a mock court scenario on the final day). The practice interviews are conducted as though they are real investigative interviews with a police officer and a social worker being given a scenario in advance and having to plan the interview accordingly.

At present, the experience trainees have is variable; not only between jurisdictions, but also within individual jurisdictions because the practice interview scenarios trainees receive are widely variable in a number of ways and trainees receive different scenarios so they are able to watch other trainees' interviews. The scenarios provided for role play vary in complexity from single incidents of physical abuse to repeated rapes by multiple perpetrators. They also vary by age (4–to 15-years-old), status (victim or witness) and willingness to disclose (willing or unwilling). Some scenarios contain a 'red herring', which involves the interviewer being misled from the start by the interviewee, who will make a false allegation. Therefore, some trainee interviewers have a much more difficult experience during training compared to some of their peers. Further, until very recently trainees in some jurisdictions were expected to scribe (keep a handwritten 'verbatim' record) of the interview, while other

jurisdictions had access to visual recording equipment, negating the need for trainees to scribe.

6.3.   The use of actors in interviewer training in Scotland

The practical component of interviewers' training- the mock interviews with adult actors- are held in high esteem by trainee interviewers. Goetzold (2015) interviewed JII trained interviewers about their self-evaluated competencies and experience of undertaking JIIT training, 13 out of the 16 interviewers spontaneously mentioned the role-play interviews and the feedback received about them as being particularly helpful in preparing them to conduct field interviews. Indeed, another survey of police officers found that 25% of their respondents believed more practice interviews would improve JII training (La Rooy et al., 2011). Training organisers also believe the actor interviews to be a very important component of the interviewer training as some commentators have attested:

"Using Interact Roleplay has given realism and quality to our scenarios. They work together with our team at Tulliallan, making the training courses realistic and enjoyable as well as professional." (Brian Rodgers, Scottish Police College). (Deadline, January, 2010)

"Interact Roleplay provides a highly professional and dependable service which is an integral part of the training we provide. The experiential learning provided in the scenarios helps our students make crucial connections between theory and practice. We consistently receive feedback which comments on the high calibre of the actors and the

190

quality of the learning experience they provide." (Mark O'Donnell, Joint

Investigative Interview, Training Coordinator, Strathclyde Police,

(http://www.interactroleplay.co.uk/casestudy5.html). (Interact Roleplay

website)

[the use of actors] *"can enhance professionals' understanding of each*

other and improve their communication skills for the benefit of the

children and families they are supporting". (Lexton et al., 2005, p. 203)

However, not all interviewer trainers agree that the experience

actors provide is always entirely realistic. For example, some have

commented that the actors appear to provide just as lengthy responses to

open-prompts as they do to closed questions (personal communication

with interview trainers R. Morris & M. Lambley, 2011). Researchers have

also voiced concerns that using adult actors to role-play children has its

disadvantages, for example, Powell et al., (2008) have claimed that adult

actors do not respond in the same as children would in a real interview.

These concerns are not necessarily surprising because actors are not

required to have any specialised knowledge of interview guidelines or any

other aspects of forensic interviewing. Studies in other countries have

found that practising with a trained respondent can increase the

interviewers' use of open questioning because these actors provide

appropriate responses and reward interviewers with information when

they ask these best-practice questions (Powell et al., 2008). Such training

studies have provided lengthy training to ensure standardised

performance from the actors (Freeman & Morris, 1999; Powell et al.,

191

2008). However, this is not the case in Scotland. Therefore, while at first glance it may indeed appear useful to have interviewers conduct interviews with actors, we do not actually know how JIIT actors in Scotland are responding to interviewers' questions and whether or not they are enhancing the skills of the interviewer as intended.

A single study to date has investigated the quality of JIIT practice interviews conducted during training. Goetzold (2015; strand a) assessed the role-play interviews conducted with adult actors on a JIIT refresher course. She reported having 'concerns' about 18 out of the 21 interviews and said that the sample showed an overall lack of open questioning and a reliance on closed questioning. Further, only two interviews contained an attempt at a practice interview and the closure phases lasted 1 minute 34 seconds on average with interviewers failing to convey all of the recommended closure practices, despite having time remaining to conduct a full closure session. It is extremely important to note however that no information was provided by the author with regards to either the categorisation or the frequency of the different question types so we do not actually know what constituted open and closed questions or how often they were used. Further, the interviewers in this sample were on their JIIT refresher training, meaning they had already passed JIIT training at an earlier time point and were returning to strengthen their interviewing skills. Therefore, we do not know if these findings are applicable to JIIT trainees because the original training gains made by participants in this study may have dissipated between training and refresher training explaining their poor performance. Finally, no analyses

192

were made on the actors' responses, therefore, to date there have been no studies published that have investigated untrained adult actors' responding to different question types in mock forensic interviews.

Despite the long-standing international consensus on how to best interview children, the current procedures for conducting interviews with children and training interviewers in Scotland contain some aspects of practice that may not be conducive to best practice. Firstly, the update to the interviewer guidelines has seen the removal of the NICHD Protocol sample. The Protocol has been shown to increase interviewers' use of open prompting and improve the quality of children's responses (Lamb et al., 2008). Its removal has the potential to decrease the level of quality of interviews that were being conducted in Scotland before the update and it is recommended that the sample protocol be re-instated. Secondly, while the use of trained adult actors can be beneficial in training (Powell, et al., 2008), the actors in Scotland are untrained for this role. Further, their responding in mock interviews has elicited concerns from interviewer trainers. Therefore, the actors could unknowingly be having a negative impact of interviewers' questioning. It is recommended that training in appropriate responding is developed for the actors and that they must complete this training before participating in JIIT.

Thirdly, inconsistencies between courses, such as access to VRI facilities, means some trainees are being disadvantaged compared to their peers, for example, scribing the interview means the interview pace is slowed and these trainees have less time to question their interviewee

than those that do not have to scribe. It is recommended that VRI facilities are mandatory on all interviewers' training courses and in all forensic interviews with children. Finally, even within interviewer training courses the scenarios used in the mock interviews are radically different. This means interviewers do not get an equal opportunity to practice their interviewing skills (e.g. if their interviewee is displaying reluctance), it is recommended that scenarios are standardised with regards to level of difficulty at this early stage.

6.3.1. *Summary*. In spite of the shortcomings of Joint investigative Interviewing Training courses, these courses are held in high regard by trainee interviewers and by the organising police bodies (Goetzold, 2015; La Rooy, Lamb & Memon, 2011; personal communications). In particular, the practical exercises, the mock interviews with adult actors, are believed by participants to have enhanced their interviewing skills and prepared them well for the task of conducting forensic interviews with children (Goetzold, 2015; La Rooy et al., 2011). While it has been demonstrated that role-play interviews with adult actors can improve interviewers use of best practice questioning, extensive training has normally been provided to the actors to prepare them for this role (Powell et al., 2008), adult actors that participate in JIIT courses receive no such preparation. To date only one study has assessed the quality of role-play interviews conducted on a JIIT course in Scotland. Here, the author reported that interviewers' substantive questioning behaviour was poor with a lack of open prompting and a reliance on focussed questions, and that most interviewers failed to conduct a practice interview and

194

appropriate closure (Goetzold, 2015). However, this study did not provide any information on how interviewer questions were coded, the frequencies of the different question types or information regarding the actors' responses. Further, the JIIT interviews were conducted on a refresher course and the findings may not be applicable to first time JIIT trainees. Therefore, we do not presently know the quality of interviewers' questioning or how adult actors respond in JIIT interviews in Scotland.

6.4. *Field research*

In addition to a lack of research on interviewers' training in Scotland, there is a paucity of field research in Scotland. The first piece of research to examine interviewers' behaviour in the field was La Rooy et al.'s (2011) survey of police officers in in Scotland's self-report interviewing behaviours. Interviewers reported beliefs that many of the practices laid out in the interviewer guidelines were ineffective and subsequently reported not following them. For example, 52% of respondents indicated that they believed that practice interviews were either 'not very effective' or 'not at all effective' and 87% admitted that they 'never' or 'rarely' conducted one.

With regards to what is arguably the most important practice in interviewing children- asking open prompts in the substantive phase-,12% of respondents stated that they believed open prompts were not effective in eliciting information from children and 20% claimed that they 'never' or 'rarely' used open prompts during their interviews. In fact, only 43% of interviewers said that they 'always' or 'almost always' used the best-

195

practice open prompts that are endorsed in their interviewer guidelines. Conversely, 85% of interviewers indicated that the riskier closed Yes/No questions were 'quite', 'very' and 'always' effective and 17% thought that leading and suggestive questions were 'quite effective' despite their guidelines discouraging interviewers from using them.

Another concern that was flagged by interviewers in the survey revolved around the scribing (a second interviewer keeping a handwritten 'verbatim' record) of interviews. The interviewers reported scribing negatively impacted the quality of their interviews for reasons including illegible handwriting, having to wait for scribes to catch up, creating pauses and scribes missing details. Importantly, interviewers reported that when scribing it is difficult to note the lengthy responses to open prompts and scribes prefer more direct and focused questions because they are quicker to write down. In 1989 the Pigot Report recommended that all child interviews be visually recorded so that video evidence could be used in court eradicating the need for children to appear in court for examination-in-chief. As a result of this, this measure has been in place since 1992 In England and Wales. In Scotland this practice has been variable between police forces, with some jurisdictions using VRI (video recorded interviewing), some using audio recording and others relying on scribing alone.

Research that has examined contemporaneous handwritten 'verbatim' interview records suggests that such records are unlikely to be accurate (Cauchi & Powell, 2009; Cauchi, Powell & Hughes-Scholes,

2010; Lamb, Orbach, Sternberg, Hershkowitz & Horowitz, 2000). For example, Lamb et al., (2000) compared 20 verbatim contemporaneous handwritten investigative interview records with the audio recording of the interview and found that 57.3% of the total interviewer utterances and 25% of the details elicited from children were missing from the handwritten account. Further, 56% of the details in children's responses were attributed to the wrong eliciting utterance type in the investigators' notes, with a tendency to misattribute details to more open rather than more focused prompts. Cauchi and Powell (2009) examined handwritten disclosure interviews (initial complaint interviews conducted with children in order to determine whether or not an official investigation will be undertaken) and found 46% did not contain any interviewer questions, instead the children's evidence was recorded in one first person account. When interviews did record the interviewer's questions, a mean proportion of .93 questions were recorded in a way that allowed the type of question asked to be distinguished, the rest could not be discriminated e.g. due to missing question stems.

The experimental research findings compliment these field findings. For example, Cauchi et al., (2010) asked police officers to record a 15-minute long live interview script, which was read out to them at a pace of approximately 2.2 words/s, in the format they would usually handwrite an interview with a child. Analysis of the interview records showed that 12% of interviews did not contain a record of a single question. Of the potential questions that could have been recorded, a mean proportion of .37 questions were recorded either entirely or partly,

197

with.39 questions truncated to one or two words. On average, only 61% of abuse-related details were recorded by the note taker and fatigue effects were evident (i.e., the number of questions and responses recorded declined as the interview progressed). Even under optimal conditions the typical rate of the spoken word is much faster than the rate of the written word (Piolat, Olive & Kellogg, 2005), therefore, interviewers that are expected to scribe interviews are facing a formidable challenge and risk losing forensically relevant information.

Despite child interviewing guidelines being in existence in Scotland since 2003, to date, only two studies have analysed the quality of field interviews conducted with child victims and witnesses by JIIT trained investigators in Scotland (La Rooy, Earheart & Nicol, 2013; La Rooy, Nicol, Halley & Lamb, 2012). La Rooy et al., (2012) analysed the quality of 37 interviews conducted with children aged 4–to 13-years-old between 2003 and 2011, under the Scottish Executive 2003 guidelines. Interviews were transcribed from either 'scribed' handwritten contemporaneous 'verbatim' notes taken during the interview by the second interviewer or from electronic recordings of the interview if the jurisdiction conducting the interview had access to such equipment at the time. Presence and absence of some ground rules were noted, the importance of telling the truth was discussed in 43% of the interviews with 30% of interviews containing a practice demonstration of the child's understanding of truth and lies. Twenty-four per cent of interviews contained the instruction to let the interviewer know if they did not understand a question, 22% of interviews contained instruction not to guess but to instead say 'I don't

198

know' and 3% of interviews conveyed that the child could and should correct interviewers if they made a mistake. Not one interview contained the recommended practice interview before asking children substantive allegation-related questions. Analysis of the substantive phase revealed that only 8% of the questions asked were open prompts. The rest of the questions were comprised of 39% directives, 36% option-posing and 17% were suggestive.

The follow-up to the 2012 field study analysed 19 investigative interviews conducted in 2012 in Scotland under the guidance of the Scottish Executive 2011 guidelines (La Rooy et al., 2013). This study found that with regard to the practice interview, no improvements had been made; there were still no practice interviews conducted in any of the interviews. The number of open prompts used had improved, rising to 15%, the rest of the questions being made up of 49% directives, 34% option-posing and 2% suggestive. The authors found modest improvements had been made in the communication of some of the ground rules, with the exception of demonstrating the child's understanding of the concept of truth and lies using a practice question, which was absent in all interviews at this time. A discussion of the importance of telling the truth was present in 74% of interviews, 58% of interviews instructed children to let interviewers know if they did not understand a question, 4% included the 'I don't know' instruction instead of guessing and 16% of interviewers they told the child to correct them if they made a mistake.

199

These findings are disappointing given that the Scottish national guidelines are based on the NICHD Protocol and emphasise all of the components of best practice interviewing according to the research consensus and that they place such a value on using open prompts. There could be many explanations for the poor field performance. Firstly, the survey revealed that interviewers do not value conducting practice interviews or using open prompts, therefore, the problem may be that interviewers are choosing purposefully to not use open prompts rather than reflecting lack of awareness of best practice. Secondly, the concerns regarding inability to scribe open prompts and their responses in real time may have pushed interviewers to use a less desirable questioning style in an effort to keep up with recording. Finally, the protocol validation studies have shown that merely providing best practice guidelines and a protocol to follow does not necessarily aid interviewers to conduct best practice interviews. Instead extensive training is required to bring about change in interviewing behaviour, thus, it may be aspects of interviewer training in Scotland that are not providing trainee interviewers with the necessary skills to conduct high quality interviews.

While an apparent improvement in use of open prompts was made between the two studies, it is important to note some limitations of the samples before drawing any conclusions. Firstly, both samples were small in size. Secondly, the 2012 study included interviews that had been conducted over a wide period of time (8 years), some of which were at the introduction of the very first set of national interviewer guidelines, while the second study included interviews conducted over a shorter

200

period of time during which national guidelines had been in place for almost a decade. Lastly, the field samples used La Rooy et al. (2012, 2013) were not random samples, the interviews were all proceeding to court and had been referred to the first author by solicitors for his expert opinion on their quality. Therefore, the interviews examined as part of the study may have rendered them particular low quality and therefore they may necessarily reflect the state of interviews on a national level.

Despite the use of such forensic aids such as anatomically correct dolls and human figure diagrams generating a lot of controversy within the scientific community, mostly centred on the issue of suggestibility and interviewers' subjective interpretations of play (Bruck et al., 2000; Bruck et al., 1995; Goodman et al., 1997; Saywitz et al., 1991) the Scottish Executive guidelines state that use of anatomical dolls and body diagrams is appropriate 'as long as adequate training has been provided'. However, in Scotland there is no standard protocol for appropriate use of anatomical dolls or HFDs and no training in their use, yet they are still used in field interviews with children. Therefore, it is unknown how interviewers are actually trained to use them and the skill level of interviewers using these aids is also unknown. Many other unsupported practices have been found in interviews conducted in Scotland including asking reluctant children to 'write down a word' when they do not disclose abuse, requesting children to draw pictures relating to the incident, presenting materials allegedly drawn/written down by the child outside of the interview room, the use of a 'secrets jar' and the use of alleged telepathic communication. While these practices are not common, they

201

can be detrimental to the quality of the interview and the credibility of the evidence.

6.4.1. Summary A number of worrying findings were elicited from the scarce pool of field research conducted in Scotland. Interviewers' self-reports showed they did not believe in the value of two of the most important techniques shown by research to improve the quality of evidence elicited from children. They did not think using open prompts and conducting a practice interview were effective and even admitted to not using these techniques when they interviewed children (La Rooy et al., 2011). Analysis of two samples of field interviews confirmed that in practice interviewers relied on focussed questions, used few open prompts and did not conduct practice interviews when questioning children (La Rooy et al., 2012; 2013). However, due to the potential skewedness of the two field studies we really cannot be sure of the quality of interviews conducted nationwide with children in Scotland. There is an urgent need for more in depth analyses of a representative sample of field interviews in Scotland in order to assess the standard of investigative interviews conducted with children in this country. Further, unlike analyses of field interviews in other countries, these studies did not conduct any analyses on the responses of the children; it is possible that the characteristics of the sample meant that reluctant children were over represented, leading interviewers to move to using focussed questions in a bid to elicit information. Further, techniques such as the use of body diagrams and other unsupported practices are still used in Scotland, therefore, future field research must also include analyses of children's

202

responses and the prominence of unsupported interview techniques as well as ones of interviewers' questioning.

6.5. Overall chapter conclusion

Research conducted in Scotland has shown that investigative interviews with children in this country are of a low quality (La Rooy et al., 2012; 2013). Interviewers in this country have reported a lack of faith in best-practice techniques such as the practice interview and the use of open questioning and found it difficult to record open prompts due to scribing interviews (La Rooy et al., 2011).The only study conducted looking at the quality of JII training interviews has revealed that even at this early stage interviewers are using inappropriate questioning strategies and infrequent use of open questions (Goetzold, 2015). No study to date has examined the quality of the adult actors' responses in Scotland, despite the fact that they receive no training to prepare them for this role. Therefore, it is of the utmost importance that in depth evaluations of both interviewers' questioning and actor responding in JII training interviews are evaluated. This has never been done before and these interviews are the first opportunity to observe 1) whether interviewers are complying with the guidelines at this stage or not, and 2) whether or not actors are reinforcing best practice questioning and responding in ways that are similar to real children in terms of length and level of detail of responses. Therefore, the next chapter, the first study of this thesis, is an analysis and a comparison of trainee interviewers' mock interviews conducted during JIIT with adult actors at 2 different training sites in Scotland. Quality analyses were conducted on both the

203

questioning behaviour of the interviewers and, for the very first time, on

the actors' responses, specifically the length and number of details

contained in their responses.

Chapter 7. Examining the quality of Scottish Investigative Interviewers' training interviews with adult actors

7.1   Study aim and rationale

There is a lack of research into the quality of investigative interviews of children in Scotland, and only one study to date on interviewers' training that did not examine interviewer's use of questions empirically (see Goetzold, 2015). In light of this, it is important to evaluate systematically whether JIIT prepares interviewers to conduct high-quality interviews. Initially, a sample of 100 field interviews of children aged between three and thirteen was requested from a Scottish police force for analysis. Although the force initially agreed to share this data, eventually they declined to provide the sample of interviews. However, researchers can analyse data from JIIT itself in order to examine interviewer competencies at training, as JIIT is the single opportunity trainees have to conduct a joint interview from start to finish before they may be called upon to interview a child in a real case. Indeed, it is essential to identify lack of adherence to best practice guidelines at training, as poor practices by interviewers in the training room can generalize to the field, such as inappropriate wording of questions when interviewers question adult actors during mock interviews and real children in field interviews(Powell, Cavezza, Hughes-Scholes&Stoove, 2010).This is the first study to both systematically examine the quality of JIIT in Scotland, and importantly, the first study to examine the responses of actors who play the role of a child during JIIT. In addition, interviewer behaviours during the non-

substantive phase of the interview (e.g., when rapport is being established with the interviewee) were also examined. Measuring actor behaviour is important in light of concerns about the utility of adult actors who play the role of a child in this context, for example, when responding to specific types of questions from the interviewer in ways that encourage and reinforce the trainees use of best-practice questions (reviewed in Chapter 5.2).

Thus, the aim of this first study was to provide the first systematic analysis of the quality of JIIT role-play interviews. This involved examining both the behaviour of the trainee and the actor and the contrasting performance across two different Scottish jurisdictions in order to test whether trainee behaviour is consistent across locations because there are suspicions it is not (see The Evidence and Procedure Review by the Scottish Courts and Tribunals Service). Here, interviewer performance was examined according to 1) whether they adhered to the introductory 'ground rules' when initially interacting with the actor, 2) whether interviewers used narrative elaborative training as a form of practice interview, 3) the types of questions used during the substantive phase of the interview (i.e. recommended open prompts versus riskier focussed prompts), 4) the proportion of questions posed to the actor before focussed prompts were used (i.e. directive and option-posing) and 5) the use of appropriate best-practice rules during the closure phase of the interview. Actor performance was examined according to 1) their use of informative versus uninformative responses to interviewer's questions, 2) the overall and average length of response and overall and average

number of details yielded per question (i.e. richness) in response to different question types (i.e. did they reinforce best practice by the interviewer), 3) the type of question that elicited the first detail from the actor and 4) the proportion of details provided before the first focussed prompt.

Given data from JIIT in Scotland has not been analysed systematically to date, many of the analyses are, by definition, exploratory and are run to confirm whether interviewers use individual best-practice rules at levels greater than would be expected by chance, whether interviewers apply these rules consistently in training in two different jurisdictions and whether actors respond in a consistent manner in training courses held at two different police jurisdictions, despite being hired from separate acting companies. However, based on reviewing the limited investigative interviewing research in Scotland from the field and wider research on the efficacy of training courses for investigative interviewers, some *a priori* hypotheses were derived. Firstly, high proportions of best-practice open prompts have been observed immediately post-training (e.g., 73% open prompts, even when interviewing untrained respondents; Powell et al., 2008), therefore it was predicted that interviewers would follow best-practice guidelines and use open prompts to a greater extent than focussed prompts in light of their training. The use of recommended ground rules at levels greater than chance was expected.

Secondly, the second force within the sample, in contrast to the first, conduct role-play interviews as part of a more intensive training

course (a continuous week-long course) and observe other trainees in role play. In light of this it was predicted that there would be an association between use of best practice guidelines in training and the force at which the training was administered. Force two interviewers were expected to conduct higher quality interviews during the substantive phase (e.g., use a greater proportion of open prompts and delay focussed questioning until relatively later in the interview) in comparison to force one interviewers. Finally, although no research to date has examined the behaviour of JIIT role-play actors, other researchers have raised potential concerns about the authenticity of adult actors who play the role of child, for example, by responding to interviewers in a way that would not be typical of a child (Powell, 2002; Powell & Wright, 2008). Indeed, JIIT trainers have reported concerns that adult actors do not respond appropriately according to question type in the same way that children might (e.g., providing lengthy and rich responses to option-posing questions and failing to provide such answers to open prompts; see Chapter 6).Thus, the length and richness of actors' responses to the different question types were compared in order to verify whether such concerns were supported by analyses of JIIT data.

7.2   Method

7.2.1   Participants. Forty-two child protection workers took part in the study. Participants (trainee interviewers) were offered the opportunity to take part in the study while undergoing JIIT, and provided written consent

at the beginning of the course to take part. Two police jurisdictions in Scotland took part in the study (Force 1: N=22, Force 2, N=20), which included police officers (N=17) and social workers (N=25), 17 of whom were male. Data were collected from five JIIT courses between 2009 and 2011. All adult actors who took part in the training (Force 1: N=3, all female. Force 2, N=2, 1 of whom was male) were professional hired actors employed to role-play in JIIT courses. All actors held a professional acting qualification and were currently working as actors. Each force hired actors from a different acting company.

7.2.2 *Detailed outline of training procedure.* Each interviewer was provided with a copy of the "Guidance on Interviewing Child Witnesses in Scotland" (Scottish Executive, 2003) at the start of their training. Over the course of their training the interviewers were familiarised with issues to do with child abuse, child development, memory and suggestibility and how to conduct best practice interviews. Specific reference is made to Appendix A of the guidelines, which provides a structured interview protocol for interviewers based on the recommendations of the NICHD. At the start of the role-play sessions, interviewers were provided with hypothetical scenarios of child abuse and were given half an hour to plan their interview. Scenarios were created by the trainers and were adapted from real cases that had occurred in their jurisdictions. Interviewers were provided with the name and age of their 'child interviewee', some information about their family background, the alleged allegation and how the abuse came to light. The actors were provided with additional information including the type of abuse that had occurred, whether the

abuse was an isolated or a repeated event, the number of prior incidents of abuse and the child's willingness to disclose. Of note, some of the scenarios provided to the actor included instructions to make a false accusation and some actors were instructed to play the role of a child who is unaware why they are being interviewed.

The Scottish Executive (2003) guidelines require that forensic interviews are conducted both by a social worker and a police officer. One interviewer takes the lead role and asks questions while the other keeps a handwritten record of what is said by the lead interviewer and the child. As Force 1 had access to Video Recorded Interviewing (VRI) equipment in their jurisdiction, these interviewers did not follow the guidelines. Instead, they were given half an hour to complete one full interview (from introduction to closure) that they conducted alone and were video recorded for evaluation. Force one used four different scenarios of hypothetical child abuse (three different scenarios in 2009, and the same scenario in 2011 and 2012). All scenarios described females ranging from 5-10 years old, with three alleging sexual abuse and one having witnessed domestic abuse.

Force 2 interviewers conducted their training in pairs with one police officer and one social worker. Trainees participated in two interviews over a two-day period, and had the opportunity to act as lead interviewer in one interview and scribe in the other interview. Interviews were audio recorded for the purpose of this study. Trainees were split into two groups and trained in two separate rooms in order that scenarios could be used without trainees becoming aware of the specific details of

210

each scenario. On day one the scenarios included a 5-year-old boy alleging physical abuse by his father and an 8-year-old girl alleging digital penetration by her father, with both children willing to talk and aware of the reason for their interview. Each group watched the other members of their group interview the actor via videolink. On the morning of day one each pair of interviewers began by conducting 15 minutes of rapport building after which they took a break to discuss their performance with colleagues who viewed the interview from another room, in order to gain peer feedback on how to proceed. Trainees then returned in the afternoon and had 20 minutes to finish the interview. At the end of the training day, interviewers received group feedback from both actors.

On the second day, the trainee pairs were given 45 minutes (without a break) to conduct an entire interview. The scenarios on the second day involved children who either did not know why they were there, were unwilling to disclose information or deliberately misled the interviewers. For example, one scenario involved a 4-year-old boy who did not know why he was being interviewed. He had drawn pictures of a man's genitals in nursery and the interviewers were told in their scenario that his nanny had recently been fired from her job, which served as a 'red herring' as he had in fact been abused by his father. One of the female scenarios consisted of a 15-year-old girl who had run away from a care home that has secretly been involved in child prostitution. At the end of the day the interviewers were again provided with group feedback. This design is typical of JIIT within this location.

211

Actors were not given specific instructions about various interview phases or how to respond to specific question types, but had informal discussions with trainers if they had any questions about their character or role.

7.2.3   *Procedure for coding interviews.* General procedure. Each interview was transcribed verbatim by the author of this thesis to Microsoft Word. Transcribing software/equipment was unavailable to the author and so this process instead involved playing the interview DVDs on VLC (VideoLAN Client) media player on a PC and manually controlling the pausing and rewinding while simultaneously transcribing each interview word for word. On average, each of the 44 interviews took approximately five hours to transcribe. Following transcription each interview was checked for accuracy and completeness by playing the interview DVD from start to finish while simultaneously reading through the transcript.

Coding the pre-substantive phase. First, the quality of the introductory phase of the interview was assessed using an 'inclusion checklist' developed by the author of this thesis to account for all of the protocols that are expected to be followed in this phase according to Scottish Executive Guidelines (2003, see Appendix 1A). Twelve ground rules: 'Listen', 'Knowledge', 'I don't know', 'I don't remember', 'Don't guess', 'Don't know demonstration', 'I don't understand', 'Don't understand demonstration', 'Correct me', 'Correct me demonstration' 'Repeated questions' and 'Tell the truth' were scored as either present or absent according to whether the interviewer used the rule.

Narrative Elaboration Training. Next, a practice interview was scored as either present or absent. To be scored as present, the interviewer must have attempted to elicit information about an event unrelated to the scenario (e.g. "Tell me about this morning at school"). In order to comply with the Scottish Executive Guidelines (see also Chapter 4), the practice interview should use open prompts in order to provide 'narrative elaboration practice' for the interviewee. Therefore, the practice interviews were further categorised if they followed this rule. Appropriately conducted practice interviews were scored 'open' and consisted of the interviewer asking open questions to elicit a narrative account (e.g. "Tell me what happened from when you arrived at school this morning until you left to come here"). Inappropriately conducted practice interviews were scored as 'closed' and consisted of the interviewer questioning the actor about the event using focussed questions rather than open prompts (e.g., when discussing events from earlier in the day, "What time did you get to school?", "Who do you sit with?", "Do you like your teacher?").

Coding the substantive phase of the interview. Next, each prompt in the substantive phase of the interviews (i.e. when discussion of the allegation or the reason for interview commenced) was categorised into one of six following categories originally developed by Lamb et al. (1996): Invitations, directives, option-posing utterances, suggestive utterances, summaries, or non-substantive utterances, (full descriptions of each category are provided in Chapter 4.2).Facilitators (utterances such as

"uh-huh" or repeating the child's last few words to encourage the child to keep talking) were coded but not analysed as a separate question category. Instead, facilitators were counted as additional questions in the same category as the question they followed. For example: ([Interviewer]: "What happened next?" [Actor]: "We went inside", [INTERVIEWER]: "You went inside uh-uh"). The latter would be classed as an additional invitation and any words and/or details that follow would be attributed to the previous question type (Lamb et al., 1996a). Of note, while some questions structurally may look like they belong to one category, the information the question is requesting takes primacy when coding. For example, if an interviewer asked "Do you have a name for the place he touched you?" the question looks like an option-posing question (i.e. it can be answered with yes or no) but the interviewer is actually requesting a name for the location, thus it is coded as a directive question.

First use of focussed prompts. As it is best practice to delay the use of focussed prompts until as late as possible in the interview(e.g. Lamb et al., 2007a, Lamb et al., 2008; Scottish Executive, 2003, 2011), the number and percentage of questions asked before the first directive question and before the first option-posing utterance were calculated.

Closure. Finally, the quality of the closure phase of the interview was assessed using an 'inclusion checklist' developed by the author of this thesis to account for all of the protocols that are expected to be followed in this phase according to Scottish Executive Guidelines (2003, see

Appendix 1a). Six closure principles: 'Summarise', 'Interviewee questions', 'Next', 'Contact', 'Thanks' and 'Neutral closure' were scored as either present or absent in each interview.

Coding the responses of actors. Every actor response in the substantive phase of the interview was scored as either 'informative' or 'uninformative' depending whether they provided information in their reply or not (see Chapter 4.2 for a full description of informative versus uninformative responses from Lamb et al., 1996). Informative responses included:

1.   Responsive. Related to the content of the interviewer's previous utterance.

2.   Responsive action. Gestures not accompanied by a verbal response.

3.   Unresponsive. Not related to the interviewer's previous utterance but do provide incident-related information.

Uninformative responses included:

1.   Providing no answer.

2.   Requesting clarification.

3.   Unclear responses.

4.   Digressions to non-substantive topics.

Three additional categories created for this study were:

5.   Saying 'I don't know or 'I don't remember'.

215

6.   Questions.

7.   Resistance, denial or unwillingness to provide information.

Coding details provided by actors. Following the NICHD coding manual, each actor response was also scored for the number of substantive (allegation-related) details that it contained. A detail is the smallest unit for analysing information provided in the interviewee's account and consists of any information pertaining to the incident that is conveyed by the interviewee. Details include naming, identifying or describing individual(s), object(s), event(s), place(s), action(s), emotion(s), thought(s), sensation(s) that are part of the alleged incident as well as any of their features (e.g. appearance, location, time, duration, temporal order, sound, smell, and texture). Details are counted in specific statements that express recollections of personal episodic memories of alleged incidents (occurring at a specific time and a specific location), as well as generic statements that refer in general to something that happened during a single incident or summarise more than one incident with the same suspect. Only new details per utterance are counted (i.e. repeated details are not counted again). Non-verbal cues (points, nods) are counted as details but not as words. If it is unclear if a detail is relevant or not, it is still counted (see NICHD codebook).

In order to examine whether actors were sensitive to the different types of questions posed by the trainee interviewer (i.e. responding authentically to that of a child, longer and more detailed responses to open questions than focussed questions), the question type that elicited

the first detail from actors was recorded for each interview. In addition, as it is best practice to delay the use of focussed prompts until as late as possible in the interview (e.g. Lamb et al., 2007a, Lamb et al., 2008; Scottish Executive, 2003, 2011), the number and percentage of details elicited from actors before interviewers moved on to directive and option-posing questions was calculated. Actor responses were also scored for the number of words they contained, with contracted words (e.g., "don't", "can't") scored as two words.

7.2.4 Coder training and inter-rater reliability. Both the lead coder (the author of this thesis) and reliability coder were trained to code forensic field interviews at a week-long intensive NICHD coding workshop held at Cambridge University. Extensive coding was completed for five full days under the supervision of Dr Yael Orbach (Research Scientist, National Institute of Health & Human Development) via a live video link from Washington DC. This coding system has been used to code field interviews with children in many publications (reviewed in Chapter 4.2). The lead coder coded all transcripts and the reliability coder coded a subset (20%) of randomly-selected transcripts, with all disagreements resolved through discussion until an agreement was reached. Reliability between raters in coding question types was high (K=.78), and reliability between raters in coding actor responses (K=.69) and actor details (K=.74) was good.

7.2.5 Outline of analyses The two forces differed in the number of questions posed across interviews ($M_{UtterancesForce1}$=47.05, *SEM*=3.69, $M_{UtterancesForce2}$=27.05, *SEM*=3.78; *t*(40)=3.78; *p*<.01, *d*=1.19) and the actors in the two forces differed in the number of words provided to the trainee ($M_{WordsForce1}$=490.14, *SEM*=43.52, $M_{WordsForce2}$=206.75, *SEM*=30.71; *t*(40)=5.22; *p*<.01, *d*=1.67). Therefore analyses are conducted on proportional data rather than absolute numbers where analysis of the latter would represent a confound (e.g. falsely attributing meaning to a higher order interaction between question type and training force). Analyses are conducted on interviewer behaviour followed by actor behaviour in the order in which they occur during the interview. Chi-square tests are conducted to test for associations between training force and the use of introductory ground rules, narrative elaboration training and closure principles at each interview. Where use of rules does not differ according to training force, binomial tests collapsed across training force are conducted to test whether the presence of rules across training sessions is greater than would be expected by chance (i.e. 0.50). Mixed design ANOVAs are conducted in order to test whether various interviewer and actor behaviours discussed differ according to the force at which the training was administered and the question type posed. Overall main effects and/or interactions are interpreted with t tests and/or Bonferroni corrected multiple comparisons of within-subjects' effects (i.e. actor responses to question type and/or interviewer usage of different question types).

7.3 Results

7.3.1 Interviewer behaviour (pre-substantive phase of interview)

When comparing force one to force two, chi-square tests revealed a significant association between the force where training was administered and the presence of recommended rules 'Don't know' (Observed usage$_{Force1}$ = 23%, Observed usage$_{Force2}$ = 60%, $X^2(1)$=6.04; p=.014, Odds ratio =5.10), 'Don't understand' (Observed usage$_{Force1}$ = 27%, Observed usage$_{Force2}$ = 65%, $X^2(1)$=6.02; p=.014, Odds ratio =4.95), 'Truth' (Observed usage$_{Force1}$ = 27%, Observed usage$_{Force2}$ = 65%, $X^2(1)$=6.02; p=.014, Odds ratio =4.95) and 'Narrative Elaboration Training' (Observed usage$_{Force1}$ = 55%, Observed usage$_{Force2}$ = 20%, $X^2(1)$=5.30; p=.021, Odds ratio=4.80). The odds ratio demonstrates that the odds of using 'don't know', 'don't understand' and 'truth' rules were, respectively, 5.10, 4.95 and 4.95 times greater in force two than force one. By contrast, the odds of using Narrative Elaboration Training were 4.80 times higher in force one than in force two. None of the expected frequencies within this analysis were less than five (see Field, 2009 for discussion). Use of all other rules at this phase of training did not differ according to force (all X^2<2.64 all p>.10). Although use of 'Don't know demonstration' (Observed usage$_{Force1}$ = 5%, Observed usage$_{Force2}$ = 35%, $X^2(1)$=6.30; p=.012) and 'Don't remember' (Observed usage$_{Force1}$ = 9%, Observed usage$_{Force2}$ = 35%, $X^2(1)$=4.18; p=.041) did differ according to the force where the training took place, expected frequencies within this analysis fell below the minimal recommended number of five, therefore this data should not be analysed separately by force (Field, 2009).

219

Separate binomial tests were then conducted (collapsed across forces) on frequency of usage of the remaining ten rules that did not differ reliably according to force (see Table 7.1). These analyses revealed that with the exception of the rule 'listen' ($p=.88$), all remaining rules were used significantly less than would be expected by chance (i.e. 0.50, all $p<.01$).

Table 7.1: Presence of remaining ten rules at pre-substantive phase, collapsed across training force

Rule	Mean proportion usage (SEM)	Binomial test result
Listen	0.48 (.08)	$p=.88$
Knowledge	0.24 (.07)	$p<.01$
Don't know demonstration	0.19 (.06)	$p<.001$
Don't understand demonstration	.02 (.02)	$p<.001$
Correct me	.19 (.06)	$p<.001$
Correct me demonstration	.07 (.04)	$p<.001$
Repeated questions	.12 (.05)	$p<.001$
Don't remember	.21 (.06)	$p<.001$
Don't guess	.17 (.06)	$p<.001$
NET Open	.19 (.06)	$p<.001$

7.3.2 Interviewer behaviour (Substantive phase of interview).

A 2x6 mixed-design ANOVA was then conducted on the outcome variable proportion of total utterances by interviewer, with the within subjects' factor *question type* (invitation, directive, option-posing, suggestive, non-substantive, summary) and the between-subjects factor *training force* (force one, force two). Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). These analyses revealed a main effect of

question type ($F(2.75,110.07)=78.63$; $p<.001$, $np^2=.66$) that was not

qualified by an interaction with *training force* ($F(2.75,110.07)=1.42$; $p=.24$,

$np^2=.03$). Post-hoc analyses collapsed across *training force*(corrected for

multiple comparisons using Bonferroni testing) demonstrated that all six

question types were used to a different extent by interviewers throughout

the course of their interviews (all $p<.013$, see Table 7.2), except for

invitations and directives, and non-substantive questions and summaries

which were used to the same extent as one another across forces and

interviews (all Bonferroni corrected $p=1.00$).

Table 7.2: Use of different question types as a proportion of total interviewer utterances during the substantive phase of the interview, collapsed across training forces (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	.34 (.02)	.30	.39
Directive	.32 (.02)	.28	.35
Option-posing	.22 (.02)	.18	.25
Suggestive	.11 (.02)	.08	.14
Non-substantive	.01 (<.01)	.001	.01
Summary	.01 (<.01)	.01	.02

Next, the first use of focussed prompts (directives and option-posing

questions) were compared across forces. A 2x2 mixed-design ANOVA,

with the between subjects factor *training force* (force one, force two) and

the within-subjects factor *focussed question type* (option-posing,

directive) revealed a main effect of *focussed question type* ($F(1,40)=6.15$;

$p=.017$, $np^2=.13$) that was not qualified by an interaction with *training

force* ($F(1,40)=0.32$; $p=.58$). There was no main effect of *training force*

($F(1,40)=.14$; $p=.71$). Paired samples t tests revealed that, across forces, trainees introduced directives earlier in the conversation ($M_{utterances}=4.21$, $SEM=.46$) than they introduced option-posing questions ($M_{utterances}=6.19$, $SEM=.72$, $t(41)=2.48$; $p=.018$, $d=0.38$).

The proportion of details elicited from the actor before the use of focussed prompts was then examined. A 2x2 mixed-design ANOVA, with the between subjects factor *training force* (force one, force two) and the within-subjects factor *focussed question type* (option-posing, directive) revealed no main effect of *focussed question type* ($F(1,40)=1.42$; $p=.24$) or interaction between *focussed question type* and *training force* ($F(1,40)=.54$; $p=.47$). There was a main effect of *training force* ($F(1,40)=11.09$; $p<.01$, $np^2=.22$). Independent samples t tests collapsed across focussed question type revealed that force one elicited a greater proportion of details from actors before introducing focussed questions ($M_{proportiondetails} = 23.86$, $SEM=4.21$), than did force two ($M_{proportiondetails} = 8.13$, $SEM=1.76$, $t(28.01)=3.45$; $p<.01$, $d=1.32$).

7.3.3 Use of the closure principles. When comparing force one to force two, chi-square tests revealed a significant association between the force where training was administered and the presence of the recommended closure rules 'Actor questions' (Observed usage$_{Force1}$ = 13.6%, Observed usage$_{Force2}$ = 55.0%, $X^2(1)=8.07$; $p<.01$, Odds ratio =7.74), 'Contact' (Observed usage$_{Force1}$ = 4.5%, Observed usage$_{Force2}$ = 60%, $X^2(1)=15.07$; $p<.001$, Odds ratio =31.50) and 'Thanks' (Observed usage$_{Force1}$ = 27.3%, Observed usage$_{Force2}$ = 85.0%, $X^2(1)=14.09$; $p<.001$, Odds ratio =15.11). Here, the odds ratio demonstrates that the odds of

222

asking the actor questions were 7.74 times greater in force two than force one, the odds of using the contact rule were 31.50 times greater in force two than force one and the odds of using the thanks rule were 15.11 times greater in force two than force one. Although use of 'Next' (Observed usage$_{Force1}$ = 4.5%, Observed usage$_{Force2}$ = 30%, $X^2(1)$=4.89; p=.027) and 'Summary' (Observed usage$_{Force1}$ = 36.4%, Observed usage$_{Force2}$ = 5%, $X^2(1)$=6.12; p=.013) rules differed according to the force where the training took place, expected frequencies within this analysis fell below the minimal recommended number of five, therefore this data should not be analysed separately by force.

Separate binomial tests were then conducted (collapsed across forces) on frequency of usage of the remaining three rules that did not differ reliably according to force. These analyses all remaining rules were used significantly less than would be expected by chance by trainees (see Table 7.3).

Table 7.3: Presence of remaining three closure rules, collapsed across training force

Rule	Mean proportion usage (SEM)	Binomial test result
Next	.17 (.06)	p<.001
Neutral	.31 (.07)	p=.021
Summary	.21 (.06)	p<.001

7.3.4 Actor behaviour. A 2x6 mixed-design ANOVA was conducted on the outcome variable proportion of total details by actor, with the within subjects' factor question type (invitation, directive, option-posing,

223

suggestive, non-substantive, summary) and the between-subjects factor training force (force one, force two). Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test (p<.001). These analyses revealed a main effect of question type ($F(2.78,111.33)=35.44$; $p<.001$, $np^2=.47$) that was not qualified by an interaction with training force ($F(2.78,111.33)=1.36$; $p=.26$, $np^2=.03$). Post-hoc analyses collapsed across training force (corrected for multiple comparisons using Bonferroni testing) demonstrated that all six question types elicited a different amount of detail from the actor over the course of the interview (all p<.01, see Table 7.4), except for invitations and directives, directives and option-posing questions, option-posing questions and suggestive questions, and non-substantive questions and summaries, which all elicited an equivalent amount of detail from the actors across forces (all Bonferroni corrected $p>.064$).

Table 7.4: Proportion of total details elicited from the actor during the substantive phase of the interview broken down by question type and collapsed across training forces (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	.41 (.04)	.34	.48
Directive	.28 (.03)	.23	.33
Option-posing	.16 (.03)	.11	.22
Suggestive	.12 (.03)	.07	.17
Non-substantive	.01 (<.01)	.002	.02
Summary	.02 (<.01)	.01	.04

Descriptive statistics suggested that across four question types (invitations, directives, option-posing, suggestive) the first detail was elicited from invitations 72.7% of the time in force one (Directives=4.5%,

224

Option-posing=9.1%, Suggestive=13.6%) and 50% of the time in force

two (Directives=10%, Option-posing=25%, Suggestive=10%). In order to

verify if invitations elicited the first detail more often than the other

question types, data were recoded to 1 if an invitation elicited the first

detail and 0 if another question type elicited the first detail. Chi-square

tests comparing force one to force two revealed no association between

the force at where the training was administered and the tendency for an

invitation to elicit the first detail (Observed tendency$_{Force\ 1}$ = 72.7%,

Observed tendency$_{Force\ 2}$ = 52.6%, $X^2(1)$=1.78; p=.18). A separate

binomial test collapsed across training force revealed that an invitation

did not elicit the first detail at levels significantly greater than chance

(p=.12).

Next, a 2x6 mixed-design ANOVA was conducted on the outcome

variable proportion of total words elicited by the actor, with the within-

subjects factor *question type* (invitation, directive, option-posing,

suggestive, non-substantive, summary) and the between-subjects factor

training force (force one, force two). Greenhouse-Geisser corrections

were used as the assumption of sphericity was violated according to

Mauchly's test (p<.001). This analysis revealed a main effect of *question

type* ($F(2.33, 93.20)$=73.58; p<.001, np^2=.65) that was not qualified by an

interaction with *training force* ($F(2.33, 93.20)$=2.89; p=.052, np^2=.07).

Post hoc analyses collapsed across training force (corrected for multiple

comparisons using Bonferroni testing) revealed that all combinations of

questions types elicited a different proportion of total words from the actor

(all p<.01, see Table 7.5), except for invitations and directives, option-

posing questions and suggestive questions, and non-substantive

questions and summaries (all Bonferroni-corrected p=1.00).

Table 7.5: Proportion of total words elicited from the actor during the substantive phase of the interview broken down by question type and collapsed across training forces (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	.42 (.03)	.37	.48
Directive	.34 (.03)	.29	.39
Option-posing	.11 (.02)	.08	.14
Suggestive	.10 (.02)	.06	.14
Non-substantive	.01 (<.01)	.002	.02
Summary	.01 (<.01)	.004	.02

Next, to test the length of actors' responses, a 2x4 mixed-design ANOVA

was run on the outcome variable average length of response, with the

within subjects' factor *question type* (invitation, directive, option-posing,

suggestive) and the between subjects' factor *training force* (force one,

force two). Participants who were not asked one of the question types

were excluded from analysis (N=7 participants in force two) and

summaries and non-substantive prompts were excluded from the overall

analysis as only twelve and seven interviewers used these question types

respectively. Greenhouse-Geisser corrections are used as Mauchly's test

revealed that sphericity was violated (p<.001). These analyses revealed a

main effect of *question type* (F(1.40,55.91)=6.15; p<.01, np^2=.13) that

was not qualified by an interaction with *training force* (F(1.40,55.91)=.26;

p=.69). There was no main effect of *training force* (F(1,40)=2.11; p=.15).

Post-hoc analyses collapsed across *training force* (corrected for multiple

comparisons using Bonferroni testing) demonstrated that invitations

226

elicited a longer response from the actor, on average, than did option-posing questions ($p<.001$) and that directive questions elicited a longer response from the actor, on average, than did option-posing questions ($p<.001$). Average length of response did not differ between any other combination of question type (all other Bonferroni-corrected $p>.21$, see Figure 7.1).
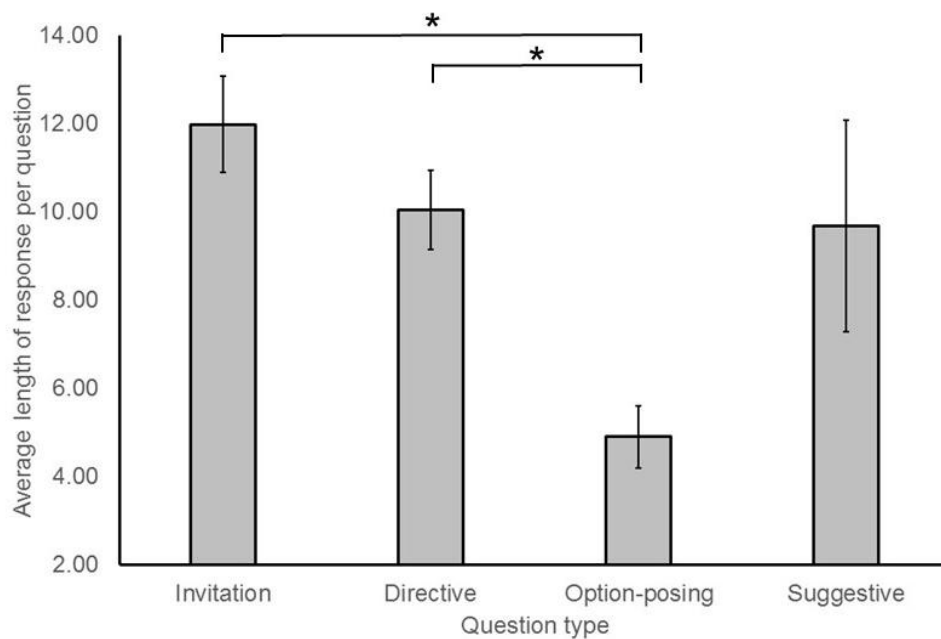
Figure 7.1 : Average length of response per question type. Error bars show +-1SEM

Average number of details yielded per question type

To test the richness of actors' responses, a further 2x6 mixed-design ANOVA was run on the outcome variable average number of details per response, with the within subjects' factor *question type* (invitation, directive, option-posing, suggestive, non-substantive, summary) and the between subjects' factor *training force* (force one, force two). Greenhouse-Geisser corrections are used as Mauchly's test revealed that sphericity was violated ($p<.001$). These analyses revealed a main effect

of *question type* ($F_{(2.38, 95.13)}=5.24$; $p<.01$, $np^2=.12$) and a main effect of *training force* ($F_{(1,40)}=29.66$; $p<.001$, $np^2=.43$). There was no interaction between *question type* and *training force* ($F_{(2.38, 95.13)}=2.05$; $p=.13$, $np^2=.05$). Independent t tests revealed that, across question types, force one generally elicited more details per response from actors (*M*=3.29, *SEM*=.32) than did force two (*M*=1.12, *SEM*=.22, $t_{(36.87)}=5.55$; $p<.001$, *d*=1.81). Post-hoc analyses collapsed across *training force* (corrected for multiple comparisons using Bonferroni testing) demonstrated that invitations elicited more details from actors than option-posing questions and non-substantive questions (both $p<.001$). Directive questions elicited more details from actors than option-posing questions and non-substantive questions (both $p<.021$). No other significant differences were found from pairwise comparisons of different question types on richness of actor response (i.e. number of details per response; all other Bonferroni-corrected $p>.055$, see Figure 7.2).
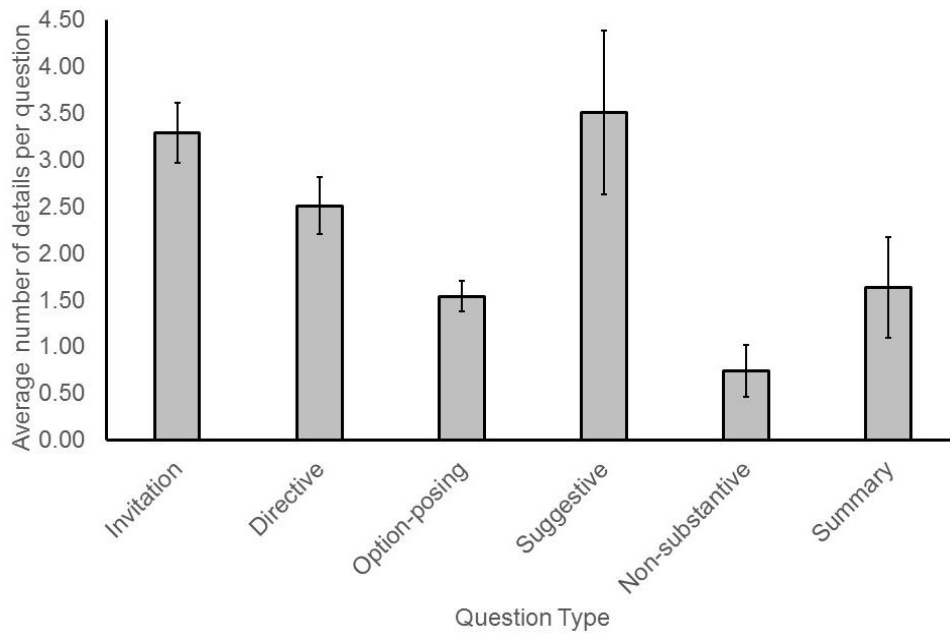
Figure 7.2: Average number of details provided per question type. Error bars show +- 1SEM

Finally, a 2x4 mixed-design ANOVA was conducted on the proportion of informative responses provided by the actor, with the within subjects factors *question type* (invitation, directive, option-posing, suggestive) and the between subjects factor *training force* (force one, force two). Greenhouse-Geisser values are reported as sphericity was violated according to Mauchly's test (p=.024). This analysis revealed a main effect of *question type* (F(2.32, 76.41)=6.81; p<.01, np^2=.17) that was not qualified by an interaction with *training force* (F(2.32, 76.41)=2.43; p=.09). There was no main effect of *training force* (F(1,33)=1.83; p=.19). Post hoc comparisons collapsed across training force (Bonferroni corrected p values) demonstrated that all combinations of question type elicited a different proportion of informative responses from the actor (all significant comparisons: Bonferroni-corrected p<.03), except for directives and option-posing questions, directives and suggestive questions and option-

229

posing questions and suggestive questions which elicited an equivalent proportion of informative responses from the actor (all non-significant comparisons: Bonferroni-corrected p =1.00, see Figure 7.3).
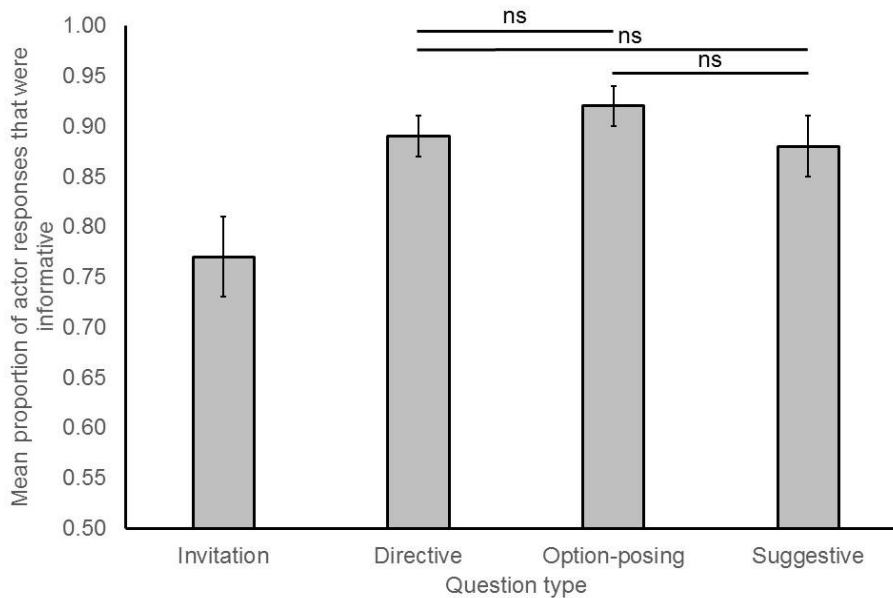
Figure 7.3: Average proportion of informative responses provided per question type. Error bars show +- 1SEM. All comparisons are significant unless indicated above (three comparisons).

7.3.5 Additional general observations. A few unexpected observations were made during the study (and hence were not predicted). Force one interviewers asked the transition question(the question that commences the substantive phase by asking the actor about the reason they are being interviewed) straight after their introductions but on obtaining a disclosure they reverted back to building rapport. Some interviewers asked further questions related to the allegation after closing the interview with neutral topics, and some offered the interviewee an opportunity to 'write down' a word about what happened when faced with resistance. Interviewees, in some cases, were

230

asked to tell or show what happened on their own body and sometimes were asked to draw or close their eyes and imagine what happened. Some trainees also asked the actor playing the child to use items in the room (e.g. a pen and book) to represent the people and objects they were describing.

Actors also exhibited some concerning behaviours e.g., some of the force two scenarios instructing actors to be unwilling to disclose information resulted in one interviewer eliciting no details even though open prompts were used and in actors swearing at interviewers. Scribing was also an issue that negatively impacted on the quality of interviews conducted by force two. Frequent interruptions and a quick pace led to some information being lost (e.g. when scribes asked for information to be repeated, often nobody could remember previous questions or answers). Unexpectedly, actors frequently refused to provide any information when interviewers attempted narrative elaboration training (e.g. saying 'nothing' or 'can't remember' in response to open prompts such as 'tell me what you did yesterday').

7.4   Discussion

Here, the first systematic analysis of both interviewer and actor behaviour during JIIT revealed several findings that speak to the quality of investigative interviewer training within Scotland. Firstly, use of three of the introductory rules and narrative elaboration training differed according to the force in which the training was administered. Consistent with predictions that force two would perform better than force one in light of the intensity of their training, interviewers in force two were 5.10 times

more likely to relay the importance to the actor of saying when they don't know something, and were 4.95 times more likely than force one interviewers to reiterate the importance of telling the interviewer when they don't understand the question and to only tell the truth when responding. By contrast, and contrary to predictions, interviewers in force one were 4.80 times more likely than interviewers in force two to use narrative elaboration training with the actor. Collectively, these findings suggest that a minority of rules were used differently in training according to the force at which training was administered. However, a majority of rules (ten rules) did not differ reliably according to the force where training took place. Moreover, analysis of this data suggests that with the exception of one rule 'listen' (an explanation that it is the interviewer's job to listen and the child's to do the talking), all of the remaining rules were used less than would be expected by chance across both forces. Collectively, the first examination of use of best-practice rules at the pre-substantive phase of an interview suggests that trainees are relatively poor at adhering to best practice, even while attending training.

Secondly, training forces were consistent in the proportion of their entire interview they devoted to asking different types of questions to the actor, using option-posing questions 22% of the time, directives 32% of the time and invitations 34% of the time. While use of invitations reflects best practice guidelines (e.g. Lamb et al., 2007a, Lamb et al., 2008; Scottish Executive, 2003, 2011), follow-up analyses of these data demonstrate that these open prompts were used to the same extent as focussed prompts such as directives which, in turn, were used to a

greater extent than option-posing questions. Moreover, despite empirical research on the risks of using suggestive questions during eyewitness interviews (e.g. Bruck, Ceci, Francoeur& Barr, 1995b; Leichtman&Ceci, 1995; Thompson, Clarke-Stewart & Lepore, 1997), these question types still constituted approximately 1 in 10 of all questions posed to the actor portraying a child during training. Contrary to predictions, force two gathered fewer details from the actor before introducing focussed questions in comparison to force one. When examining the number of utterances more generally during the conversation, both forces introduced directives earlier in the conversation than they introduced option-posing questions. Collectively, trainees on JIIT courses appear to understand the merits of using invitations during an interview but also contradict best practice guidelines by using directives to the same extent, introducing them relatively early into the conversation and, among force two, gathering relatively little detail from the actor before posing a directive question.

Finally, when examining the use of closure principles, force two was more likely to use three of the closure rules than force one: they were more likely to ask the actor questions at the end (OR=7.74), to use the contact rule (OR=31.50) and were more likely to thank the actor (OR=15.11) than force one. Of the remaining three rules, both forces were less likely to use these closure rules than would be expected by chance. Collectively, with the exception of three rules used by force two, trainees were poor at using the closure rules during training.

As highlighted previously, the current study provides one of the first systematic examinations of the behaviour of adult actors who play the role of a child during JIIT. The data demonstrate that across both forces, actors reinforced best practice by talking more in response to invitations than option-posing or suggestive questions, however, when examining both proportion of total details and proportion of total words provided, actors responded equally to invitations and directives. Moreover, when examining the first detail provided by the actor across training forces, invitations were no more likely to elicit the first detail from the actor than other question types were. On average, actor responses were longer to invitations than option-posing questions (reinforcing best practice) but did not differ from length of responses to directive questions which, in turn, were longer than responses to option-posing questions, suggesting inconsistency in the actor's adherence to best practice guidelines. Although force one actors provided more details per response than did force two actors (i.e. higher yield), actors across forces provided richer responses (i.e. more details) in response to invitations than option-posing questions however, of concern, provided equally rich responses in response to invitations, directives and suggestive questions. Indeed, although the majority of responses from the actor to each question type were informative across training forces (ranging from 77% to 92%), actor responses to invitations were statistically more likely to be uninformative (i.e. the actor did not provide any information in response to the question) than were responses to focussed prompts or suggestive questions. Collectively, these findings raise concerns about the authenticity of adult

234

actors who play the role of a child, even when JIIT actors are provided from separate acting companies.

There are several implications of these findings and potential explanations for these patterns of results. Firstly, even during training when interviewers are fully immersed in the intensive learning process and being supported by and receiving feedback from their trainers, trainees are not practicing all aspects of conducting interviews. Interviewers' lack of adherence to interviewing guidelines during the introductory phase of the interviews is concerning because research has shown that delivering ground rules and practicing narrative elaboration training are important features of forensic interviews with children (see Chapter 4) and this is the only practice interviewers have before conducting field interviews. It may be that interviewers believe that knowledge alone of the introductory elements of the interview without practice is sufficient during training and that when they are conducting real interviews with children they would be sure to include ground rules and NET. However, field studies (La Rooy, Earheart & Nicol, 2013; La Rooy, Nicol, Halley & Lamb, 2012) have shown that these aspects are routinely ignored in the field, therefore, not practicing during JIIT courses may give the impression these aspects are not as important as the substantive phase, in turn diminishing their use in the field. Indeed, interviewers have expressed beliefs that NET is not useful, e.g. La Rooy et al. (2011) found in their survey that 52% of police interviewers of children indicated that they believed the practice interview to be either not

very effective or not at all effective and 87% admitted that they never or rarely conducted one.

This may stem from a lack of understanding of what a properly conducted NET practice actually is, of the 38% of interviewers that attempted to conduct a practice narrative about a personally experienced event, of these, only half of the practice interviews were conducted in the correct manner using open-prompts. Interviewers did however attempt to establish rapport by asking about things the interviewee liked to do but they did so using focussed questions. This suggests that interviewers do not understand the purpose of NET and are confusing the concept with establishing general rapport and getting the interviewee talking, compounded with not engaging with NET during training this may further reinforce the idea that it is not an important component of child interviewing.

Encouragingly, with regards to interviewers' questioning behaviour in the substantive portion of the interviews, the proportions of invitations, directives and option-posing questions used by these trainees all fall in the range of best practice field interviews conducted using the NICHD protocol (Invitations range from 30-48%, directives 26%-44%, option-posing 18%-26%, Cyr et al., 2006; Lamb et al., 2006; Orbach et al., 2000; Sternberg et al., 2001). Further, the proportion of open prompts used by trainees was greater than the proportion of focussed (option-posing and suggestive) questions. On one hand this is unsurprising given that interviewers were using an example protocol based on the NICHD protocol and that they had just undergone training in how to conduct high

quality interviews. However, it is surprising that interviewers in Scotland are conducting such high quality interviews during training given the fact that the finding that the quality of field interviews in Scotland is low. It has been shown that the quality of interviews is highest during training and that over time the gains made during training dissipate (Lamb et al., 2002a) and therefore low quality interviews in the field could reflect increasing length of time since undertaking interviewers training and conducting field interviews. Alternatively, during training interviewers know that they are being observed and may make more of an effort to 'stick to the rules' than when they are in the field. Unfortunately, interviewers used invitations and directives to the same extent as one another therefore, while interviewers' training has conveyed the importance of relying on recall prompts as opposed to recognition prompts, it has not managed to underscore the superiority of open prompts.

Some behaviours were observed in the interviews that are not evidence based. Interviewers at force one asked the transition question straight after introducing themselves and after gaining a disclosure they explained that they wanted to get to know the interviewee better before discussing the allegation and reverted back to completing the rapport phase. While adult actors responded to the transition question with an allegation this may not be the case with children. The guidelines recommend rapport and practice narratives before transitioning into the substantive phase in order to familiarise the interviewee with the interviewer, to put them at ease and importantly to give them an

237

opportunity practicing responding with narratives. Children may not be ready to discuss the allegation at this early stage having just met the interviewer and even if they do respond with a disclosure early on they have not been familiarised with open promoting and may not yet be able to respond with narratives. Further, it may be detrimental to children's motivation that after disclosing an instance(s) of abuse that the interviewer ignores this and they are diverted back to rapport. Some interviewers offered interviewees an opportunity to "write down" a word or to draw aspects of things they were talking about, some were asked to close their eyes and imagine and some interviewers asked the interviewee to use items in the room (e.g. a pen and a book) to represent people and objects they were describing. These techniques are not recommended in the guidelines and should be avoided because information elicited in such ways may be difficult to interpret and distract the child from their task of narrating events from their lives.

Scribing negatively impacted both interviewers' questioning behaviour and adult actors responding. During the substantive phase, interviewers at force one used almost double the amount of utterances used by interviewers at force two and force one actors provided three and a half times more words and more than three times more details overall than force two actors. This result is likely to be because force one had access to VRI equipment and recording was mandatory in this jurisdiction. No such measures were in place at the time in force two. Instead, the interviewers had to record the interview verbatim by hand, which slowed the pace at which questions could be asked and answered.

Therefore, the lack of VRI facilities disadvantaged force two interviewers in that they did not have the same opportunity to practice asking and maintaining the use of open prompts as force one interviewers. La Rooy et al. (2011) have reported previously interviewers' concerns about having to wait for scribes to catch up, though not part of the analysis, it was noted that during some interviews information was lost when scribes requested repetition of questions and answers, neither the lead interviewer nor the interviewee could remember them.

Adult actors exhibited some behaviours that were potentially counterproductive to best practice. Actors refused to respond to interviewers' attempts to initiate NET, when asked for example about things they had done earlier they often replied with "nothing" or "don't know". As the actors responded to closed questions during rapport, instead of purposefully obstructing interviewers, it is likely that actors, having received no training or instruction to play these roles, were unaware that the interviewers were attempting to elicit narratives about events from them and did not know how to respond appropriately. Therefore, unknowingly actors may have diminished the interviewers' attempt to conduct NET and hurried them towards the substantive phase.

With regards to question type, both invitations and directive questions naturally elicited longer and more detailed average responses from actors than option-posing questions but no other significant differences were found. This speaks to the fact that actors are not sensitive to question types because we would want actors to provide longer and more detailed responses to invitations than directives to show

239

superiority of open prompts. Further, the best practice invitations elicited the smallest proportion of informative responses from actors (77%) of all the question types, vice versa actors provided a significantly greater proportion of uninformative responses to invitations than to all other question types. This means actors were more likely to be 'difficult' or resistant when asked best practice open questions and interviewers were more likely to receive information when they asked directives, option-posing and suggestive questions. Therefore, actors may have been unknowingly reinforcing the use of these questions and undermined the superiority of the open prompts, decreasing the likelihood of interviewers maintaining their use of open questioning. It is of the utmost importance then that actors receive appropriate training that includes learning about different question types and about how real children respond to them in order to undertake this role.

Interestingly, when the average number of words and details yielded per question type were examined there was a main effect of training force for details, but not for words. Force one actors provided more details per response to every type of question than did force two actors but actors at both forces gave responses similar in length, that is, force 2 actors provided the same average length of responses but they were less detailed. Actors are therefore providing a different experience at different locations because force one interviewers' questions were rewarded to a greater extent. It is likely that the scenarios used at the different forces played a role in the inferior responding of force two actors as some of their scenarios provided instruction to withhold information.

Finally, the frequency of use of some interview components was variable across training sites, despite the fact that all JIIT courses are expected to follow a national standard and they rely on the same interview guidelines. It is likely the differences in interviewers' performance between forces found were because of lack of standardisation of aspects of the course. Standardisation of the learning materials (scenarios), learning outcomes, the expectations of trainee interviewers and the facilities and resources available, such as VRI, on JIIT courses will improve the quality and standardise JIIT across the country.

Given that the current study demonstrates limitations in the use of standard issue hired adult actors who play the role of a child alleging abuse during JIIT courses in Scotland, the next study aims to examine a major concern relating to this issue. Namely, a study was conducted to see whether an actor with knowledge of evidence-based practice in investigative interviews would result in responses that reflected and reinforced the benefits of open prompts, by providing longer and more detailed responses to these questions in contrast to responses to cued and recognition prompts.

Chapter 8: The accuracy of Scottish investigative interviewers' handwritten verbatim interview records compared with electronically recorded versions

8.1 Study rationale and aims

The findings of Study 1 raised concerns about the negative effects that 'scribing' (keeping a real-time handwritten verbatim interview record) imposes on the quality of trainee's JIIT role-play interviews and the accuracy of their interview records. Interviewers that had to rely on a fellow-trainee to scribe their interviews asked 43% fewer total questions compared to their peers that had access to VR (visual recording) facilities during training. Moreover, information provided by the interviewee was lost from the interview record as the 'scribe' was unable to keep up with the pace of the dialogue within the dyad. Indeed, when the scribe asked for information to be repeated, often the lead interviewer or actor could not remember their questions or answers.

This is not the first occasion that such concerns have been highlighted. For example, La Rooy et al. (2011) surveyed child interviewers in Scotland and found that the majority of respondents expressed concerns that scribes failed to write down details elicited from children during field interviews. Moreover, these concerns have been validated in both field and experimental research. Lamb et al. (2000) found that 57% of the questions asked by interviewers and 25% of the details elicited from children in investigative interviews were missing from the handwritten versions of the interviews when they compared them to the audio taped transcribed versions. An experimental study where

242

conditions were tailored in the interviewers' favour (a live interview script was read out at 2.2 words per second), similarly showed that interviewers still failed to record 39% of the abuse-related details provided by adult actors.

In addition to the absence of interviewers' questions and interviewee' details in scribes' accounts, other problems have been identified with handwritten interview records. The scribes in Lamb et al. (2000) distorted the questions that had really been asked in their handwritten account in a manner that made them look like a higher quality question had been asked (that is focussed questions were recorded as being more open). Cauchi and Powell (2009) found in interviewers' handwritten interview records their questions were uncategorisable due to missing question stems and questions being truncated to one or two words.

Almost 30 years ago it was recommended that all child interviews should be visually recorded (Pigot Report, 1989), however this practice was not standardised throughout the different police jurisdictions in Scotland, and it was not until the end of 2011 that it became mandatory for all interviews in Scotland to be visually recorded (Scottish Executive Guidelines, 2011). This means that, in Scotland, the only records that exist for an incalculable number of interviews conducted with children alleging abuse over the course of the last few decades are in handwritten format. Not only is it likely that these records are incomplete through scribes failing to write everything down, but they are also likely to contain inaccuracies. For example, one study demonstrated that 56% of

children's details were attributed to a different utterance type in the investigators' notes than the one that had actually elicited the detail, and these utterances were usually changed from a focussed prompt to a more open one (Lamb et al., 2000). As children's interviews are often the only evidence that investigators have to build a case against alleged perpetrators of child abuse, these handwritten interview records are often presented as evidence when cases appear in courts. Inaccuracies and missing information can have severe consequences on case outcomes, for example, if children say something in court that interviewers fail to record or that contradicts what has been written down, this may harm children's credibility. Factors such as these may further compound the time taken for a case to reach the courts. In addition, illegible handwriting of scribed transcripts may also result in further complications when evidence is used years later (La Rooy et al., 2011).

As VR equipment had only just been introduced to some police forces in Scotland during data collection for this Study, the equipment was still being trialled and so VR facilities were not fully in place yet for conducting field interviews. This meant that while it was possible to visually record trainee interviewers' JIIT role-play interviews, the trainees were also required to complete a practice at scribing an interview in real-time due to the possibility that they would have to do so future. This provided an opportunity to compare the scribed and the electronically recorded versions of trainee investigative interviewers' JIIT role-play interviews conducted with adult actors. It was predicted that there would be less information in the scribed versions of the interviews compared to

244

the electronic recorded versions. Specifically, it was predicted that there would be fewer interviewer questions, fewer actor details and fewer actor words in the handwritten interviews than in the recorded transcript versions. It was also predicted that some questions would be incompletely recorded given the trouble of keeping up pace with speech within the dyad. Potentially this was expected to affect the open prompts to a greater degree than any other question type as interviewers report that lengthy open prompts are most difficult to write down in real time (La Rooy et al., 2011). Finally, based on prior work(Lamb et al., 2000) it was predicted that most of the incorrectly recorded questions would be recorded in a more open manner than they had actually been asked.

As the interviewers in this study are still in training, it is of the utmost importance that we find out whether or not interviewers are capable of keeping accurate verbatim handwritten interview records at this stage before they progress to interviewing real children. If interviewers fail to record information, do not accurately record the type of question that elicited information from children or incompletely record questions, then the interview record is not an accurate reflection of children's accounts and we cannot be sure how information was elicited from the child. Even though Scotland now has implemented VR facilities, it can take years for cases to progress to court, therefore, it is likely that handwritten evidence will still be used in the Scottish court system for the foreseeable future. Thus, it is crucial to confirm whether evidence obtained in this format is reliable, in order to raise awareness among the legal profession.

245

8.2 Method

8.2.1 Participants. Thirty child protection workers took part in this study. Participants (trainee interviewers) were offered the opportunity to take part in the study while undergoing JIIT, and provided written consent at the beginning of the course to take part. One police jurisdiction in Scotland participated in the study, consisting of both police officers (N = 11) and social workers (N= 19). Seven of the total sample were male. Data were collected from three JIIT courses between 2011 and 2012**.**All adult actors who took part in the training (N=3, 1 female, 2 males) were professional hired actors employed to role-play in JIIT courses. All actors held a professional acting qualification and the 2 males were currently working as actors. The female actor was the author of this thesis.

8.2.2 Outline of training procedure. Interviewers were provided with a copy of the Scottish Executive (2011) guidelines at the start of their training ('Guidance on Joint Investigative Interviewing of Child Witnesses in Scotland'). Data were collected from one of the police forces (Force 2) that participated in first study (See Chapter 7.2.2 for details of the course and the conduct of the interviews). The study/training procedure was identical to Chapter 7, except that in addition to the second interviewer scribing the interview by hand, VRI equipment was available and so the interviews were also visually recorded (as VRI became mandatory in Scotland at this point).

8.2.3 *Procedure for coding interviews.* DVD versions of the interviews were provided to the author of this thesis and were then transcribed verbatim to Microsoft Word. The transcription process and the interview coding system was the same as Chapter 7 except for the following changes:

When categorising question types, facilitators were coded as a separate question category rather than being attributed to the previous question type because the focus of this study was to compare whether questions were recorded faithfully. Thus, it was of interest which types of questions were recorded and omitted. Any words and details that followed facilitators were still attributed to the preceding question type.

Three additional supplementary question codes were created for this study:

1. Omitted. When a question was present in the recorded interview but missing entirely from the verbatim record.

2. Incorrect. When the question appeared in both the recorded and the verbatim records but the way it was written down in the handwritten record reflected the question in such a way that its question code was different from the code it was allocated to in the recorded version. For example, "Tell me about where you were yesterday" is coded as a directive but when written as "Tell me about yesterday" it is coded as an invitation. When question codes were changed as a reflection of incorrect recording, the handwritten question was further classed as either 'worse'

(the handwritten question code was less open than it had been asked in recorded interview) or 'better' (the handwritten question code was assigned a more open category, in contrast to the category assigned in the recorded interview).

3. Incomplete. When the handwritten question was not faithfully recorded due to shortening but the question code remained the same in both interviews. For example, "Did he speak to you?" versus "Did he speak to you before he did that?" (both of which are option-posing questions).

Details and words were coded the same way as in Chapter 7, however, each detail and word was coded in both the verbatim and recorded version and any details and words that were recorded but not written down in the scribed version were allocated the supplementary code 'omitted'. Conversely, any details or words that were written down but had not been communicated in the recorded interview were allocated the supplementary code 'added'.

8.2.4 *Coder training and inter-rater reliability.* The lead coder coded all transcripts and the reliability coder (the same reliability coder from Study 2) coded a subset (20%) of randomly selected transcripts, with all disagreements resolved through discussion until an agreement was reached. Reliability in both the identification of question type K = .85, and actor details K =.80 was high.

8.2.5 Outline of analyses. Analyses are conducted on interviewer behaviour followed by actor behaviour in the order in which they occurred during the interview. Wilcoxon tests are used to compare the presence of ground rules and closure rules in the recorded versus scribed versions of interviews. Mixed design ANOVAs are conducted in order to test whether various interviewer and actor behaviours discussed differ according to the interview format and the question type posed. Overall main effects and/or interactions are interpreted with t-tests and/or Bonferroni corrected multiple comparisons of within-subjects' effects (i.e. actor responses to question type and/or interviewer usage of different question types).

8.3   Results

8.3.1   Interviewer behaviour (use of ground rules and closure rules)

For each interview, the presence of ground rules was averaged across the fourteen ground rules. A Wilcoxon Signed Rank Test on the difference between use of ground rules in the handwritten and recorded version revealed an effect of version, where fewer ground rules were present in the handwritten version (*Mdn*=0.25) than the recorded version (*Mdn*=0.36; $z = 3.78$, $p<.001$, *r*=0.49). A Wilcoxon Signed Rank Test on the difference between use of closure rules in the handwritten and recorded version revealed an effect of version, where fewer ground rules were present in the handwritten version (*Mdn*=0.57) than the recorded version (*Mdn*=0.71; $z = 3.57$, $p<.001$, *r*=0.46).

249

8.3.2 Interviewer behaviour (Substantive phase of interview). A two-

way repeated measures ANOVA was conducted on the outcome variable number of total utterances by interviewer, with the factors question type (invitation, directive, option-posing, suggestive, introductory comment, summary, facilitator) and interview format (handwritten, recorded). Greenhousse-Geisser values were used as the assumption of Sphericity was violated. These analyses revealed a main effect of question type (F(3.44, 99.72)=30.66; p<.001, np^2=.51), that was qualified by an interaction with interview format (F(1.53, 44.35)=12.46; p<.001, np^2=.30). The main effect of interview format was also significant (F(1.00, 29.00)=42.97; p<.001, np^2=.60).

Paired samples t tests to interpret this interaction demonstrated that the handwritten interviews contained fewer invitations (t(29)=2.72; p=.011, d=0.50), directives (t(29)=3.19; p<.01, d=0.58), introductory comments (t(29)=4.24; p<.001, d=0.77), summaries (t(29)=5.51; p<.001, d=1.01) and facilitators (t(29)=4.09; p<.001, d=0.75) than the recorded versions. There were no differences in the numbers of option-posing or suggestive questions recorded across different interview formats (both t<1.86, both p>.074, see Table 8.1).

250

Table 8.1: Descriptive statistics: Use of question types according to two different versions (handwritten, recorded)

Question type	Mean Handwritten (SEM)	Mean Recorded (SEM)
Invitation	14.23 (1.14)	15.60 (1.27)
Directive	11.27 (1.28)	12.57 (1.38)
Option-posing	6.97 (.85)	7.53 (.79)
Suggestive	1.87 (.30)	2.00 (.31)
Introductory comment	2.43 (.47)	14.57 (2.90)
Summary	.80 (.19)	5.03 (.76)
Facilitator	.93 (.31)	4.93 (1.16)

A repeated measures ANOVA with *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary, facilitator) as the repeated measure was conducted on the outcome variable proportion of omitted utterances by interviewer. The Greenhouse-Geisser correction was used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed a main effect of *question type* ($F(2.80, 81.09)=17.87$; $p<.001$, $np^2=.38$).Post-hoc analyses (corrected for multiple comparisons using Bonferroni testing) demonstrated that a greater proportion of introductory comments were omitted compared to invitations, directives, option-posing and suggestive questions (all $p<.01$). A greater proportion of summaries were omitted than directives, option posing and suggestive questions (all $p<.01$) and a greater proportion of facilitators were omitted than option-posing and suggestive questions (both $p<.046$). No other comparisons were significant (all Bonferroni-corrected $p>.054$, see Table 8.2).

Table 8.2: Proportion of different question types omitted from the substantive phase in the handwritten interviews (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	.11 (.04)	.03	.19
Directive	.03 (.01)	.01	.04
Option-posing	.02 (.01)	.01	.04
Suggestive	.02 (.01)	<.001	.04
Summary	.20 (.03)	.13	.27
Facilitator	.18 (.05)	.09	.27
Introductory comment	.44 (.05)	.33	.55

A repeated measures ANOVA with *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary) as the repeated measure was conducted on the outcome variable number of incompletely recorded utterances in the handwritten interview. The Greenhouse-Geisser correction was used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed a main effect of *question type* ($F(2.11, 61.07)=9.49$; $p<.001$, $np^2=.25$).Post-hoc analyses (corrected for multiple comparisons using Bonferroni testing) demonstrated that a greater number of invitations were incompletely recorded compared to suggestive questions, introductory comments and summaries (all $p>.001$). Also a greater proportion of option-posing questions were incompletely recorded compared to summaries (all $p>.01$). No other comparisons were significant (all $p<.055$),see Table 8.3.

Table 8.3: Number of different question types incompletely recorded in the substantive phase of the handwritten interviews (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	1.47 (.26)	0.93	2.00
Directive	1.20 (.40)	0.39	2.01
Option-posing	0.60 (.15)	0.30	0.90
Suggestive	0.13 (.08)	-0.03	0.30
Introductory comment	0.17 (.08)	-0.01	0.34
Summary	0.03 (.03)	-0.04	0.10

A repeated measures ANOVA with *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary) as the repeated measure was conducted on the outcome variable number of incorrectly recorded utterances in the handwritten interview. The Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed a main effect of *question type* ($F(2.13, 61.68)=5.51$; $p<.01$, $np^2=.16$). Post-hoc analyses (corrected for multiple comparisons using Bonferroni testing) demonstrated that a greater number of invitations were incorrectly recorded compared to summaries and introductory comments (all $p<.019$). No other comparisons were significant (all $p>.094$), see Table 8.4.

253

Table 8.4: Number of different question types incorrectly recorded in the substantive phase of the handwritten interviews (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	0.67 (.15)	0.35	0.98
Directive	0.77 (.27)	0.22	1.31
Option-posing	0.23 (.09)	0.05	0.42
Suggestive	0.20 (.07)	0.05	0.35
Introductory comment	0.03 (.03)	-0.04	0.10
Summary	0.13 (.08)	-0.03	0.30

Next, the first use of focussed prompts (directives and option-posing questions) were compared across interview format. A two-way repeated measures ANOVA was conducted on the outcome variable proportion of questions asked before focussed prompts with the factors *question type* (directive, option-posing) and *interview format* (handwritten, recorded). This analysis revealed a main effect of *focussed question type* ($F(1,28)=10.32$; $p<.01$, $np^2=.27$) that was not qualified by an interaction with *interview format* ($F(1,28)=2.78$; $p=.11$). There was also a main effect of *interview format* ($F(1,28)=11.37$; $p<.01$, $np^2=.29$). Across interview formats, trainees introduced directives earlier in the conversation (12.01% of utterances) than they introduced option-posing questions (22.70% of utterances; $t(26)=2.73$; $p=.011$, $d=0.54$). To interpret the effect of interview format, paired samples t tests revealed that focussed prompts were recorded as being introduced earlier into the conversation in the recorded version ($M_{utterances}=15.10\%$, $SEM=1.87\%$) than they were in the handwritten version ($M_{utterances}=19.24\%$, $SEM=2.21\%$; $t(28)=3.37$; $p<.01$, $d=0.64$).

254

8.3.3   Actor behaviour. A two-way repeated measures ANOVA was then conducted on the outcome variable number of total details provided by actor, with the factors question type (invitation, directive, option-posing, suggestive, introductory comment, summary) and interview format (handwritten, recorded). The Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test (p<.001). These analyses revealed a main effect of question type (F(1.22,35.42)=41.63; p<.001, np^2=.59) that was not qualified by an interaction with interview format (F(1.43,41.41)=3.26; p=.064, np^2=.10). A main effect of interview format was also significant (F(1.00,29.00)=17.71; p<.001, np^2=.38). Post-hoc analyses collapsed across interview format (corrected for multiple comparisons using Bonferroni testing) demonstrated that all six question types elicited a different amount of detail from the actors over the course of the interview (all p<.047), except for option-posing and suggestive questions, and summaries and introductory comments, which all elicited an equivalent amount of detail across interview formats (all p>.69). Paired samples t tests collapsed across question type revealed that handwritten versions of interviews contained fewer details from the actor (*M*=17.41, *SEM*=2.15) than recorded versions (*M*=19.43, *SEM*=2.37; *t*(29)=4.21; *p*<.001, *d*=0.78).

A repeated measures ANOVA was conducted on the outcome variable number of details omitted in the handwritten interviews, with *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary) as the repeated measures factor. Greenhouse-

Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed no main effect of *question type* (F(1.43, 41.41)=3.26; p=.064, np^2=.10).

A repeated measures ANOVA was conducted on the outcome variable number of details added in the handwritten interviews with the factor *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary). Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed no main effect of *question type* (F(2.78, 80.53)=0.68; p=.56).

Next, the first use of focussed prompts (directives and option-posing questions) were compared across interview format. A two-way repeated measures ANOVA was then conducted on the outcome variable proportion of details elicited before focussed prompts with the factors *question type* (directive, option-posing) and *interview format* (handwritten, recorded). This analysis revealed a main effect of *focussed question type* (F(1,26)=6.56; p=.017, np^2=.20) that was not qualified by an interaction with *interview format* (F(1,26)=.004; p=.953). Across interview formats, trainees elicited fewer details before directives were introduced into the conversation ($M_{utterances}$=12.44, *SEM*=4.02), than they elicited before option-posing questions were introduced into the conversation ($M_{utterances}$=25.07, *SEM*=5.01).

A two-way repeated measures ANOVA was then conducted on the outcome variable number of total words provided by actor, with the factors *question type* (invitation, directive, option-posing, suggestive,

introductory comment, summary) and *interview format* (handwritten, recorded). Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). These analyses revealed a main effect of *question type* ($F(1.20, 34.87)=67.82$ $p<.001$, $np^2=.70$) that was qualified by an interaction with *interview format* ($F(1.35, 39.11)=14.77$, $p<.001$, $np^2=.34$). There was also a main effect of *interview format* ($F(1,29)=63.32$; $p<.001$, $np^2=.69$). Paired samples t tests to interpret this interaction demonstrated that the handwritten interviews contained fewer words from invitations ($t(29)=4.85$; $p<.001$, $d=0.89$), directives ($t(29)=3.69$; $p=.001$, $d=0.67$), option-posing questions ($t(29)=2.75$; $p=.01$, $d=0.50$), and introductory comments ($t(29)=3.30$; $p<.01$, $d=0.60$), than the recorded versions. There were no differences in the number of words recorded between the two interview formats for actor responses to suggestive questions or summaries (all $t<1.14$, all $p>.26$),see Table 8.5.

Table 8.5: Number of words from different question types during the substantive phase of the interview

Question type	Mean Handwritten	Mean Recorded	95% CI of the difference
Invitation	256.07	356.33	(-142.51, -58.02)
Directive	90.23	121.87	(-49.17, -14.09)
Option-posing	24.33	32.03	(-13.42, -1.98)
Suggestive	14.53	19.30	(-13.37, 3.84)
Summary	7.07	10.37	(-10.21, 3.61)
Introductory comment	5.40	15.40	(-16.20, -3.80)

A repeated measures ANOVA with conducted on the outcome variable number of words omitted in the handwritten interviews, with *question type*

257

(invitation, directive, option-posing, suggestive, introductory comment, summary) as the factor. Greenhouse-Geisser corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed a main effect of *question type* (F(1.31, 38.02)=18.12; $p<.001$, np^2=.39). Post-hoc analyses (corrected for multiple comparisons using Bonferroni testing) demonstrated that words elicited by invitations were omitted more often than words elicited by option-posing, suggestive, introductory comments and summaries (all $p<.002$). Words elicited by directives were omitted more often than words elicited by option posing questions, suggestive questions, introductory comments and summaries (all $p<.037$ see Table 8.6). No other comparisons between question type were significant (all $p>.078$).

Table 8.6: Number of words omitted from the substantive phase of the handwritten interviews, collapsed across interview format (adjusted means reported)

Question type	Mean (SEM)	95% CI (Lower bound)	95% CI (Upper bound)
Invitation	104.10 (19.80)	63.61	144.59
Directive	35.47 (7.51)	20.10	50.83
Option-posing	8.27 (2.72)	2.70	13.83
Suggestive	8.77 (3.22)	2.17	15.36
Introductory comment	11.27 (2.76)	5.62	16.92
Summary	6.83 (2.70)	1.30	12.36

A repeated measures ANOVA with *question type* (invitation, directive, option-posing, suggestive, introductory comment, summary) as the repeated measure was conducted on the outcome variable number of words added in the handwritten interviews. Greenhouse-Geisser

corrections were used as the assumption of sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed no main effect of *question type* ($F(2.73, 79.30)=.64$; $p=.58$, $np^2=.02$),

8.4 Discussion

Using in-depth, systematic coding, this study finds several differences in how someone may interpret the quality of JIIT interviews (or field interviews) depending on whether they analyse a handwritten verbatim record or a video recording of the same session. The analyses reported here suggest that handwritten versions of JIIT include, on average, 11% fewer ground rules and 14% fewer closure principles from the interviewer than video-recorded versions of the interview, which could affect the appearance of the quality of an interview and perceptions of adherence to Scottish Executive (2011) guidelines. Moreover, handwritten versions of interviews captured fewer invitations and directive questions from the interviewer. Despite this, when examining the total proportion of omissions in the handwritten interview, introductory comments, summaries and facilitators were more likely to be omitted than more substantial question types of this nature, which may suggest a deliberate strategy used during the transcription process with regards to these utterances. Indeed certain interviewer utterances were recorded faithfully, such as suggestive and option-posing questions. Of concern however, is that invitations were more likely to be incompletely and incorrectly recorded in handwritten interviews compared most other question types. In addition, a subtle but significant difference was observed in the introduction of focussed prompts in scribed versus

recorded versions of interviews, where the interviewer made 4% more utterances in the handwritten than recorded version before introducing focussed prompts.

Analyses of differences between scribed and recorded versions of actor responses suggest a somewhat more promising pattern of results. Both formats were accurate in capturing the proportion of details elicited by the actor before the introduction of focused prompts. Across question types, scribes were not likely to add words from the actor and any omissions or additions when scribing details provided by the actor were not specific to question type. Scribes were also accurate in recording words in response to certain question types, such as suggestive questions and summaries. However, words elicited from invitations were omitted more than other question types, except for directives, where words elicited from this category were also more likely to be omitted than other categories. Finally, regardless of question type, scribes failed to record details from the actor, omitting approximately two details from their handwritten interviews.

The findings reported here suggest that practical problems with scribing questions and answers in real time can have an effect on the quality of evidence when researchers or legal practitioners refer to field interviews. Fewer recordings of introductory comments, summaries and facilitators may reflect a conscious strategy on the part of the 'scribe'. Scribes may omit such 'procedural' or repeated questions if they feel they are unlikely to elicit new incident-related details and if this helps them to keep up with the pace of the interview. As these records are supposed to

be verbatim (Scottish Executive, 2003), this may have implications if additional lines of questioning are of utility for cases later. Consistent with prior concerns raised by Scottish interviewers (La Rooy et al., 2011), invitations may be omitted from the handwritten record as they are the longest questions to write down, particularly when the invitations are time segmented (e.g. "Tell me everything that happened from the time you walked into his house until the time that you left his house and went home") or cued (e.g. "You said that he was always doing things to you that you didn't like, tell me about the things he did to you that you didn't like"). This is concerning from the point of view of establishing how allegation related information was elicited from a child and for establishing the extent to which best practice questions (i.e. open prompts) are used in interviews.

Although absolute numbers were small, invitations were more likely than other types of questions to be incorrectly written down. In contrast to the interviewers in Lamb et al. (2000), interviewers' errors in recording question type were not attempts to allocate the question to a more desirable type (i.e. to improve the appearance of the interview). Here, invitations could only be afforded a less open code when recorded incorrectly and invitations accounted for the largest proportion of incorrectly recorded questions. It is likely that incorrectly recorded questions resulted both from genuine confusion (attempting to listen and scribe simultaneously) and attempts to shorten questions to keep up with the pace of speech (invitations were one of the two types of question most likely to be *incompletely* recorded). In contrast to prior work where

261

incomplete recording rendered question types indistinguishable (due to missing question stems; Cauchi and Powell, 2009) trainee interviewer's incomplete questions included the stem but not the content of the question. For example trainees would record "tell me…" as opposed to "tell me about him touching you?"

These findings have other important implications for practice. While this may be a time saving technique, the stem of the question does not always reflect the correct code and the *content* of the question must be taken into account because an additional observation was that attempts made to note questions in a shorter format ended up changing the question code in some instances. For example, if the child had not mentioned being touched it would be coded as a suggestive question, but this is impossible to gauge without complete recording of the question. This is of concern as a child's testimony could be perceived as contradictory if they answer a question differently in a subsequent interview or courtroom, despite this discrepancy actually being a reflection of incorrectly or incompletely recorded questions in the handwritten interview record. Incomplete questions are difficult to rely on and provide a poor-quality evidence base for effective justice in child abuse cases. In conclusion, incompletely recorded questions may appear to be a useful time saving strategy, however, the only way of ensuring that handwritten interviews are reliable records is if they are complete verbatim records of all questions and answers during an interview.

Differences in the details and/or words recorded from actors raise further issues. It is concerning that 74% of the words lost from

interviewee responses had been elicited by recall prompts. Consistent with La Rooy et al. (2011), these omissions may reflect difficulties by the scribe in recording lengthier average response to invitations. Conversely, focussed questions can be answered with a few words (e.g. yes/no questions) making them quicker to write down and more likely to be scribed accurately. Clearly this has strong implications for accurate recording of the details and/or responses children provide in field interviews to best-practice questions, and in establishing an evidence base that is admissible in court. Moreover, interviewers whose interviews are scribed may pick up on this in an effort to assist the scribe and, in turn, pose fewer open questions to children and conduct a poorer-quality interview as a result.

Critically, although absolute numbers were small, 71% of the details lost through scribing were elicited from recall prompts, which is problematic as open prompts elicit accurate information (Lamb &Fauchier, 2001; Lamb et al., 2007b; Orbach & Lamb, 1999, 2001; Orbach, Lamb, La Rooy & Pipe, 2012). In other words, details that are most likely to be accurate were more likely to be lost due to arbitrary practical concerns of scribing lengthy answers. Encouragingly, however, details or words added to a transcribed account did not vary according to question type, suggesting that 'scribes' did not try to manipulate responses in their favour (i.e. to present the interview as better quality than it was).

The loss of information in this study was smaller than other field studies. For example, while this study found 10% fewer details and 38%

263

fewer questions were recorded, Lamb and colleagues (2000) found 25% fewer details and 57% fewer questions recorded through scribing. It is important to note, however, that the interviewers in this study are 1) still under observation during this training period and 2) are interviewing adult actors who may purposefully try to assist the scribe in keeping up with the interview pace. The trainee interviewers are aware that their accounts will be examined by their trainers following the interview role-plays and that there is a mock 'court' scenario on the final day of their JIIT course where some interviews will be subject to scrutiny. This may ensure that interviewers try to adhere as closely as possible to the rule that their scribed record must be verbatim whereas in the field there is no way of assessing the accuracy of the scribed account. The actors may speak slower and tolerate long pauses during which they know they should remain silent to allow the scribe to catch up. Scribes may not be afforded this type of assistance with real children. Indeed, experimental research that has assessed interviewers' ability to scribe a live interview script that was read at an optimal pace found that interviewers still failed to record 39% of the abuse-related details (Cauchi et al., 2010).

In sum, even under more favourable conditions than a real interview with a child, on some key dimensions (e.g. related to open prompts) interviewers did not keep entirely faithful handwritten interview records. Questions and information from responses were lost and questions were sometimes incompletely or incorrectly recorded. When expert witnesses testify in child abuse cases often they assess the quality of children's evidence by examining the quality of the interview, which is

264

often done by analysing the types of questions that elicited information from children. Thus, a poorly scribed interview could potentially jeopardise a case if information is recorded as being elicited from a different utterance type than it actually had been. Therefore relying on such records when making legal decisions is potentially hazardous and caution should be exercised when basing legal decisions on handwritten 'verbatim' records of forensic interviews with children. Prior legal decisions based on handwritten interviews may have been made based on inaccurate records of questions and answers during an interview. Moreover, if trainee interviewers find recording invitations and the responses they elicit difficult, in contrast to shorter focussed questions and their shorter responses, then, at an early stage in their career they may opt to use shorter questions in real interviews in an effort to keep accurate handwritten records. The final empirical Chapter of this thesis examines the quality of an actual sample of Scottish field interviews with children.

Chapter 9. Examining the quality of Scottish Investigative Interviewers' training interviews with adult actors

9.1 Study rationale and aims

Given the shortcomings of using hired actors to role-play abused children in JIIT interviews found in the first study, it is clear that such 'untrained' adult actors (i.e. professional actors who receive no prior training or guidance) are exhibiting undesirable behaviours that are not conducive to encouraging best practice interviewing skills. Interviewers are taught to rely on open prompts, yet actors' lack of sensitivity to question type was evident in a number of ways in the prior study. Actors often responded comparably to open prompts and directive questions (e.g. the proportions of both total details and total words provided and the average number of details and words provided per response). In some cases, actor behaviours may have discouraged the use of open questions by the interviewer. For example, invitations elicited the highest proportion of uninformative responses from actors compared to all other question types and open prompts were no more likely to elicit the first detail than any other question types. In addition, it was reported, anecdotally, that actors refused to engage in narrative elaboration training by refusing to respond to open prompts during the pre-substantive phase.

Therefore, the focus of this study is to test whether it is possible to improve the efficacy of the role-play interview practice sessions for trainee investigative interviewers by implementing an adult actor with expertise in interviewing best practice on JIIT. This was examined by

266

comparing interviews conducted with a 'trained' adult actor (a postgraduate researcher with knowledge of best-practice questioning and child responses, who had been instructed to responded preferentially to open prompts) against interviews conducted with an 'untrained' actor. It was predicted that the role-play interviews conducted with the trained actor would be of a higher quality than those conducted with the 'untrained' actor. Firstly, it was predicted that the trained actor would provide longer responses to open prompts during the pre-substantive phase than the untrained actor. Secondly, it was predicted that the trained actor would provide a greater proportion of details and greater average number of details within a response to open prompts during the substantive phase compared to the untrained actor. Third, it was predicted that the trained actor would provide a greater proportion of words and a greater number of words within a response to open prompts during the substantive phase compared to the untrained actor. Fourth, it was predicted that the trained actor would provide a greater proportion of informative responses to open prompts during the substantive phase compared to the untrained actor. Finally, the trained actor was expected to provide their first substantive detail in response to an open prompt more often than the untrained actor.

An effect of expertise on preferential responding to open prompts by the actor was expected to be related to interviewers' questioning behaviour. Previous research has shown that adult actors with knowledge of forensic interviewing and question type have improved interviewers' use of open questions (Freeman & Morris, 1999; Powell et al., 2008).

267

Therefore, it was hypothesised that interviewers in the trained actor condition compared to the untrained actor condition would 1) ask a greater proportion of open prompts during the substantive phase of the interviews, 2) delay their use of focussed prompts until later in their interviews and 3) have elicited a greater proportion of details before asking focussed prompts.

9.2 Method

9.2.1 Participants. Forty-four child protection workers took part in the study (17 police officers and 27 social workers, 15 of the sample were male. The sample consisted of members of force two from the first study (Chapter 7). Participants were offered the opportunity to take part in the research while undergoing JIIT. At the beginning of their training, the trainer(s) obtained written consent from trainees. One police jurisdiction in Scotland participated in the study. Data were collected from five JIIT courses that ran between 2010 and 2012.

The untrained adult actor was a professional hired actor who was employed to participate in JIIT role-play training interviews. The actor was not provided with specific instructions about different phases of an eyewitness interview (e.g. including Narrative Elaboration Training) or how to respond to different types of questions. Instead, the actor only had an informal discussion with trainers where they could ask any questions about their character or role. The trained actor was a psychology PhD student whose research focuses on forensic interviews of children and has extensive knowledge of question types and age-appropriate

responses to questions, and holds a professional acting qualification[1]. In preparation for this role the trained actor became familiar with question type and realistic responses during interviews by watching DVDs of real forensic interviews with children, attending a week-long interview coding workshop and CPD courses on assessing the quality of interviews and practice coding of samples of field interviews. The trained actor was instructed to 1) engage with interviewer attempts to elicit narratives and selectively reward open prompts with responses of a greater average length compared to all other question types during the pre-substantive phase, 2) to provide a greater number of details per average response to the open prompts compared to all other prompt types during the substantive phase and 3) to provide a greater number of words per average response to the open prompts compared to all other prompt types during the substantive phase.

9.2.2 *Detailed outline of training procedure and coding of interviews.* Each interviewer was provided with a copy of the current Scottish child interviewing guidelines (either the 2003 "Guidance on Interviewing Child Witnesses in Scotland" or the 2011 "Guidance on Joint Investigative Interviewing of Child Witnesses in Scotland") at the start of their training. The method of training was identical to that received by force two (see Method, Chapter 7). The method for transcribing and

---

[1] The trained actor in this study is also the author of this thesis. Any potential limitations of having the author involved in this stage of the research were controlled for by ensuring a random sample of the data were coded by a second coder to ensure reliability in the coding.

coding interviews was identical to Chapter 7, except that interviewer questions and length of actor responses were also coded during the pre-substantive phase of the interview. Interviewer questions in the pre-substantive phase were coded either as invitation, directive, option-posing, suggestive, facilitator or introductory comment (any remark about the procedural aspects of the interview). The number of words spoken by the actor in response to each question type was recorded and the average number of words per response from actors to each question type was recorded (any words following facilitators were attributed to the question type the facilitator followed).

9.2.3 Coder training and inter-rater reliability. The lead coder trained a different coder to the previous Chapter on a set of independent transcripts. The lead coder coded all transcripts and the reliability coder independently coded 20% of the transcripts that were randomly selected, with all disagreements resolved through discussion. Reliability in identifying question type was high (K = .82) and reliability in identifying actor response type (K=.73) and actor details (K=.76) were good.

9.2.4 Outline of analyses. Analyses conducted were identical to Chapter 7, except that the between subjects' factor was the expertise of actor (trained, untrained). Given that trainees were all from the same force, this was no longer a factor in the current study. In addition, interviewer and actor behaviour during the pre-substantive phase of the interview was also measured (the same outcome variables for interviewer questions and actor behaviour were measured for the pre-substantive

270

phase as the outcome variables for interviewer questions and actor

behaviour during the substantive phase of the interview).

9.3 Results

9.3.1 Interviewer behaviour (pre-substantive phase of interview)

Binomial tests run across interviews demonstrated that the ground rules

'don't understand', and 'truth' were used at levels greater than would be

expected by chance. The ground rules 'don't understand demonstration',

'correct me demonstration', 'don't remember', 'NET open' were all used at

levels significantly less than would be expected by chance. All other

ground rules were not used more or less than would be expected by

chance (see Table 9.1 for values).

**Table 9.1: Presence of ground rules at pre-substantive phase,
across all interviews**

Rule	Mean proportion usage (*SEM*)	Binomial test result
Listen	0.48 (.08)	*p*=.88
Knowledge	0.36 (.07)	*p*=.10
Don't know	0.61 (.07)	*p*=.18
Don't know demonstration	0.34 (.07)	*p*=.05
Don't understand	0.73 (.07)	*p*<.01
Don't understand demonstration	.07 (.04)	*p*<.001
Correct me	.43 (.08)	*p*=.45
Correct me demonstration	.27 (.07)	*p*<.01
Repeated questions	.41 (.08)	*p*=.29
Don't remember	.32 (.07)	*p*=.02
Don't guess	.36 (.07)	*p*=.10
Truth	.73 (.07)	*p*<.01
NET	.34 (.07)	*p*=.05
NET Open	.20 (.06)	*p*<.001

271

9.3.2 Interviewer behaviour (Substantive phase of interview). First, a mixed-design ANOVA was conducted on the outcome variable proportion of total utterances by interviewer, with the within-subjects factor question type (invitation, directive, option-posing, suggestive, summary, non-substantive) and the between subjects' factor expertise of actor (trained, untrained). Greenhouse-Geisser corrections were used as Sphericity was violated according to Mauchly's test ($p<.001$). This analysis revealed a main effect of question type ($F(2.79, 117.22)=132.78$; $p<.001$, $np^2=.76$) that was not qualified by an interaction with expertise of actor ($F(2.79, 117.22)=1.29$; $p=.28$). Post-hoc analyses to interpret the main effect of question type demonstrated that all combinations of question types were used to a different extent during interviews, except for invitations and directives, suggestive questions and non-substantive questions, and summaries and non-substantive questions which were all used by interviewers to the same extent as one another during interviews (all non-significant Bonferroni-corrected p values >.07, see Figure 9.1).

272
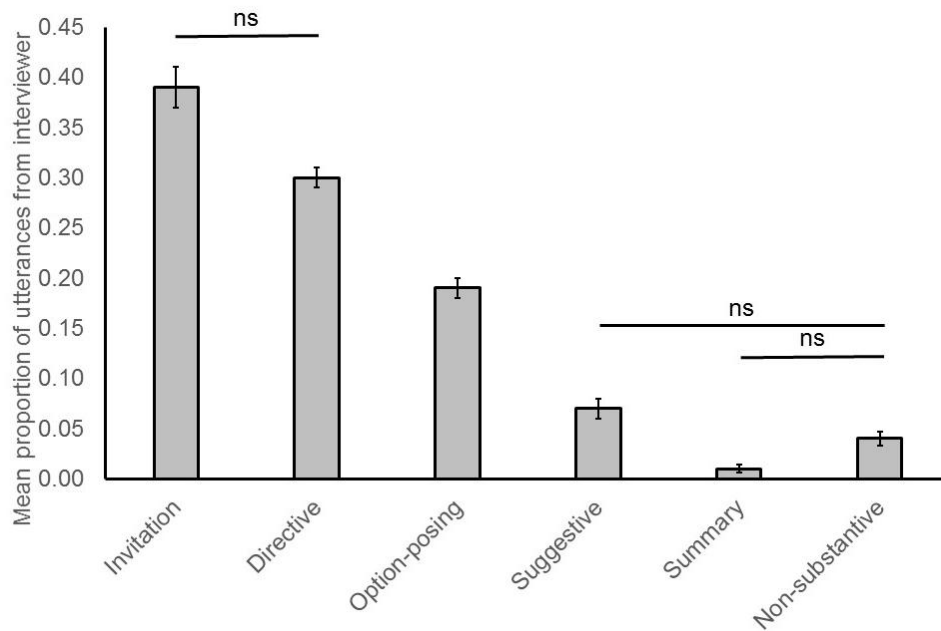
Figure 9.1: Mean proportion of total interviewer utterances according to question type Error bars show +- 1SEM.

Next, an ANOVA was run on the proportion of questions used before the introduction of focussed prompts, with the within subjects' factor *question type* (option-posing, directive) and the between subjects' factor *expertise of actor* (trained, untrained). This analysis revealed a main effect of *focussed prompt* ($F(1,42)=14.59$; $p<.001$, $np^2=.26$) that was not qualified by an interaction with *expertise of actor* ($F(1,42)=.87$; $p=.36$). There was no main effect of expertise of actor ($F(1,42)=.41$; $p=.52$). Paired samples t tests to interpret the main effect of focussed prompt revealed that, regardless of actor, interviewers introduced directives earlier into the conversation ($M=0.14_{questions}$, $SEM=.02$) than they introduced option-posing questions ($M=0.25_{questions}$, $SEM=.03$, $t(43)=3.76$; $p<.01$).

273

9.3.3 *Closure phase of interview.* Binomial tests run across interviews demonstrated that the ground rules 'thanks', and 'child questions' were used at levels greater than would be expected by chance. The ground rule 'summary' was used less than would be expected by chance. All other ground rules were not used more or less than would be expected by chance (see Table 9.2 for values).

Table 9.2: Presence of ground rules at closure phase, across all interviews

Rule	Mean proportion usage (*SEM*)	Binomial test result
Next	.39 (.07)	*p*=.18
Contact	.57 (.08)	*p*=.45
Neutral	.61 (.07)	*p*=.18
Thanks	.77 (.06)	*p*<.01
Child questions	.77 (.06)	*p*<.01
Interviewer 2 questions	.39 (.07)	*p*=.18
Summary	.07 (.04)	*p*<.001

9.3.4 *Actor behaviour (Pre-substantive phase of interview).* As this data was not suitable for ANOVA (e.g. in some cases, too few instances of specific question types posed during this phase to both actors), independent samples t tests were conducted separately on the average length of responses to invitations, directives, option-posing questions and introductory comments, to test for an effect of expertise of actor. Data from suggestive questions were not analysed given that the number of

274

suggestive questions across participants was particularly low. The trained

actor provided longer average responses to invitations (M =29.61, SEM=

3.41) than the untrained actor (M =10.91, SEM = 2.60, $t(38)$=4.10;

p<.001, d=1.32, see Table 9.3). No other comparisons between trained

and untrained actors were significant (all t< all p>.052).

Table 9.3: Average number of words (SD) elicited from the actor during the introductory phase of the interview

Question type	Untrained	Trained
Invitation	10.91 (10.74)	29.61 (16.37)
Directive	7.05 (6.59)	16.92 (19.23)
Option-posing	3.98 (4.06)	6.79 (11.86)
Introductory comments	1.75 (1.34)	1.21 (1.28)

9.3.5 Substantive phase of interview. A mixed-design ANOVA was

conducted on the outcome variable proportion of informative responses,

with the within subjects' factor question type (Invitation, directive, option-

posing, suggestive) and the between subjects' factor expertise of actor

(trained, untrained). Greenhouse geisser values were used as the

assumption of Sphericity was violated. This analysis revealed a main

effect of question type ($F(2.32, 81.10)$=3.21; p<.05, np^2=.08) that was not

qualified by an interaction with expertise of actor ($F(2.32, 81.10)$=2.26;

p=.10, np^2=.06).Paired samples t tests to interpret the main effect of

question type revealed that, regardless of actor, both directives (M=.91,

SEM=.02) ($t(36)$= 2.31; p<.05) and option-posing questions (M=.93,

SEM=.02), ($t(36)$=2.43, p<.05) elicited a greater proportion of informative

275

responses than suggestive questions ($M_{Suggestive}$=.81, *SEM*=.05). No other comparisons were significant (all *t*<, all *p*>.08).

A mixed-design ANOVA was conducted on the outcome variable proportion of details elicited from actors, with the within subjects factor *question type* (invitations, directives, option-posing, suggestive, summaries, introductory comments) and the between subjects factor *expertise of actor* (trained, untrained). Greenhouse Geisser corrections are reported as the assumption of Sphericity was violated. This analysis revealed a main effect of *question type* (F(1.93, 79.10)=137.68; *p*<.001, np^2=.77) that was qualified by an interaction with *expertise of actor* (F(1.93, 79.10)=4.99; *p*=.01, np^2=.11). Independent samples t tests to interpret this interaction demonstrated that the trained actor provided more details than the untrained actor in response to invitations (*t*(41)=2.75; *p*<.01, *d*=0.85). The untrained actor provided more details than the trained actor in response to option-posing questions (*t*(24.59)=2.20; *p*=.037, *d*=0.90). Expertise of actor did not have an effect on number of details elicited for the remaining question types (all *t*<1.97, all *p*>.064, see Figure 9.2).

Figure 9.2: Proportion of details provided by different actors. The trained actor provided more details than the untrained actor in response to invitations but fewer details than the untrained actor in response to option-posing questions.

The data for average number of details per actor response was not suitable for ANOVA as too few trainees used all six prompt types. Therefore, independent samples t tests were conducted individually on the average number of details per response to invitations, directives, option-posing, suggestive, summaries and introductory comments, to test for an effect of expertise of actor. The trained actor($M = 4.87$, $SEM = .46$) gave richer responses to invitations than the untrained actor ($M = 2.23$, $SEM = .42$, $t(42)=4.17$; $p< .001$, $d=1.29$).Average number of details did not differ between trained and untrained actors for any remaining prompt types (all $t<1.73$, all $p>.09$).

An ANOVA was run on the outcome variable proportion of times the first detail was elicited from actor, with the within subjects' factor *question type* (invitation, directive, option-posing, suggestive) and the

277

between subjects' factor *expertise of actor* (trained, untrained). Greenhouse geisser values were used as Sphericity was violated. This analysis revealed a main effect of *question type* ($F(2.37, 99.43)=20.94$; $p<.001$, $np^2=.33$) that was not qualified by an interaction with *expertise of actor* ($F(2.37, 99.43)=2.69$; $p=.063$, $np^2=.06$). Bonferroni-corrected pairwise comparisons collapsed across actor revealed that invitations ($M_{Invitations}=.67$, *SEM*$=.07$) elicited the first detail more often than all the three other question types ($M_{Directives}=.09$, *SEM*$=.05$, $M_{OptionPosing}=.13$, *SEM*$=.05$, $M_{Suggestive}=.11$, *SEM*$=.05$, all $p<.001$). No other comparisons were significant (all $p=1.00$).

An ANOVA was then conducted on the percentage of details elicited before the introduction of a focussed prompt, with the within-subjects factor *focussed prompt* (directive, option-posing question) and the between subjects' factor *expertise of actor* (trained, untrained). This revealed a main effect of *question type* ($F(1,41)=15.10$; $p<.001$, $np^2=.27$) that was not qualified by an interaction with *expertise of actor* ($F(1,41)=.02$; $p=.89$). There was no main effect of *expertise of actor* ($F(1,41)=.09$; $p=.76$). Paired samples t tests revealed that across actors, a smaller proportion of details were elicited before directives ($M=13.02\%$, *SEM*$=2.75\%$) were introduced into the conversation, in contrast to the proportion of details elicited before option-posing questions were introduced into the conversation ($M=24.43\%$, *SEM*$=3.53\%$, $t(42)=3.98$; $p<.001$, $d=0.61$).

An ANOVA was conducted on the proportion of words elicited by the actor, with the within subjects' factor *question type* (invitations,

278

directives, option-posing, suggestive, summaries, introductory comments) and the between subjects' factor *expertise of actor* (trained, untrained). Greenhouse Geisser values were used as the assumption of Sphericity was violated. This analysis revealed a main effect of *question type* (F(1.54, 64.85)=196.10; p<.001, np^2=.82) that was qualified by an interaction with *expertise of actor* (F(1.54, 64.85)=11.72; p<.001, np^2=.22). Independent t tests to interpret the interaction revealed that the trained actor (*M*=.65, *SEM*=.03) provided a greater proportion of words to invitations than the untrained actor (*M*=.46, *SEM*=.04, t(42)=4.22; p<.001, *d*=1.32), and the untrained actor provided a greater proportion of words to directives (*M*=.37, *SEM*=.04) than the trained actor (*M*=.23, *SEM*=.02, t(42)=2.98; p<.01, *d*=0.93). No other differences in length of response per question type were found between actors (all t<1.86, all p>.07).

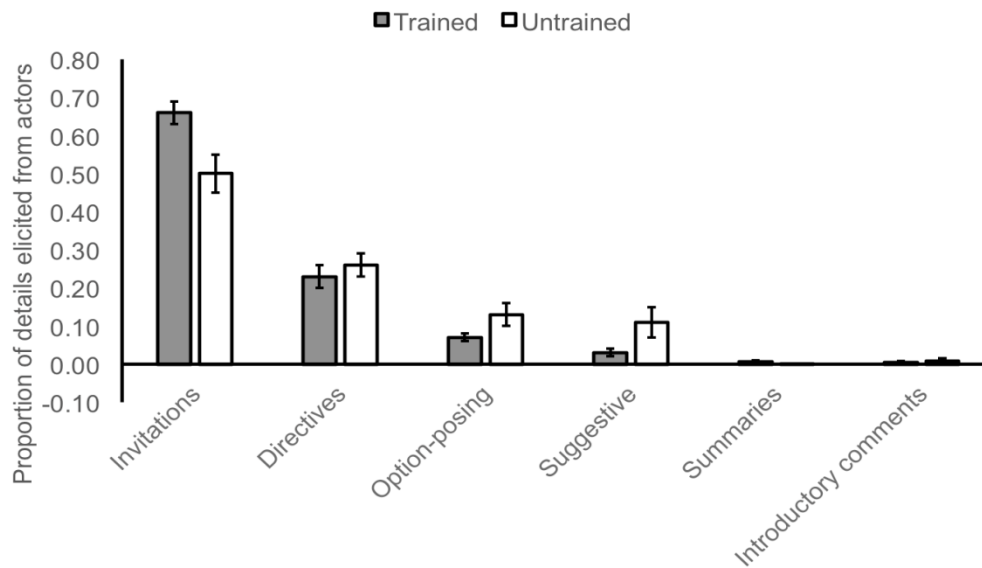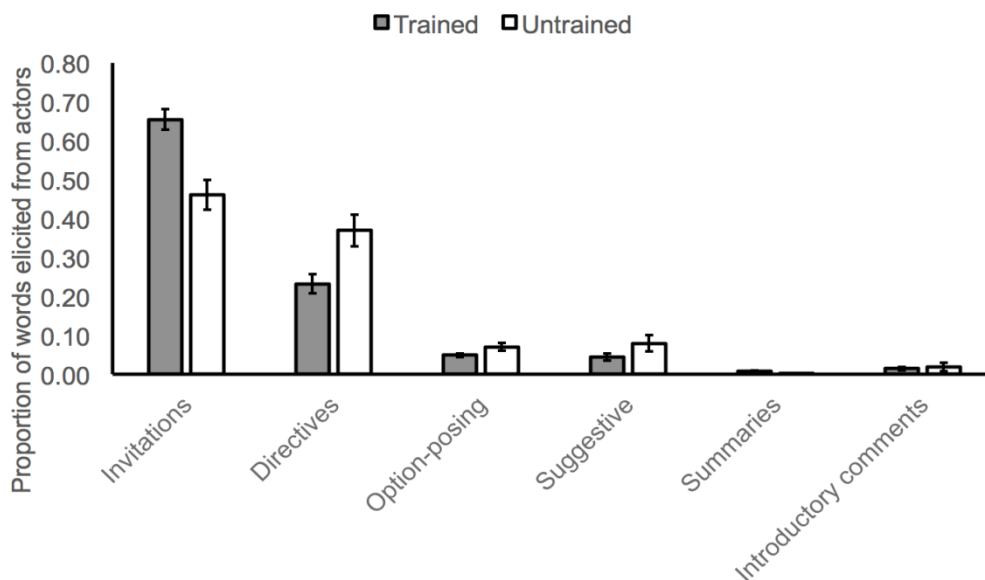**Figure 9. 3**: Proportion of words provided by different actors. The trained actor provided more words than the untrained actor in response to invitations but fewer words than the untrained actor in response to directive questions.

The data on average length of actor response were not suitable for ANOVA as too few trainees used all six prompt types across both actors. Therefore, independent samples t tests were conducted separately on the average length of response to invitations, directives, option-posing, suggestive, summaries and introductory comments, to test for an effect of expertise of actor. The trained actor gave longer responses to each invitation from the interviewer $t(42) = 6.01$, $p< .001$, ($M = 16.97$, $SD = 5.27$) than the untrained actor ($M = 7.96$, $SD = 4.53$). The trained and untrained actor did not differ from one another in their average length of response to remaining prompt types (all $p>.15$).

9.4 Discussion

The current study examined the performance of trainee investigative interviewers who interacted with either a trained (having expertise in investigative interviews of children) or an untrained adult actor (holding no such expertise) playing the role of a child. Firstly, trainees were relatively poor at using both ground rules and closure rules. Interviewers only used two ground rules at levels greater than would be expected by chance and used four rules at levels less than would be expected by chance. Similarly, only two closure rules were used across interviews at levels greater than chance and one rule was used at levels less than would be expected by chance. During the substantive phase of the interview, and regardless of the actor's expertise, invitations and directives were used to the same extent by interviewers (consistent with findings in Chapter 7), who were just as likely to use suggestive questions

and summaries as they were to use non-substantive questions. Consistent with the first study in this thesis, trainees also introduced directives earlier into the conversation than option-posing questions, regardless of the expertise of the actor they were questioning.

The behaviour of the two actors revealed some evidence that knowledge of research on investigative interviews of children may help the actor to reinforce best-practice lines of questioning by the trainee. As predicted, the trained actor provided longer responses on average (2.7 times longer) to invitations during the pre-substantive phase of the interview than the untrained actor. Moreover, the trained actor provided greater numbers of words and details overall, and on average to invitations during the substantive phase than the untrained actor. Conversely, the untrained actor provided a greater number of details in response to option-posing questions (i.e. a focussed prompt) and a greater proportion of words to directives then the untrained actor.

Regardless of the expertise of the actor, invitations elicited the first detail from actors and fewer details were elicited from the actors before the introduction of directives versus option-posing questions. Collectively, training in the scientific literature in investigative interviews of children appears to be related to better-quality interviews for trainees on some dimensions (i.e. reinforcing the use of invitations) while both actors performed similarly on some dimensions (the question that led them to provide the first detail). These latter findings may suggest relatively good performance from the untrained actor in this instance.

281

In sum, these findings suggest that despite behaviours by the actor that may reinforce the use of best-practice lines of questioning, the interviewers did not respond to this. All effects of question type on interviewer questioning behaviour were not qualified by the expertise of the actor whom they were responding to. Interviewers' open prompts were rewarded to a greater degree during the pre-substantive phase by the trained actor, which was expected to encourage the use of open questions during the substantive phase. However, there was no effect of expertise of actor on interviewers' questioning behaviour during this phase. This suggests that interviewers were not sensitive to feedback from the actor in response to question type. Indeed, although the interviewer's line of questioning was of high quality in some domains (e.g. invitations used more than option-posing questions), consistent with Chapter 7, open prompts were used to the same extent as directives across both actors and, perhaps of concern, suggestive questions were used to the same extent as non-substantive questions. Thus, JIIT, in practice, does not appear to promote the use of open-prompts over directive lines of questioning.

The findings of this study have implications in light of prior research. Prior work suggests that interviewers who conduct mock interviews with trained respondents use a greater proportion of open questions than their peers who are interviewed by a fellow participant (Powell et al., 2008). The results of the current study may differ from Powell and colleagues as the interviewers in this study were undergoing their initial training, rather than 'refresher training' among a sample who

had conducted, on average, 54.50 field interviews with children. Moreover, interviewers in this earlier study received two practice sessions with either a trained actor or one of their peers undergoing training, before conducting their post-training assessment interview. No such practice sessions were offered for the sample here, who instead only received a single opportunity to conduct an interview, which was analysed here. Further work, with practice interviews conducted before the main mock interview (for analysis), and with data on performance in field interviews after JIIT, may shed light on these discrepancies with prior work.

Further research should aim to engage police forces in a longitudinal study and analyse field interviews of newly qualified interviewers in order to establish the long term effects of JIIT on interview respondents use of questions as higher quality questioning has been demonstrated among individuals that practiced with a trained actor during a post-training interview (Powell et al., 2008). The current study suggests that further work may be of priority in Scotland. If the current training regime (one practice interview) does not impact trainee interviewers' questioning behaviour then JIIT requires revision in order to reflect the needs of the trainees and provide them with more than one opportunity to conduct practice interviews before they progress to field interviews with real children. As was the case in the sample of interviews in Chapter 7, training is failing to promote interviewers' adherence to best practice in the pre-substantive and closure phases of training interviews. Interviewers' training must be updated to underscore the importance of

communicating the ground rules, practicing NET, and appropriately closing interviews.

During the period of data collection for the present study, the Scottish Executive (2003) guidelines were replaced by the updated 2011 guidelines. One significant change here was the mandatory requirement that all joint interviews are visually recorded. Up until this point, interviewers in Scotland had kept handwritten records of their interviews with children, which could be (and have been) used as evidence in child protection cases in courts, despite concerns raised about the merits of this practice (Lamb et al., 2000; La Rooy, et al., 2011, see also Study 1 of this thesis). This pivotal change in Scotland presented the opportunity for the next study. As the practice of visually recording child interviews it still new in Scotland, interviewers are still being trained to keep handwritten records of their interviews during training. Therefore, the next study compares visually recorded versions of JIIT interviews with their handwritten verbatim records, to examine whether interviewers are keeping faithful handwritten records of their interviews.

Chapter 10: The quality of forensic interviews conducted with children in Scotland

10.1 Study rationale and aims

As emphasised earlier, there is an urgent need for research to assess the quality of field interviews conducted with children in Scotland. The quality of child interviews has been studied in many countries worldwide (e.g. Canada, England & Wales, Israel, Norway, Sweden, and the USA), however, no study to date has examined a representative sample of field interviews conducted with children alleging abuse in Scotland. The limited field research that has analysed the quality of field interviews in Scotland (La Rooy et al., 2012, 2013) showed that interviews contained low proportions of open prompts and interviewer questioning was predominantly comprised of focussed questions. The results of the first three studies of this thesis are at odds with these findings; JIIT trainees in Scotland asked a large proportion of open prompts during their mock interviews with adult actors. As the field interviews (La Rooy et al., 2012, 2013) had been provided for the purposes of expert witness analysis in court, the samples were potentially of low quality and therefore may not be representative of the quality of child interviews in Scotland as a whole.

In an effort to ascertain whether interviewers in Scotland ask large proportions of open prompts in the field (as they do during training) or few open prompts as is the case in the field samples, a sample of 100 field interviews with children were requested for analysis from a single police jurisdiction. Despite agencies agreeing to provide this sample, problems

285

obtaining the transcripts were apparent and the police force declined to take part further in the research. However, a small sample of interviews ($n = 9$) had been obtained during an initial period of data collection and the quality of those interviews is the focus of this study.

The aims of this study were to assess the quality of interviewers' questioning and the information elicited from children. Interviewer behaviour was assessed by 1) calculating the frequency of the introductory principles communicated, 2) calculating the frequency of NET practice interviews 3) calculating the numbers and proportions of the different types of questions asked during the substantive phase, 4) calculating how soon interviewers used their first focussed prompts (directive and option-posing questions) and 5) calculating the frequency of closure principles communicated by interviewers. Based on the previous field research (La Rooy et al., 2013, 2013) it was predicted that interviewers would not use all of the introductory principles, would not conduct NET practice interviews, would ask predominantly focussed questions and few open prompts in the substantive phase and that focussed prompts would appear early in the interviews. It was also predicted that the closure principles would be infrequently communicated as interviewers in Scotland tend to ignore other best practice recommendations when conducting field interviews (La Rooy et al., 2013, 2013).

The information elicited from children and their engagement with the interview was also of interest. With regards to children's behaviour in the substantive phase, the remaining measures examined were: 1) the

286

proportion of 'informative' responses elicited by different question types, 2) the total number and average level of detail provided in response to the different question types, 3) the type of question eliciting the first substantive detail, 4) the proportions of details provided before focussed prompts (directive and option-posing) were asked and 5) the total number and average number of words provided in response to the different question types by children. As interviewers' questioning was expected to be poor, it was predicted that that the greatest proportions of details and words would be elicited by focussed prompts because they were expected to comprise the bulk of the questioning. As it was anticipated that interviewers would not delay their use of focussed questioning, it was expected that few details would be elicited from children before focussed prompts were asked. With regards to the type of question that would elicit the first substantive detail and the proportion of responses that would be 'informative' to the different question types, no specific hypotheses were made as these measures have not been the focus of previous research with children.

10.2 Method

10.2.1 Participants.
Interviews were conducted jointly by a police officer and social worker, with the majority of interviews (N=8) lead by the police officer. The social worker led the remaining interview. Three lead interviewers were male and six were female. Eight interviews were conducted on female children and one interview was conducted on a male child. Children (between 3.92 and 14.67 years of age) made various

allegations including allegations of sexual touching (N=5), allegations of some form of penetration (N=2) and one allegation of lewd and libidinous conduct. Alleged perpetrators ranged from someone familiar to the child (N=7), an immediate family member (one case) and a stranger (one case). Three perpetrators were adults, five perpetrators were children and one perpetrator was of unknown age. Analyses are of interviews conducted in 2007 (one interview), 2008 (two interviews), 2009 (one interview) and 2010 (five interviews).

10.2.2 Procedure, materials, analyses and coding of transcripts

Interviews were conducted under the "Guidance on Interviewing Child Witnesses in Scotland" published by the Scottish Executive (2003). These guidelines require that forensic interviews are conducted jointly both by social workers and police officers. Interviewers had access to VRI and interviews were audio recorded. Transcribing was identical to Chapter 7, with each interview taking on average 6.5 hours to transcribe. Interviews were coded in an identical manner to Chapter 7, except that summaries and non-substantive utterances were grouped into a separate category ('other'). Consistent with Chapter 8, analyses of behaviour during the pre-substantive phase of the interviews were also conducted. Reliability coding (same procedure as Chapter 7) of transcripts revealed that reliability in identifying question types (K=.84) and identifying child details (K=.79) were both high. For analyses, identical variables were examined to Chapter 7 (except that training force was not a relevant factor in this dataset).

10.3 Results

10.3.1 Interviewer behaviour (pre-substantive phase). Analyses of interviewer behaviour during this phase revealed that nine of the ground rules were used significantly less than would be expected by chance across interviews. None of the ground rules were used at levels greater than would be expected by chance (see Table 10.1).

Table 10.1: Presence of ground rules at pre-substantive phase

Rule	Mean proportion usage (SEM)	Binomial test result
Listen	0.44 (.18)	$p=1.00$
Knowledge	0.11 (.11)	$p=.039$
Don't know	0.22 (.15)	$p=.18$
Don't know demonstration	0 (0)	$p<.01$
Don't understand	.44 (.18)	$p=1.00$
Don't understand demonstration	0 (0)	$p<.01$
Correct me	.22 (.15)	$p=.18$
Correct me demonstration	.11 (.11)	$p=.039$
Repeated questions	0 (0)	$p<.01$
Don't remember	0 (0)	$p<.01$
Don't guess	.11 (.11)	$p=.039$
Truth	.67 (.17)	$p=.51$
NET	0 (0)	$p<.01$
NET Open	0 (0)	$p<.01$

10.3.2 Interviewer behaviour: Substantive phase. A repeated measures ANOVA was conducted on the outcome variable proportion of total interviewer utterances, with the within-subjects factor question type (invitation, directives, option-posing, suggestive, other). Greenhouse-Geisser corrections were used as Sphericity was violated according to Mauchly's test (p<.001). This analysis revealed a main effect of question

289

type ($F(1.41, 11.26)=14.52$; $p=.001$, $np^2=.65$). Post-hoc analyses to interpret the main effect of question type demonstrated that a smaller proportion of invitations were used than directives, option-posing and other questions. A greater proportion of other questions were asked than suggestive questions (all non-significant Bonferroni-corrected p values >.12, see Figure 10.1).
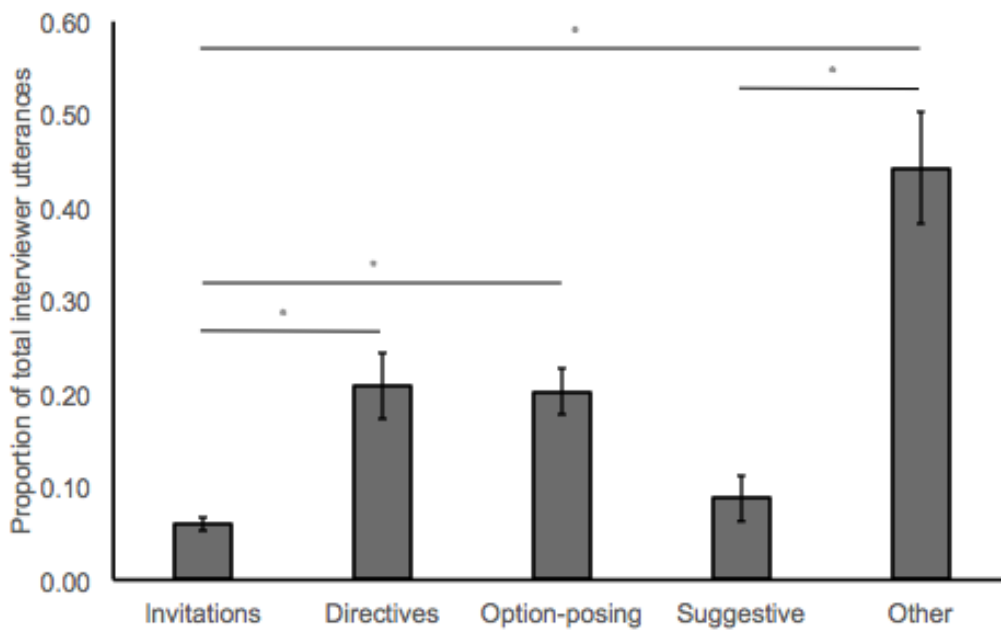
Figure 10.1: Proportion of total interviewer utterances for each question type during the substantive phase of the interview. Error bars show standard error of the mean.

Next, an ANOVA was run on the outcome variable proportion of questions used before the introduction of focussed prompts, with the within subjects' factor *question type* (option-posing, directive). This analysis revealed a main effect of *focussed prompt* ($F(1,8)=5.58$; $p<.05$, $np^2=.41$). A paired samples t test was used to interpret the main effect of focussed prompt that revealed interviewers introduced directives earlier

into the conversation ($M=0.037_{questions}$, $SEM=.009$) than they introduced option-posing questions ($M=0.056_{questions}$, $SEM=.01$, $t(8)=2.36$; $p<.05$).

10.3.3 Presence of closure principles. Analyses of the use of closure principles revealed that three of the principles ('next', 'thanks' and 'contact') were used less than would be expected by chance. None of the principles were used at levels greater than would be expected by chance (see Table 10.2).

Table 10.2: Presence of closure principles across interviews (N=9)

Rule	Mean proportion usage (SEM)	Binomial test result
Interviewer two questions	0.67 (.17)	$p=.51$
Summary	0.22 (.15)	$p=.18$
Child questions	0.44 (.18)	$p=1.00$
Neutral	0.33 (.17)	$p=.51$
Next	0.11 (.11)	$p=.039$
Thanks	0 (0)	$p<.01$
Contact	0 (0)	$p<.01$

10.3.4 Child behaviour (pre-substantive phase). Due to the absence of any practice interviews, the average number of words per response by children to interviewer questions during the pre-substantive phase was calculated for each question type. Descriptive statistics show that invitations tended to elicit the longest response from children ($M=15.61$, $SD=12.92$), followed by directives ($M=8.41$, $SD=4.62$), option-posing questions ($M=5.68$, $SD=3.67$), suggestive questions ($M=3.44$, $SD=2.48$) and introductory comments ($M=2.38$, $SD=2.97$).

10.3.5 Child behaviour (substantive phase). A repeated measures ANOVA was conducted on the outcome variable proportion of informative responses, with the within subjects' factor question type (Invitation, directive, option-posing, suggestive). Greenhouse-Geisser values were used as the assumption of Sphericity was violated. This analysis revealed no main effect of question type ($F(3, 24)=2.16$; $p=.149$, $np^2=.21$).

10.3.6 Richness of responses. A repeated measures ANOVA was conducted on the outcome variable proportion of details elicited from children, with the within subjects factor question type (Invitation, directive, option posing, suggestive, other). This analysis revealed a main effect of question type ($F(4,32)=4.16$; $p<.01$, $np^2=.34$). Bonferroni-corrected pairwise comparisons revealed that no specific question type elicited more details from children than another question type (all Bonferroni-corrected $p>.11$, see Table 10.3).

An identical ANOVA on the outcome variable average details per prompt type (corrected for violating the assumption of Sphericity) revealed a main effect of question type ($F(1.13, 9.01)=5.51$; $p=.041$, $np^2=.41$). Bonferroni-corrected pairwise comparisons revealed that option-posing questions elicited more details per response from children than other questions ($p=.01$), and suggestive questions elicited more details per response from children than other questions ($p<.01$). No other comparisons were significant (all other Bonferroni-corrected $p>.12$, see Figure 10.2).

Table 10.3: Descriptive statistics for richness of child responses

Question type	Mean details (SEM)	Mean proportion details (SEM)
Invitations	44.89 (14.67)	.24 (.06)
Directives	75.56 (22.71)	.36 (.06)
Option-posing	33.33 (8.37)	.18 (.03)
Suggestive	14.11 (3.48)	.11 (.04)
Other	17.67 (5.83)	.11 (.03)
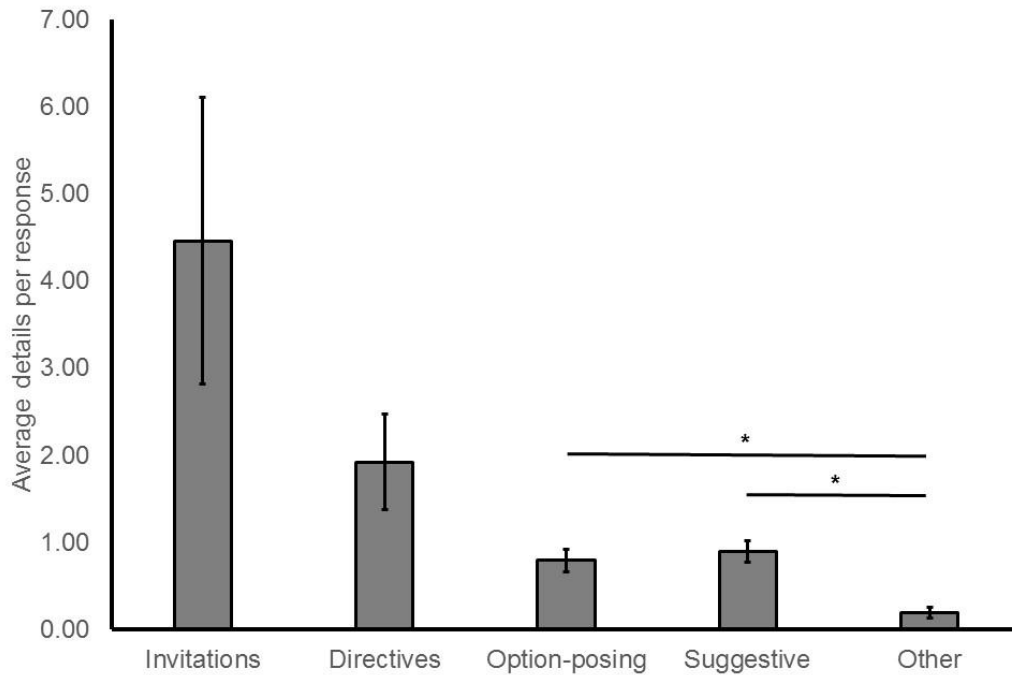
Figure 10.2: Main effect of question type on average number of details per response from children.

Descriptive statistics (on a small sample of interviews) suggest that suggestive questions elicited the first detail from children (in 4 interviews), followed by invitations (in 3 interviews) and directives (in 2 interviews). Moreover, only 7.63% of overall details (M=19.33 details, SD=45.04 details) were elicited from children before a directive question was introduced and 12.73% of overall details (M=30.00 details, SD=49.40 details) were elicited before an option-posing question was introduced.

293

10.3.7 Length of response. A total of 12,532 words were elicited from children (M=1932.44, SD=1210.24). A within-subjects ANOVA was conducted on the outcome variable proportion of total words elicited, with the within subjects factor question type (invitations, directives, option-posing, suggestive, other). Greenhouse-Geisser corrections were used as the assumption of Sphericity was violated according to Mauchly's test. This analysis revealed a main effect of question type (F(1.51,12.04)=4.49; p=.043, np^2=.36).Bonferroni-corrected pairwise comparisons revealed no significant comparisons (all Bonferroni-corrected p>.24) except that 'other' questions elicited a greater proportion of total words from children than suggestive questions did (p=.037, see Table 10.4).

Table 10.4: Proportion of total words elicited from children per question type

Question type	Mean proportion of total words (SEM)
Invitations	.14 (.02)
Directives	.29 (.06)
Option-posing	.18 (.02)
Suggestive	.09 (.03)
Other	.31 (.06)

Finally, an ANOVA was conducted on the outcome variable average words elicited per response, with the within subjects factor *question type* (invitations, directives, option-posing, suggestive, other), corrected for violating the assumption of Sphericity. This analysis revealed a main effect of *question type* (F(1.18,9.45)=11.70; p<.01, np^2=.59). Bonferroni-corrected pairwise comparisons revealed that

invitations elicited a longer response on average from children than directives (p=.049). No other comparisons were significant (all Bonferroni-corrected p>.058, see Figure 10.3).
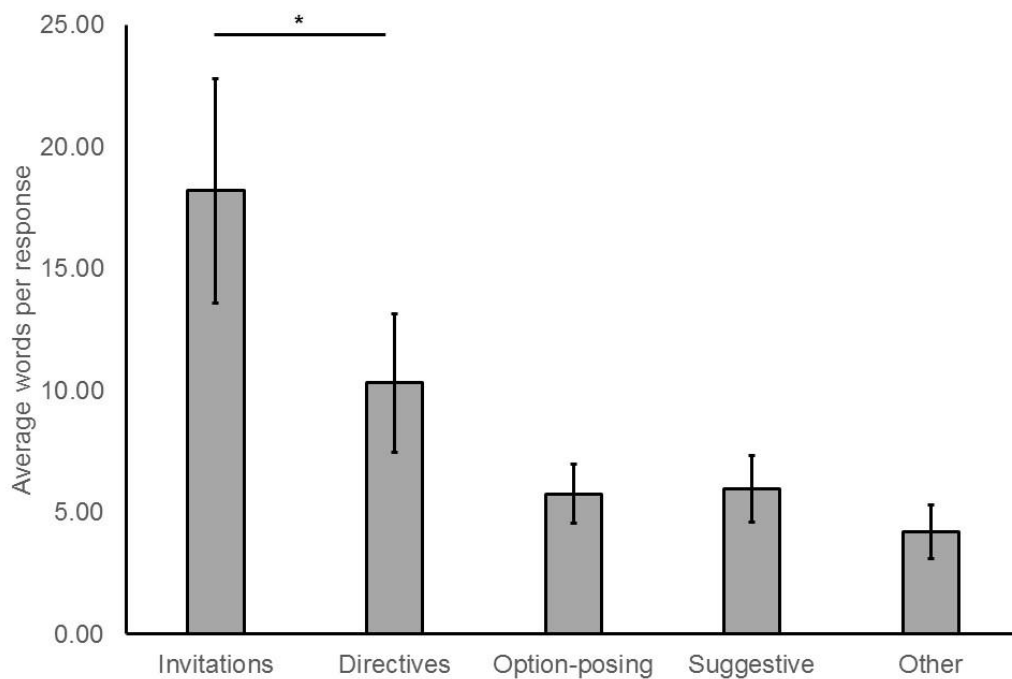
Figure 10.3: Main effect of question type on average number of words per response from children.

10.4 Discussion

Only a single ground rule, the instruction to tell the truth, was present in more than half of the interviews in this study. None of the ground rules were used at levels greater than would be expected by chance and, conversely, nine of the ground rules were used significantly less than would be expected by chance across interviews. Of concern, six ground rules were never communicated in a single interview. In other words, one quarter of the ground rules listed in the Scottish Executive

(2003) Guidelines were not present in any of the field interviews conducted in this study.

No NET (practice interviews) was attempted in this sample of interviews. As this was also found to be the case in both of the small field samples obtained by La Rooy et al. (2012, 2013), this finding is of concern but consistent with field research. As these interviews were conducted between 2007 and 2010 under the Scottish Executive (2003) guidelines, which included a sample protocol based on the NICHD Protocol, these findings cannot be attributed to lack of awareness among interviews that a practice interview is one of the fundamental preparatory procedures that should always be done. Children in Scotland are not adequately prepared for an interview of this nature, as they are being denied an opportunity to practice responding to open prompts before discussing substantive issues. The few open prompts that were provided to children in this early phase of the interview elicited the longest average responses compared to all other pre-substantive question types, as is the case when open prompts are asked in the pre-substantive (Price, Ahern & Lamb, 2016) and the substantive phases of interviews with children (Lamb et al., 1996a; Lamb et al., 2007a, 2007b;Orbach& Lamb, 2000, 2001;Orbach et al., 2000;Sternberg et al., 2001a). Of note, this had no bearing on interviewers' behaviour as the substantive phase was dominated by focussed prompts, although they elicited shorter answers from children early on.

In the substantive phase of the interviews, a substantial proportion of interviewers' questioning was off-topic, with 'other' questions

comprising 44% of the utterances in this phase. As the substantive phase begins once the transition question is asked, this means that predominantly non-substantive questions, that were unrelated to the incident, were being asked once discussion of substantive events had already begun. These off topic discussions were interviewer driven and therefore reflect interviewers' biases as they dictated the direction of questioning instead of encouraging children to continue with their narrative of events. For example, in response to the transition question, one child provided a lengthy narrative about why they thought they were being interviewed - they provided an allegation that someone had been staying at their house and touched them in a sexual manner during the night. Instead of eliciting further narrative about the incident the interviewer spent a considerable amount of time asking the child to describe the layout of the entire house and the room that it happened in. As the location had already been confirmed in the child's narrative and it was a familiar location (and therefore not a location that the police needed information about in order to identify it) it is unnecessary to ask the child to describe this sort of information. If such confirmation of details were required then this could have been asked at the end of the questioning phase once the free narrative had been exhausted. Instead, interviewers risk fatiguing children by requesting such specific information at the start of the substantive phase before they have the chance to discuss the allegation itself.

As predicted, the remaining substantive utterances (that focussed on allegations related topics) consisted of mainly directive and option-

297

posing questions (21% and 20% respectively). Interviewers also moved to asking focussed questions very early on in the interviews. Only 5.89 utterances were asked (only 7.63% of overall details elicited) before the first directive question and 8.78 utterances were asked (12.73% of overall details elicited) before an option-posing question was introduced. The majority of children's first details were elicited using a suggestive question (suggestive questions elicited the first detail from children in four interviews, followed by invitations in three interviews and directives in two interviews). This was not due to reluctance on the part of the children, as they responded informatively to the majority of questions asked by interviewers (range 84-91%) and there was no difference in informative responding to any question type. Instead, this is likely to be the result of interviewers' transitioning into the substantive phase of their interviews in a suggestive manner, for example, by naming the suspect ("I understand you're here today to talk about something that happened with C last week…"). This means that interviewers are not following best practice recommendations and waiting until as late as possible in the interview after exhausting free narrative before asking focussed questions and instead are contaminating children's accounts early on, in some cases before they have had the chance to provide their first detail.

Open prompts accounted for only 6% of substantive questions and were significantly less likely to be asked than directive, option-posing and 'other' questions. Even though the proportion of open prompts used in the interviews was extremely low, this is not unusual in practice around the word. Indeed, this figure is comparable to the number of open

prompts found in other field studies, with typical values, for example, such as two per cent (Korkman et al., 2006; Lamb et al., 1996b), 4.7% (Sternberg et al., 1996) and six per cent (Cederborg et al., 2000). Directive questions elicited the largest number of details and largest proportion of words from children in the interviews. This is not surprising because out of all the questions that requested allegation related details from children (not including 'other' questions as these were non-substantive or information summaries), directives comprised the largest proportion. Overall numbers of details and words elicited did not differ significantly according to question type, however, despite being the least-used used question type, invitations elicited the second largest number and proportion of details from children with almost a quarter of total details elicited provided in response to invitations.

Closure principles were infrequently communicated. Offering the second interviewer an opportunity to ask the child any questions was the most often communicated closure principle and the only principle to appear in more than half of the interviews. Analyses of the use of closure principles revealed that three of the principles ('next', 'thanks' and 'contact') were used less than would be expected by chance ('thanks' and 'contact' were not used by any interviewers in this sample), and none of the principles were used at levels greater than would be expected by chance. Neutral closure appeared 33% of the time despite its recommendation in the guidelines that this should always be done whether a disclosure is elicited or not.

299

Finally, of the nine interviews in the sample, two interviewers used body diagrams to elicit information from children and five interviewers requested the child draw pictures of aspects of allegation related information (e.g. of the room the incident took place in). The guidelines state that body diagrams may be used to clarify information the child has already mentioned, however, the diagrams were used during attempts to elicit such information. Further, no training in the use of body diagrams is provided on JIIT courses and therefore we do not know how the interviewers were trained to use such aids. With regards to asking children to draw things they were describing, it was unclear why this was necessary and in some cases children actually refused to do so.

Collectively, these findings suggest that interviews with children in Scotland are of an extremely low quality. Ground rules and closure principles were never used at levels greater than chance. During substantive questioning directive prompts were the information seeking prompt type most frequently used by interviewers, followed closely by option-posing questions. The largest proportion of questions used were 'other' questions. These questions did not request allegation-related information from children and therefore the majority of the interview time was not spent on topic. Open prompts, which are the best practice questions, comprised the smallest proportion of questions in the interviews (6%), which is comparable to the proportions of open prompts found in field interviews in other countries. Even during the pre-substantive phase invitations elicited the longest average responses from children than any other question type, yet this still did not prompt

interviewers to maintain their use. Instead, interviewers appeared insensitive to the differences in children's responses and persisted with focussed questioning.

Interviewers also behaved in ways that could contaminate children's evidence. Interviewers used focussed questions very early on in the interviews, were most likely to elicit children's first details using suggestive questions, failed to let children provide free narrative before asking focused questions, persisted with requesting information from children when they provided expressions of ignorance and introduced body diagrams as a way of eliciting information. The guidelines that the interviewers in this sample followed contained an interview protocol based on the NICHD protocol, which has been shown to improve interviewers' questioning (Cyr & Lamb, 2009; Lamb et al., 2009; Orbach et al., 2000; Sternberg et al., 2001b), this study provides further evidence that providing interviewers with a protocol alone does not improve the quality of their interviews and instead interviewers must receive specialised training in its use if they are to be able to benefit from a protocol.

Although the predicted effects of using open questions (longer, more detailed responses from children) were not apparent in this study it is important to note that this is likely due to the methodological limitations of the study. The sample was small (9 children in total) and the spread of age was uneven; the sample consisted of two 3-to 4-years-olds, three 5-to 6-year-olds, one 10-year-old and three 13-to 14-year-olds.Previous research has shown that the youngest children, (3-6 years of age)

301

provide shorter and less detailed responses to open prompts than older children and that the ability to provide information to open prompts increases with age (Brown & Lamb, 2015; Gagnon & Cyr, 2017; Lamb et al., 2003). Just over half of the sample was comprised of very young children that may not yet have developed the ability to provide narratives to open prompts. Unfortunately there were not enough participants to compare the differences in responding between younger and older children but this is an avenue for future research in Scotland.

Chapter 11: General discussion

This thesis is the first research to provide in-depth evaluation of the effectiveness of JIIT courses in Scotland on trainee interviewers' ability to conduct appropriate forensic interviews in a role-play context with adult actors. Previous research that has evaluated trainee interviewers' post-training interviewing skills have usually evaluated their questioning behaviour with real children following training (e.g. Aldridge & Cameron, 1999; Craig et al, 1999; Warren et al, 1999). Those that have examined post-training questioning behaviour using adult actor role-play interviews have not provided quantified the different types of questions interviewers asked, instead providing either a general comment with regards to the quality of questioning (Geotzold, 2015; Stevenson et al., 1992) or have only been looked at specific facets of questioning (Powell et al., 2008). This programme of research provided in depth analyses of the entire training interviews from start to finish conducted by trainee investigative interviewers as if they were field interviews. As these interviews are the interviewers' first attempt at conducting an interview following training and the only practice interviewers in Scotland will have before they conduct real forensic interviews with children they provide insight into whether or not JIIT has adequately prepared interviewers to conduct appropriate joint investigative interviews. (Please see Appendix 1B for the trainee interviewers' perceptions of their preparedness to conduct interviews as a result of their JIIT experience). Finally, the specific responses made by untrained adult actors role-playing children for the purpose of training

303

investigative interviewers has never been subject to previous research, untrained adult actors responses were compared across different actors and against how real children respond in forensic interviews.

11.1 Overview of findings

The inclusion of 2 separate jurisdictions in Study 1 provided evidence that training is not standardised across Scotland, a sentiment shared by the Scottish Courts and Tribunals Service who have previously stated "There are differing approaches in different parts of the country" (see 'The Evidence and Procedure Review', pp.23, Feb 2016). Force 2 interviewers were more likely to communicate 3 ground rules and 3 closure principles than Force 1 interviewers. Conversely, Force 2 interviewers were more likely to conduct a practice interview (NET) than Force 1 interviewers. This study also demonstrated that across Scotland some aspects of JIIT are lacking. Disappointingly, across forces the majority of introductory rules and closure principles were used less than would be expected by chance. Force 2 interviewers were severely disadvantaged by a lack of VR facilities, asking 43% fewer questions overall and eliciting fewer details from actors before moving to asking focussed questions, compared to Force 1 interviewers who did not have to rely on a scribe to handwrite their interview. The substantive questioning of trainee interviewers was comprised predominantly of free-recall and cued-recall questions, however, interviewers did not use these questions to a different extent. While interviewers' training has conveyed the importance of relying on recall prompts as opposed to recognition

prompts, it has not managed to underscore the overall superiority of open prompts. Further, 1 in 10 questions asked was of a suggestive nature.

Adult actors' responses showed a similar lack of awareness; there was no difference between the proportions of details and words they provided in response to invitations as they did to directive questions. Worryingly, invitations elicited the greatest proportion of uninformative responses (i.e. no information was provided), that is, directive, option-posing and suggestive questions elicited greater proportions of informative responses than the best practice questions. Actors in both forces gave longer and more detailed average responses to invitations than option-posing questions but responded equally to invitations and directives. Although actors at different forces didn't differ in the average length of responses, the actors at force 1 provided richer responses in general. This could reflect the scenarios that actors role-played, another aspect of the training that was not standardised across forces, because force 2 scenarios often contained instruction for the actor to be difficult or resistant to interviewers' questioning. In order to best prepare child interviewers to conduct interviews, actors should reinforce all aspects of best practice. This study demonstrates that regardless of force they were working with, actors did not encourage interviewers consistently when they used open prompts. Actors have no knowledge of how real children respond in interviews because they are not trained for this role, so this is unsurprising.

Study 2 implemented a 'trained' actor that had been instructed to respond favourably to open prompts compared to all other question types in an effort to increase the perception of their value and their use by interviewers. During the pre-substantive phase, the trained actor was successful in providing longer responses selectively to open prompts compared to the untrained actor (there were no differences in response length to any other question type). During the substantive phase, the trained actor provided a greater proportion of details and words, and longer and richer average responses to open prompts, compared to the untrained actor. Conversely, the untrained actor provided a greater proportion of details in response to focussed questions (option posing questions) and a greater proportion of words to directives than the trained' actor. Trained actors are useful for reinforcing best practice while untrained actors may unknowingly reinforce the use of riskier more focussed question types. Training in scientific literature can promote a better experience for JIIT trainees on some dimensions. Despite interviewers being rewarded to a greater extent by the trained actor when they asked open questions, actor condition had no effect on interviewers' questioning behaviour. Interviewers did not differ in the proportions of open prompts they asked or how soon they asked focussed questions regardless of which actor they interviewed, demonstrating trainees are not sensitive to feedback from trained actors. As was the case in the previous study, interviewers did not differ in the proportion of open questions and directives they asked in their interviews.

306

Study 3 compared the handwritten verbatim interview records kept by the second interviewer (as was the standard procedure for recording interviews with children in some jurisdictions at that time) against the visually recorded DVD versions of the same interviews. Eleven per cent fewer ground rules and 14% fewer closure principles were present in the handwritten interviews than the recorded interviews. In the substantive phases of the verbatim interviews significantly fewer invitations, directives, introductory comments, summaries and facilitators were present than in their recorded counterparts. Introductory comments, summaries and facilitators were more likely to be omitted completely from the handwritten account than other types of question while invitations were more likely than other question types to be both incorrectly written down (in a manner that changed their question code) and incompletely written down. Scribes were accurate in noting down option-posing and suggestive questions.

Handwritten accounts contained fewer details than recorded interviews but any omissions or additions when scribing details provided by the actor were not specific to question type. Handwritten accounts also contained fewer words than the recorded versions, specifically fewer words from invitations, directives, option-posing questions and introductory comments. Words that had been elicited by invitations and directives were the only question types more likely than any others to be omitted, while any added words were not specific to question type. While the handwritten accounts differed from the recorded versions in that focussed prompts appeared later, both formats were accurate in

307

capturing the proportion of details elicited by the actor before the introduction of focused prompts. These findings taken together suggest that scribes were most accurate when noting down focussed questions (option-posing and suggestive) and their responses compared to recall prompts and their responses. This may tempt interviewers to ask recognition questions ahead of recall prompts in an effort to ensure an accurate interview record. However, open prompts are more likely to elicit a greater amount of information and more accurate information from children than other question types. It also appears that interviewers were better at faithfully recording details rather than interviewer questions and words. Cauchi and Powell (2010) have previously suggested interviewers are more concerned with detail recording than question recording.

Study 4 analysed a small sample of field interviews conducted with allegedly abused children in Scotland under the guidance of the Scottish executive (2003) guidelines. The findings showed that during the pre-substantive phase interviewers did not consistently communicate the ground rules to the children, they also failed to give children an opportunity to practice recalling a real life event, no child received a practice interview. How often interviewers communicated the closure principles also varied greatly. Interviewers spent the majority of their substantive questioning time exploring peripheral and non-substantive topics rather than discussing incident-related topics. When discussing allegation-related topics, interviewers probed children using predominantly focussed questions. Open prompts accounted for only 6% of the total substantive questions, despite the fact that when open

prompts were asked they elicited the longest average responses of all question types from children. Even during the pre-substantive phase, open prompts elicited the longest average responses from children, yet this has no bearing on interviewers' subsequent questioning (as was shown to be the case for the interviewers in study 2). Lastly, the majority of interviews also included some form of interviewer aid, such as a body diagram, or requests for children to draw aspects of the allegation, techniques that are not recommended as they can elicit inaccurate information.

Both the training and the field interviews provide evidence to support some of the conclusions the limited field research conducted in Scotland (La Rooy et al., 2012; 2013) has drawn. Firstly, the introductory and closure phases of interviews are not being conducted appropriately by investigative interviewers. Delivery of the ground rules and conducting appropriate practice interviews that provide narrative elaboration practice were not consistently featured in either format of interview. It is possible that interviewers and interview trainers may feel that the one (or two in some cases) role-play interviews that trainees conduct before interviewing real children are better spent focussing on substantive questioning and knowledge alone of best practice aspects will ensure their use in the field. Instead, the opposite has not been shown to be the case, not a single practice interview (NET) was present in any either of the field interviews obtained to date in Scotland (La Rooy et al., 2012; 2013) or the small sample of interviews analysed in Study 4. Failure to practice these elements during training potentially depreciates their value

309

to trainee interviewers making them less likely to include them in their future field interviews. Not conducting appropriate practice interviews may also serve to perpetuate the beliefs that interviewers in Scotland have expressed that open prompts are not effective ways of obtaining information from children (La Rooy et al., 2011). If children that have not practiced responding to open prompts before being them in relation to allegation(s) they may not be aware they are being prompted to provide lengthy narratives because they are unfamiliar with being prompted in such a way. Further, the 3 training studies and the field sample converge on the most consistently used ground rules ("Truth", "Don't understand", "Don't know", "Listen" and "Don't guess") showing that training interviews are potentially a good measure of how interviewers will behave in the field.

Interviewer behaviour with regards to the introduction of focussed prompts was similar in the training and field interviews, regardless of setting interviewers asked directive question earlier than option-posing questions. Unfortunately, interviewers did not delay their use of focussed questions until as late as possible in either the field or the training interviews. During training, when the amount of time interviewers have to conduct the interview is limited, interviewers may be eager to elicit details that actors have not yet mentioned because they may worry that they must obtain such information in the limited time they have. This may promote bad habits in the field because asking focussed questions early on means that interviewees have not been given the opportunity to tell

everything they know before being asked riskier prompts, this has the potential to contaminate interviewees accounts from very early on.

The most striking difference between the training and the field interviews regards what is arguably the most important aspect of eliciting high quality information from children; interviewers' use of open prompting. Open prompts comprised the smallest proportion (6%) of all question types in the field interviews, infrequent use of open prompts in the field interviews is not surprising given the similar findings in samples of forensic interviews conducted in many countries worldwide. It is surprising however that during training open prompts accounted for 34-39% of interviewers' questions, making them the largest proportion of questions compared to all other question types. These contrasting findings could reflect either field interviewers' gains made during training had dissipated or during training interviewers' adhere better to the guidelines, because they are aware they are being observed. Even though interviewers used a high proportion of open prompts during training it is important to note that they still did not differ significantly between the proportions of invitations and directives they used.

An unexpected finding from study 2 was the insensitivity that interviewers displayed to interviewees' differential responding to question types. It could be argued that no differences were found in the proportions of open questions asked to the trained and the untrained actor because interviewers know they should ask open questions and do so regardless of who they are interviewing. However, the sample of field

311

interviews showed that in the field interviewers chose to ask predominantly focussed prompts during substantive questioning even though children provided longer and more detailed responses to open prompts during the pre-substantive phase. Taken together this indicates that interviewers are probably not sensitive to the length and level of detail provided by interviewees and instead follow their own line of questioning in an effort to obtain the information that they deem to be important.

There was variation in the responding of adult actors used in Scotland, which is unsurprising given that they are given no instruction in how to play these roles, however, some patterns of similarity with how real children respond emerged. When the average number of details per response was examined, both adult actors and children provided richer and lengthier responses on average to invitations then to option posing questions. However, unlike children, who provided longer responses on average to invitations than they did to directive questions adult actors' responses to open prompts and cued prompts did not differ. While providing longer and more detailed responses to recall over recognition prompts was natural, responding differentially to open prompts and directives was not. It was possible to train an adult actor to provide longer and more detailed responses to open prompts compared to directives and therefore it should not be an arduous task to develop training for all actors that participate in JIIT. Adult actors were highly regarded and with training could be a high quality resource for interviewers.

11.2 Implications

This thesis predominately focussed on investigative interviewers' training interviews, interviewers under observation may not behave the way that they really would in the field, therefore, the results of the training studies may not be generalisable to field interviewing behaviour. This research showed JIIT trainees that had opportunities to practice interviewing adult actors were competent during training to use and maintain open prompting. The reason interviewers can conduct high quality interviews during training but not in the field may be in fact be because interviewers under observation may not be behaving the way that they really would in the field and that interviewers are not 'sticking to the rules' when not under direct observation. It is however possible that post-training gains dissipate, the standard of training in Scotland could potentially have improved since the sample of field interviews was obtained and that the trainees in this study will go on to conduct high quality field interviews. We did not analyse post-training field interviews conducted by these trainees and follow-up research is needed to analyse the field interviews to evidence whether training gains dissipated and whether interviewers that conduct high quality training interviews can maintain these standards in the field.

Due to the actor variation found within this country we do not know if actors across the country are all responding in the same manner as the actors included in this research. The interview trainers at the force that participated in Studies 1-3 have all attended CPD courses by an international expert in child interviewing and were very aware of the

313

NICHD protocol and therefore possibly impacted actors' behaviour. We would need to sample all actors working country wide in an effort to determine whether it is the case that all actors participating in all JIIT courses are responding similarly.

A limitation of this programme of research is related to the sample sizes in these studies. Due to the way interviewers' training is conducted in this country there are limited places for participants on each JIIT course. The jurisdiction that participated in studies 1-3 train 16 participants per JIIT course. The field sample was small and had limited power to find differences, however, a larger sample of interviews was requested and instead issues related to accessing field data prevented this from being obtained. This is an ongoing issue in Scotland and future research in this country with larger and up to date samples of interviews is necessary in order to obtain an accurate picture of the field interviewing situation.

In study 3 the actual numbers of omitted, incorrect and incomplete utterances found are small, however, these findings were produced during training when interviewers are most likely to stick to the rules as they are still under observation. In the field when interviewers are no longer being supervised there is the potential for even greater numbers of questions to be omitted, incorrectly recorded and incompletely recorded. This has been shown to be the case in field studies that have looked at verbatim interview records (Lamb et al., 2000). Further it is also important to note that because these are training interviews the scribe was likely

better supported to record the interview verbatim than is the case in a real forensic interview with a child. The adult actors can purposefully help the scribe by speaking slower and leaving pauses for the scribe to catch up. The lead interviewers in this study also used a lot of summaries and non-substantive questions which allowed the scribe catch up. In a forensic interview real children may not be sensitive to the need to speak slower and the lead interview may not be able to initiate as many pauses. Therefore, the amount of information lost in training interviews may actually underestimate the amount of information that would be lost from a real interview with a child.

Finally, It is important to note that the trained actor in this study is also the author of this thesis. Any potential limitations of having the author involved in this stage of the research were controlled for by ensuring a random sample of the data were coded by a second coder to ensure reliability in the coding.

11.3 Future directions and recommendations

In order to enhance the training experience for trainee interviewers and in turn the quality of their later field interviews the following recommendations are put forward:

1) *The Scottish Executive (2011) guidelines need to be updated and the sample protocol reinstated.*

The current guidelines need revision because there are areas that are unclear and that contradict best practice research. They recommended the ground rules are dispersed throughout the interview claiming they are

"more effective" this way and that the "the litany approach is, in fact, counter-productive and unnecessary" (p. 27). As discussed in Chapter 4, the paucity of research on this topic means the best way to deliver the ground rules is yet to be agreed on and the guidelines provide no instruction on how to achieve a dispersed delivery approach. It is still recommend that interviewers check children have understood the ground rules by using demonstrations, however, unlike the 2003 guidelines, the updated version does not provide any examples of such demonstrations.

The 2011 guidelines advocate the use of a practice interview 'only where appropriate', for example, if the child appears they want to discuss the allegation straight away then it is recommended that the interviewer should skip practice and proceed directly to substantive questioning. While there is no common consensus on what the appropriate course of action should be if this situation arises, it fails to take into consideration the value of any alternative strategies such as confirming with the child that they will discuss substantive issues soon, but would like to get to know them a bit better first. Recommended topics for the practice interview include 'television programmes', this is incongruous with the research that advocates avoiding any discussion of fantasy (television, computer games, anything 'made up') (e.g., Lamb et al., 2008). It is emphasised at the start of the interview that the child is expected to tell the truth, and to tell the interviewer only about things that have really happened. Discussing fantasy, especially with younger children, may confuse and undermine the point of the interview.

The terminology and definition of question types is also inconsistent, for example, interviewers are advised to use 'open-ended questions' in the practice interview, yet given directive questions ("Who, What, Where, When, How") as examples. Later in the guidelines, when discussing eliciting information in the substantive phase, "Tell me more" is described as an 'open-ended prompt' and 'wh-'questions now deemed to be 'specific questions'. Despite heavily emphasising open prompting throughout the majority of the guidelines, when attention is turned to eliciting accounts of repeated abusive experiences the guidelines now advocate that specific questions are recommended to explore each incident separately.

The Scottish Executive (2003) Guidelines included a structured interview protocol, a shortened version of the NICHD protocol, for interviewers to follow (see Appendix A). In 2011 when the Scottish Executive updated their national guidelines on conducting child interviews this protocol was removed. This has not gone unnoticed with the Scottish courts and tribunals' service that expressed "the current guidelines for interviewing children, which were issued in 2011, should be revised explicitly to require the use of the NICHD protocol." (pp. 29, The Evidence and Procedure Review, 2016). The shortened protocol guides interviewers through conducting an interview from start to finish and provides an example of an appropriate practice interview about an experienced event using open prompts and time segmentation.

2) JIIT training courses should be updated according to investigative interviewers' training research recommendations

Previous research has shown that introducing a structured interview protocol alone has not been effective in improving the quality of investigative interviewers' field interviews (e.g. Sternberg, Lamb, Davies & Westcott, 2001). Instead extensive training on using the NICHD protocol that includes multiple practice opportunities, feedback on interviews, exercises that promote a deeper understanding of question type, close supervision and refresher training will be needed in order to support interviewers effectively to conduct high quality interviews with children. Short intensive training courses have not been able to impact interviewers long-term interviewing behaviour (e.g. Aldridge & Cameron, 1999; Warren et al., 1999), instead spaced learning opportunities where interviewers receive refresher sessions have been shown to be effective in improving the quality of interviews, as post-refresher training is when interviewers conducted higher quality interviews than immediately post-training (Rischke, Roberts & Price, 2010). The quality of interviews conducted by interviewers that received on-going supervision and interview feedback were shown to have improved (Lamb, Sternberg, Orbach, Hershkowitz, Horowitz & Esplin, 2002) while the quality of interviews declined upon termination of support (Lamb, Sternberg, Orbach, Esplin & Mitchell, 2002). JIIT courses need to be updated to incorporate these research recommendation, this thesis has shown that training is not standardised across Scotland, therefore, this update needs to be on a nationwide scale.

318

3) Training should be provided for the interview trainers

The results of the three training studies in this thesis have shown that while trainee interviewers are being adequately supported to conduct appropriate substantive questioning phases, other phases of the interviews are not being conducted appropriately during training. For example, the interviewers are not routinely conducting practice interviews that contain narrative elaboration practice, not being expected to do so during training may give the impression this is not an important aspect of the overall interview and it has been shown that in the field interviewers go on to neglect to conduct practice interviews (La Rooy et al, 2011; 2012; 2013; Study 4 of this thesis). The lack of practice interviews conducted during training implies that interviewer trainers are not insisting that trainees conduct them, as is the case for some other aspects of their role-play interviews. Coupled with the studies in this thesis showing that training is not standardised across Scotland, therefore, providing training and support for the interview trainers themselves is the only way to ensure that JIIT is both standardised and of the highest quality it can be. Training the trainers in the use of the NICHD protocol, setting out a standardised structured course, standardising learning points and expectations of trainees and advising on how to provide feedback will help ensuring that JIIT courses are standardised and that trainees will be receiving the same experience across country. Refresher training should also be available for trainers in order to update them on recent research in order to ensure that they are fully aware of new developments in the field and give them the opportunity to evaluate how well their training

319

courses are working with other interview trainers working in different parts of the country.

4) Actors participating in JIIT need to receive training

Adult actors receive no preparation for their role in JIs and unsurprisingly actor responding was not standardised across Scotland in Study 1. In order to be able to support interviewers appropriately and to provide a practice experience that is comparable to interviewing real children, actors should also be adequately prepared for their roles as interviewees. The most appropriate way of preparing them would be to engage them in a bespoke training course. The adult actors have no specific knowledge of forensic interviewing and therefore are not sensitive to question type. Previous research has shown that when actors are trained how to respond in interviews that they can improve interviewers' questioning behaviour (Powell, Fisher & Hughes-Scholes, 2008). A training programme that educates adult actors to recognise the different types of question and teaches them how to respond appropriately in terms of number of words and details should allow actors to reinforce the best practice open prompts and diminish the use of riskier focussed questions. This thesis noted that adult actors often refused to engage with the practice interview, training must also emphasise the importance of engaging with all phases of the investigative interview. Actors also currently provide feedback to interviewers after the role-plays, as actors have no knowledge of how to conduct a best-practice interview and are not briefed on how to provide feedback, we do not know what they are

320

advising interviewers to do and whether or not their recommendations are in line with how we would want interviewers to behave. Therefore, actor training would not only need to prepare them for their role as interviewees but also advise on what constitutes and how to deliver appropriate feedback.

5) The scenarios used in JIIT should be standardised

The role-play scenarios used on JIIT courses varied greatly in terms of complexity (number of perpetrators/incidents) and difficulty (compliant/reluctant). This means that currently the JIIT experience trainee interviewers are having differs depending on the scenario they are assigned. The scenarios also contributed to some of the differences found between actors responding as scenarios that requested actors to behave in a difficult manner lead to uninformative responding. The scenarios used in training should be standardised in order to give interviewers the same opportunity to practice their questioning skills. If interviewers are asking best practice questions but receiving no information then it may compel them to move to asking focussed questions, as it the case with reluctant children in the field (Hershkowtitz et al., 2006). While it is the case that some children will be reluctant to disclose and interviewers should be prepared for such situations, this level of difficulty may be too high for trainees' first interview. If spaced learning is applied to JIIT courses then multiple practice opportunities can allow for scenario complexity and difficulty to be increased over time once

interviewees have had an opportunity to see their open prompts eliciting lengthy and detailed responses and have practiced maintaining their use.

6) *VRI facilities need to be made available for all JIs*

The next recommendation is that VRI facilities should be available on all interviewers' training courses and at all field interview sites. Interviewers that trained without VRI facilities were disadvantaged, scribing slowed the rate at which questions could be asked and answered and as a result they asked fewer questions and importantly fewer invitations than those that trained with VRI facilities. Further, information was lost during the interview as evidenced by scribes asking for repeats of information and their fellow-interviewer and the interviewee being unable to remember what had been said. If this is the case during training when interviewing an adult who may try to provide shorter responses in an effort to 'help the scribe out' then it is likely that in the field important information relevant to the case is being lost and interviewers do not have the opportunity to ask as many questions in the time they are given. The 2011 La Rooy et al., survey highlighted interviewers' concerns that scribing negatively impacted the quality of their interviews as it was difficult to note the lengthy responses to open prompts and thus they preferred more direct and focused questions. Using open prompts to elicit information from children is the most important aspect of interviewing, if the methods that are available to interviewers directly prevents them from applying best practice principles and instead makes using inappropriate questioning the preferred method then it is really important that policies change to implement methods that assist rather than hinder best practice.

While it is now part of the guidelines that all JIs are recorded, there are still sites in Scotland that do not have the facilities available and mobile recording kits or scribing may be used in such cases. If kits are unavailable or fail to work it is not recommended that the interview proceeds, instead the interview should be delayed until equipment can be accessed.

7) *Awareness on the unreliability of scribed evidence should be circulated to decision makers in the wider legal context*

The findings of Study 3 have added to the existing literature demonstrating the unreliability of scribed evidence (Lamb et al., 2000; Cauchi & Powell, 2009; Cauchi et al., 2010). The findings may even underestimate the full extent of the difficulty of scribing because adult actors had an awareness of the difficulty of scribing the interview real-time and were sympathetic to long pauses in the conversation while real children may not be. Also, the interviewers in this thesis were under observation while scribing these interviews, this may have had an effect on how motivated they were to faithfully record the interviews. Scribed JI interviews can be used as evidence in court, therefore, solicitors, sheriffs and jurors need to be made aware of the dangers of relying on scribed evidence and to exercise caution when making legal decisions based on them.

8) *More field research is needed in Scotland*

More field research is urgently needed in Scotland. Study 4 of this thesis and the 2 field studies by La Rooy et al. (2012, 2013) have all

concluded that the quality of interviews conducted with children alleging abuse in Scotland are low with open prompts comprising the smallest proportion of substantive questions. However, the first three studies of this thesis found that trainee interviews in Scotland were asking a large proportion of open prompts during training courses that had been completed at a later point in time than the field samples. The field research is limited in Scotland and the samples in the La Rooy et al. studies were potentially skewed (were all currently involved in court proceedings). A representative sample of recently conducted interviews with children in Scotland is needed in order to determine 1) the overall quality of interviews being conducted in the field in this country. 2) Whether the high proportion of open prompts used in the training interviews will continue in the field or whether these training gains will dissipate as has been the case post-training in other countries. It would be of particular interest if future research compared interviewers' training interviews and their first few field interviews conducted post-training.

9) Implementing specialist child interviewers in Scotland

A final suggestion is tentatively put forward; child interviewing may not be a skill that can or should be taught to all police officers and social workers. Indeed, recent research by Lafontaine and Cyr (2017) found that some personal characteristics (Emotional intelligence, Extraversion, Agreeableness, Conscientiousness and Neuroticism) of interviewers were associated with their adherence to the NICHD Protocol and the ratio of open-ended questions they asked. Interviewers' cognitive abilities were

related to the amount of details children alleging sexual abuse provided in these interviews. Further, some interviewers' attitudes and motivation towards interviewing are problematic. The 2011 La Rooy et al., survey found that many interviewers did not believe the techniques in the guidelines to be effective and are resistant to using them.

11.4 Concluding remark

Past failings of organisations involved in the handling of child abuse allegations have been so serious and widespread that currently in England and Wales alone13 investigations are underway by the Independent Inquiry into Child Sexual Abuse panel in an effort to explain how so many child abuse allegations have been handled so poorly. Those child victims are now adults and at the time of making their allegations child interviewing research and guidelines for interviewers were in their infancy. However, the past decade of field research has been conducted after the introduction of guidelines and formal interviewer training programmes yet the quality of interviews still remains low. The success of the NICHD Protocol coupled with specialised training has shown that the only way of maintaining high quality interviewing is by providing refresher training, continued practice interview opportunities, feedback on interviews conducted and continued supervision. New research suggests this can be achieved in a cost-effective manner (e.g. through online training exercises). If interviewers and professional legal organisations continue to ignore research recommendations and best practice guidelines then history is doomed to repeat itself and miscarriages of justice that we are currently investigating are likely to

continue in the future generations will be conducting investigations of their own into the mistakes our generation is making right now.

Appendices

Appendix 1A

Field interview checklist		**Coder ID:**	
Interview ID:			
	Present/NA	**Demo(Grs)**	**Line Numbers/comments**
Phase 1: Introduction			
Introduce self			
Explain job			
Introduce interviewer 2			
Explain interviewer 2 job			
Reference note taking/recording			
Obtain child's consent			
Phase 2: Rapport			
Ice-breaker			
Neutral topics about child			
Ground rules-			
Interviewer there to listen/child to talk			
Interviewer was not there/child is knowledgable			
Don't know			
Don't remember			
Don't understand			
Don't guess			
Correct me			
Repeated Q's			
Tell the truth			
Phase 3: Narrative elaboration training			
Practice interview of personally experienced event			
Use open ended prompts and avoid focussed questions to get account from start to end			
Phase 4: Questioning			
Transition question			
Use open ended prompts and avoid focussed questions to get account from start to end			
One time or more than one time			
Explain biological terms			
Phase 5: Closure			
Summarise account			
Interviewer 2 opportunity to ask questions			
Child opportunity to ask questions			
Inform what happens next			
Contact details			
Thank child			
Neutral topics about child as closure			
Additional Comments about interview			

Questionnaire data

On the final day of their JII training, the trainee interviewers were offered the opportunity to complete a role-play interview feedback questionnaire (designed by the author of this thesis) that inquired about their perceptions of their experience of interviewing adult actors role-playing children. Permission was requested from 80 course participants of 5 JII training courses that took place during 2012 and 2013, only 1 interviewer declined to participate and therefore 79 provided their evaluations of the role-plays. (One participant gave the 'trained' and 'untrained' actor a different rating in question 2, which is why n = 78 for this question (the 'trained' actor was rated higher with 'very realistic' and then 'untrained' 'realistic') and n = 77 for question 3 as 2 participants ticked more than 1 answer in response to this question). The questions and number and percentage of each answer chosen by course participants is detailed below:

1) How well do you feel the mock interview sessions with actors have equipped you to conduct interviews?

Very well equipped	Well equipped	Neither equipped nor unprepared	Unprepared	Very unprepared
52	27	0	0	0
65.8%	34.2%	0%	0%	0%

2) How realistic did you feel the mock interview sessions with actors were?

Very realistic	Realistic	Neither realistic nor unrealistic	Unrealistic	Very unrealistic
41	36	1	0	0
52.6%	46.2%	1.3%	0%	0%

3) How much did you enjoy the mock interview sessions with actors?

Enjoyed very much	Enjoyed	Neither enjoyed nor un-enjoyed	Did not enjoy	Very much did not enjoy
40	30	5	1	1
51.9%	39.0%	6.5%	1.3%	1.3%

(The one participant that chose 'very much did not enjoy' for Q3 wrote the following comment next to his answer "worth it for sure but just never been a fan of role-play")

Participants were also provided a blank space at the end of the questionnaire and invited to provide any additional feedback that they wished. Their comments reflected a few overarching themes

1) Participants perceived the role-play practice interviews as being important in preparing them for conducting interviews with children :

"*This was the part of the course which I learned most from. Practice experience much more needed than theory.*"

"*Very important part of the course to be able to put into practice what has been taught.*"

2) Participants thought the adult actors provided a realistic interview experience:

"*Even though the actors were adults it was very realistic and you forget they are adults. Gives very good example and experience for doing them in real life. Very good actors.*"

329

"*Even though the actors were adults they very quickly made it believable that you were talking to a child/young person.*"

"*Having the actors was great for the mock interviews, made a huge difference as they were very believable: really helped promote a sense of being in a real interview- how we engage with the children, cues to pick up on. Good experience for the real thing! Excellent.*"

3) Interviewers preferred training with actors to using fellow-participants:

"*From training at the Scottish Police College I have participated in practical exercises where fellow students play 'actors' and there is very little benefit from this as for various reasons (e.g. I'll do what I'm told so we finish earlier, I don't want my fellow classmate to look bad if they can't deal with me) it is nothing like real life scenarios. Bringing actors in to do this helps far more as they can provide a more realistic experience and their feedback is also probably more useful as it is more constructive.*"

"*Actors very lifelike and offered answers similar to expected age group, much better using actors than fellow course goers.*"

4) Interviewers felt that they benefitted from practicing using open prompting:

"*Very good & encouraged to use 'Tell me' so should set us up for the real thing.*"

330

"*Felt the mock interviews were very useful to help get used to using specific questions i.e. "Tell me…" feedback from mock interviews were also helpful.*"

5) Watching other participants role-play interviews promoted observational learning:

"*Very good and worthwhile for learning. Very realistic. Not only good to do interviews but to watch others as a learning tool.*"

6) Interviewers would like the opportunity to engage in more role-play interviews during training:

"*Good that the actors provided feedback. Would have liked more perhaps. 40 mins instead of 30 mins, but appreciate the time constraints. Found it very helpful*"

"*Good consolidation of theory. Encouraging feedback received increasing confidence in abilities. Perhaps one more interview would have been beneficial.*"

The only negative feedback received was with regards to one particular interview scenario. One of the scenarios features a teenager who is instructed to be wary of the police and the interviewer commented

331

that it was "*Particularly difficult to get rapport with a very evasive and uncommunicative teenager.*

List of References

Ahern, E. C., Andrews, S. J., Stolzenberg, S. N., & Lyon, T. D. (2015).The productivity of wh-prompts in child forensic interviews. *Journal of Interpersonal Violence*, 1-9.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. N. (1978). *Patterns of attachment: A psychological study of the strange situation*. Oxford, England: Lawrence Erlbaum Associates Publishers.

Aldridge, J., & Cameron, S. (1999). Interviewing child witnesses: Questioning techniques and the role of training. *Applied Developmental Science*, *3*, 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*, 304-316.

Alexander, L., & Dearsley, A. (2013). Using standardized patients in an undergraduate mental health simulation: A pilot study. *International Journal of Mental Health*, *42*, 149-164.

Alink, L. R., Cicchetti, D., Kim, J., & Rogosch, F. A. (2012). Longitudinal associations among child maltreatment, social functioning, and cortisol regulation. *Developmental Psychology*, *48*, 224-236.

Allen, B., & Tussey, C. (2012). Can projective drawings detect if a child experienced sexual or physical abuse? A systematic review of the controlled research. *Trauma, Violence, & Abuse*, *13*, 97-111.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*, 20–39.

American Professional Society on the Abuse of Children. (1990). Guidelines for psychosocial evaluation of suspected sexual abuse in young children. Chicago: Author.

Anderson, D. D., Gundersen, C. B., Banken, D. M., Halvorson, J. V., & Schmutte, D. (1989). Undergraduate role players as "clients" for graduate counselling students. *Teaching of Psychology*, *16*, 141-142.

Andersen, S. L., Tomada, A., Vincow, E. S., Valente, E., Polcari, A., & Teicher, M. H. (2008). Preliminary evidence for sensitive periods in the effect of childhood sexual abuse on regional brain development. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *20*, 292-301.

Andrews, S. J., Ahern, E. C., Stolzenberg, S. N., & Lyon, T. D. (2016). The Productivity of Wh-Prompts when children testify. *Applied Cognitive Psychology*, *30*, 341-349.

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*, 253-261.

Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence*, *16*, 427-454.

Armsworth, M. W., & Holaday, M. (1993).The effects of psychological trauma on children and adolescents. *Journal of Counseling and Development*, *72*, 49-56.

Aspelmeier, J. E., Elliott, A. N., & Smith, C. H. (2007). Childhood sexual abuse, attachment, and trauma symptoms in college females: The moderating role of attachment. *Child Abuse & Neglect*, *31*, 549-566.

Atance, C. M., & Sommerville, J. A. (2014). Assessing the role of memory in preschoolers' performance on episodic foresight tasks. *Memory*, *22*, 118-128.

Ayoub, C. C., O'connor, E., Rappolt-Schlichtmann, G., Fischer, K. W., Rogosch, F. A., Toth, S. L., & Cicchetti, D. (2006). Cognitive and emotional differences in young maltreated children: A translational application of dynamic skill theory. *Development and Psychopathology*, *18*, 679-706.

Barahal, R. M., Waterman, J., & Martin, H. P. (1981).The social cognitive development of abused children. *Journal of Consulting and Clinical Psychology*, *49*, 508.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6-to 24-month-old infants. *Infant Behavior and Development*, *19*, 159–170.

Bartlett, F. C. (1920). Some experiments on the reproduction of folk-stories. *Folklore*, *31*, 30-47.

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University Press.

Bauer, P. J. (2004). Getting explicit memory off the ground: Steps toward construction of a neuro-developmental account of changes in the first two years of life. *Developmental Review*, *24*, 347-373.

Bauer, P. J. (2006). Constructing a past in infancy: A neuro-developmental account. *Trends in Cognitive Sciences*, *10*, 175-181.

Bauer, P. J. (2007). *Remembering the times of our lives: Memory in infancy and beyond*. Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

Bauer, P. J. (2015). A complementary processes account of the development of childhood amnesia and a personal past. *Psychological Review*, *122*, 204-231.

Bauer, P. J. (2015). Development of episodic and autobiographical memory: The importance of remembering forgetting. *Developmental Review*, *38*, 146-166.

334

Bauer, P. J., & Hertsgaard, L. A. (1993). Increasing Steps in Recall of Events: Factors Facilitating Immediate and Long-Term Memory in 13.5-and 16.5-Month-Old Children. *Child Development*, *64*, 1204-1223.

Bauer, P. J., Kroupina, M. G., Schwade, J. A., Dropik, P. L., & Wewerka, S. S. (1998). If memory serves, will language? Later verbal accessibility of early memories. *Development and Psychopathology*, *10*, 655-679.

Bauer, P. J., & Larkina, M. (2014a). The onset of childhood amnesia in childhood: A prospective investigation of the course and determinants of forgetting of early-life events. *Memory*, *22*, 907-924.

Bauer, P. J., & Larkina, M. (2014b). Childhood amnesia in the making: Different distributions of autobiographical memories in children and adults. *Journal of Experimental Psychology: General*, *143*, 597-611.

Bauer, P. J., Larkina, M., & Doydum, A. O. (2012). Explaining variance in long-term recall in 3-and 4-year-old children: The importance of post-encoding processes. *Journal of Experimental Child Psychology*, *113*, 195-210.

Bauer, P. J., & Leventon, J. S. (2013). Memory for one-time experiences in the second year of life: Implications for the status of episodic memory. *Infancy*, *18*, 755-781.

Bauer, P. J., & Lukowski, A. F. (2010). The memory is in the details: Relations between memory for the specific features of events and long-term recall during infancy. *Journal of Experimental Child Psychology*, *107*, 1-14.

Bauer, P. J., & Mandler, J. M. (1989). One thing follows another: Effects of temporal structure on 1-to 2-year-olds' recall of events. *Developmental Psychology*, *25*, 197-206.

Bauer, P. J., & Shore, C. M. (1987). Making a memorable event: Effects of familiarity and organization on young children's recall of action sequences. *Cognitive Development*, *2*, 327-338.

Bauer, P. J., Wenner, J. A., Dropik, P. L., Wewerka, S. S., & Howe, M. L. (2000). Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monographs of the Society for Research in Child Development*, 65, i-213.

Bauer, P. J., & Wewerka, S. S. (1995). One-to two-year-olds′ recall of events: The more expressed, the more impressed. *Journal of Experimental Child Psychology, 59*, 475-496.

Bauer, P. J., Wiebe, S. A., Carver, L. J., Lukowski, A. F., Haight, J. C., Waters, J. M., & Nelson, C. A. (2006). Electrophysiological indexes of encoding and behavioral indexes of recall: examining relations and developmental change late in the first year of life. *Developmental Neuropsychology*, *29*, 293-320.

Baxter, J., Charles, K., Martin, M., & McGroarty, A. (2013). The relative influence of leading questions and negative feedback on response change on the Gudjonsson

Suggestibility Scale (2): Implications for forensic interviewing. *Psychology, Crime & Law*, *19*, 277-285.

Bays, J., & Chadwick, D. (1993). Medical diagnosis of the sexually abused child. *Child Abuse & Neglect*, *17*, 91-110.

Benson, M. S., & Powell, M. B. (2015). Evaluation of a comprehensive interactive training system for investigative interviewers of children. *Psychology, Public Policy, and Law*, *2*, 309.

Bergen, P. V., Salmon, K., Dadds, M. R., & Allen, J. (2009). The effects of mother training in emotion-rich, elaborative reminiscing on children's shared recall and emotion knowledge. *Journal of Cognition and Development*, *10*, 162-187.

Berliner, L., & Conte, J. R. (1993). Sexual abuse evaluations: Conceptual and empirical obstacles. *Child Abuse & Neglect*, *17*, 111-125.

Besharov, D. J. (1984). Liability in child welfare. *Public Welfare*, *42*, 28-33.

Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Brown, R. D., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology*, *14*, 421-433.

Bjorklund, D. F., & Douglas, R. N. (1997). The development of memory strategies. In N. Cowan & C. Hulme (Eds.), *The development of memory in childhood* (pp. 201-246). Psychology Press Ltd. Taylor & Francis group.

Blanchard-Dallaire, C., & Hébert, M. (2014). Social relationships in sexually abused children: self-reports and teachers' evaluation. *Journal of Child Sexual Abuse*, *23*, 326-344.

Bögels, S. M. (1994). A structured-training approach to teaching diagnostic interviewing. *Teaching of Psychology*, *21*, 144-150.

Bogo, M., Shlonsky, A., Lee, B., & Serbinski, S. (2014). Acting like it matters: A scoping review of simulation in child welfare training. *Journal of Public Child Welfare*, *8*, 70-93.

Bolger, K. E., & Patterson, C. J. (2001).Developmental pathways from child maltreatment to peer rejection. *Child Development*, *72*, 549-568.

Bow, J. N., Quinnell, F. A., Zaroff, M., & Assemany, A. (2002). Assessment of sexual abuse allegations in child custody cases. *Professional Psychology: Research and Practice*, *33*, 566-575.

Bowlby, J. (1958). The nature of the child's tie to his mother. *The International Journal of Psycho-analysis*, *39*, 350-373.

Bowlby, J. (1988). *A secure base.* New York: Basic Books.

Brady, M. S., Poole, D. A., Warren, A. R., & Jones, H. R. (1999). Young children's responses to yes-no questions: Patterns and problems. *Applied Developmental Science*, *3*, 47-57.

Braha-Zeitoun, S., Schenin-King, P., Bouaziz, N., Isaac, C., Moulier, V., Gaudeau-Bosma, C., & Januel, D. (2015). Source monitoring. *AnnalesMe´dico-Psychologiques*, *173*, 419–423.

Brainerd, C. J. (1985). Three-state models of memory development: A review of advances in statistical methodology. *Journal of Experimental Child Psychology*, *40*, 375-394.

Brainerd, C., & Ornstein, P. A. (1991). Children's memory for witnessed events: The developmental backdrop. In J. Doris (Ed.), *The suggestibility of children's recollections: Implications for eyewitness testimony* (pp. 10-20). Washington, DC: American Psychological Association.

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: a review of data and theory. *Psychological bulletin*, *134*, 343-382.

Briere, J., & Runtz, M. (1990). Differential adult symptomatology associated with three types of child abuse histories. *Child Abuse & Neglect*, *14*, 357-364.

Brown, D. A., & Lamb, M. E. (2015). Can children be useful witnesses? It depends how they are questioned. *Child Development Perspectives*, *9*, 250-255.

Brown, D. A., Lamb, M. E., Lewis, C., Pipe, M. E., Orbach, Y., & Wolfman, M. (2013). The NICHD investigative interview protocol: an analogue study. *Journal of Experimental Psychology: Applied*, *19*, 367-382.

Brown, D. A., Pipe, M. E., Lewis, C., Lamb, M. E., & Orbach, Y. (2007). Supportive or suggestive: Do human figure drawings help 5- to 7-year-old children to report touch? *Journal of consulting and clinical psychology*, *75*, 33-42.

Brubacher, S. P., Glisic, U., Roberts, K. P., & Powell, M. (2011). Children's ability to recall unique aspects of one occurrence of a repeated event. *Applied Cognitive Psychology*, *25*, 351-358.

Brubacher, S. P., Malloy, L. C., Lamb, M. E., & Roberts, K. P. (2013). How do interviewers and children discuss individual occurrences of alleged repeated abuse in forensic interviews? *Applied Cognitive Psychology*, *27*, 443-450.

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., Powell, M. B., & Roberts, K. P. (2014). Recommendations for interviewing children about repeated experiences. *Psychology, Public Policy, and Law*, *20*, 325-335.

Brubacher, S. P., Powell, M., Skouteris, H., & Guadagno, B. (2015). The effects of e-simulation interview training on teachers' use of open-ended questions. *Child Abuse & Neglect*, *43*, 95-103.

337

Brubacher, S.P., Roberts, K.P., & Powell, M. B. (2011). Effects of practicing episodic versus scripted recall on children's subsequent narratives of a repeated event. *Psychology, Public Policy, and Law, 17, 286-314.*

Brubacher, S. P., Roberts, K. P., & Powell, M. B. (2012). Retrieval of episodic versus generic information: Does the order of recall affect the amount and accuracy of details reported by children about repeated events? *Developmental Psychology*, *48*, 111-122.

Bruck, M., & Ceci, S. J. (1995). Amicus brief for the case of state of New Jersey v. Michaels presented by committee of concerned social scientists. *Psychology, Public Policy, and Law*, *1*, 272-322.

Bruck, M., & Ceci, S. J. (1999).The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419-439.

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (1995). "I hardly cried when I got my shot!" Influencing children's reports about a visit to their pediatrician. *Child Development*, *66*, 193-208.

Bruck, M., Ceci, S. J., Francouer, E., & Renick, A. (1995). Anatomically detailed dolls do not facilitate preschoolers' reports of a pediatric examination involving genital touching. *Journal of Experimental Psychology: Applied*, *1*, 95-109.

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology*, *18*, 947-996.

Burns, P., Russell, C., & Russell, J. (2015). Preschool children's proto-episodic memory assessed by deferred imitation. *Memory*, *23*, 1172-1192.

Cantón-Cortés, D., Cortés, M. R., & Cantón, J. (2015). Child sexual abuse, attachment style, and depression: The role of the characteristics of abuse. *Journal of Interpersonal Violence*, *30*, 420-436.

Carter , Y. H., Bannon, M. J., Limbert, C., Docherty, A., & Barlow, J. (2006). Improving child protection: A systematic review of training and procedural interventions. *Archives of Disease in Childhood*, *91*, 740-743.

Carver, L. J., & Bauer, P. J. (2001). The dawning of a past: the emergence of long-term explicit memory in infancy. *Journal of Experimental Psychology: General*, *130*, 726-745.

Cauchi, R., & Powell, M. B. (2009). Examination of Police Officers' Notes of Interviews with Alleged Child Abuse Victims. *International Journal of Police Science & Management*, *11*, 505-515.

Cauchi, R. T., Powell, M. B., & Hughes-Scholes, C. H. (2010). A controlled analysis of professionals' contemporaneous notes of interviews about alleged child abuse. *Child Abuse & Neglect*, *34*, 318-323.

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, *113*, 403-439.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. American Psychological Association.

Ceci, S. J., Huffman, M. L. C., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition, 3*, 388-407.

Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs amongpreschoolers. *International Journal of Clinical and Experimental Hypnosis, 42*, 304-320.

Ceci, S. J., Ross, D. F., & Toglia, M. P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116*, 38-49.

Cederborg, A. C., 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: An International Journal, 14*, 242-254.

Cederborg, A. C., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse & Neglect, 24*, 1355-1361.

Chae, Y., Goodman, G. S., Eisen, M. L., & Qin, J. (2011). Event memory and suggestibility in abused and neglected children: Trauma-related psychopathology and cognitive functioning. *Journal of Experimental Child Psychology, 110*, 520-538.

Chamberland, C., Lacharité, C., Clément, M. È., & Lessard, D. (2015). Predictors of development of vulnerable children receiving child welfare services. *Journal of Child and Family Studies, 24*, 2975-2988.

Clark, D. B., Thatcher, D. L., & Maisto, S. A. (2004). Adolescent neglect and alcohol use disorders in two-parent families. *Child Maltreatment, 9*, 357-370.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature, 395*, 272-274.

Cleveland, E. S., & Reese, E. (2008). Children remember early childhood: Long-term recall across the offset of childhood amnesia. *Applied Cognitive Psychology, 22*, 127-142.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*, 37-46.

Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior, 4*, 201-210.

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science, 210*, 207-210.

Collie, R., & Hayne, H. (1999). Deferred imitation by 6-and 9-month-old infants: More evidence for declarative memory. *Developmental Psychobiology, 35*, 83-90.

Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *Journal of Personality and Social Psychology*, *58*, 644-663.

Connolly, D. A., & Lindsay, D. S. (2001).The influence of suggestions on children's reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology*, *15*, 205-223.

Connolly, D. A., & Read, J. D. (2006). Delayed prosecutions of historic child sexual abuse: Analyses of 2064 Canadian criminal complaints. *Law and Human Behavior*, *30*, 409-434.

Conway, M. A., & Pleydell-Pearce, C. W. (2000).The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261-288.

Cook, E. T., Greenberg, M. T., & Kusche, C. A. (1994). The relations between emotional understanding, intellectual functioning, and disruptive behavior problems in elementary-school-aged children. *Journal of Abnormal Child Psychology*, *22*, 205-219.

Cordón, I. M., Pipe, M. E., Sayfan, L., Melinder, A., & Goodman, G. S. (2004). Memory for traumatic experiences in early childhood. *Developmental Review*, *24*, 101-132.

Coster, W. J., Gersten, M. S., Beeghly, M., & Cicchetti, D. (1989).Communicative functioning in maltreated toddlers. *Developmental Psychology*, *25*, 1020-1029.

Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. H. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science*, *3*, 77-85.

Cummings, E. M. (1987). Coping with background anger in early childhood. *Child Development*, *58*, 976-984.

Cummings, E. M., Hennessy, K. D., Rabideau, G. J., & Cicchetti, D. (1994). Responses of physically abused boys to interadult anger involving their mothers. *Development and Psychopathology*, *6*, 31-31.

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*, 516-524.

Cyr, M., & Lamb, M. E. (2009). Assessing the effectiveness of the NICHD investigative interview protocol when interviewing French-speaking alleged victims of child sexual abuse in Quebec. *Child Abuse & Neglect*, *33*, 257-268.

Curci, A., Bianco, A., & Gudjonsson, G. H. (2017). Verbal ability, depression, and anxiety as correlates of interrogative suggestibility in children exposed to life adversities. *Psychology, Crime & Law*, *23*, 445-458.

da Silva Ferreira, G. C., Crippa, J. A., & de Lima Osório, F. (2014). Facial emotion processing and recognition among maltreated children: A systematic literature review. *Frontiers in psychology*, *5*, 1460-1470.

Dadds, M. R., & Salmon, K. (2003). Punishment insensitivity and parenting: Temperament and learning as interacting risks for antisocial behavior. *Clinical Child and Family Psychology Review*, *6*, 69-86.

Dahl, J. J., Kingo, O. S., & Krøjgaard, P. (2015). The magic shrinking machine revisited: The presence of props at recall facilitates memory in 3-year-olds. *Developmental Psychology*, *51*, 1704-1716.

Daignault, I. V., & Hébert, M. (2008). Short-term correlates of child sexual abuse: An exploratory study predicting girls' academic, cognitive, and social functioning 1 year later. *Journal of Child & Adolescent Trauma*, *1*, 301-316.

Dale, P. S., Loftus, E. F., & Rathbun, L. (1978). The influence of the form of the question on the eyewitness testimony of preschool children. *Journal of Psycholinguistic Research*, *7*, 269-277.

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*, 446-458.

Danby, M. C., Brubacher, S. P., Sharman, S. J., Powell, M. B., & Roberts, K. P. (2017). Children's reasoning about which episode of a repeated event is best remembered. *Applied Cognitive Psychology*, *31*, 99-108.

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*, 81-97.

Davies, G. M., & Wilson, C. (1997). Implementation of the Memorandum: An overview. *Perspectives on the Memorandum: Policy, Practice and Research in Investigative Interviewing*, 1-12.

De Bellis, M. D., Clark, D. B., Beers, S. R., Soloff, P. H., Boring, A. M., Hall, J., Kersh, A., & Keshavan, M. S. (2000). Hippocampal volume in adolescent-onset alcohol use disorders. *American Journal of Psychiatry*, *157*, 737-744.

Deblinger, E., McLeer, S. V., Atkins, M. S., Ralphe, D., & Foa, E. (1989). Post-traumatic stress in sexually abused, physically abused, and nonabused children. *Child Abuse & Neglect*, *13*, 403-408.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Denham, S. A., McKinley, M., Couchoud, E. A., & Holt, R. (1990). Emotional and behavioral predictors of preschool peer ratings. *Child Development*, *61*, 1145-1152.

Dent, H. R. (1986). An experimental study of the effectiveness of different techniques of questioning mentally handicapped child witnesses. *British Journal of Clinical Psychology*, *25*, 13-17.

Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning child witnesses. *British Journal of Social and Clinical Psychology*, *18*, 41-51.

Desgranges, B., Baron, J. C., & Eustache, F. (1998). The functional neuroanatomy of episodic memory: the role of the frontal lobes, the hippocampal formation, and other areas. *Neuroimage*, *8*, 198-213.

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*, 87-97.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, *65*, 649-665.

Dodge, K. A., Pettit, G. S., Bates, J. E., & Valente, E. (1995). Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. *Journal of Abnormal Psychology*, *104*, 632-643.

Dozier, M., & Bernard, K. (2015). Attachment: normal development, individual differences, and associations with experience, In A. Thapar, D. S. Pine, J. F. Leckman, S. Scott, M. J. Snowling, & E. Taylor (Eds.), *Rutter's Child and Adolescent Psychiatry*, (pp. 65-78). John Wiley & Sons: Chichester, UK.

Duke, M. C., Uhl, E. R., Price, H., & Wood, J. M. (2016). Avoiding problems in child abuse interviews and investigations. In W. T. O'Donohue & M. Fanetti (Eds.), *Forensic Interviews Regarding Child Sexual Abuse* (pp. 219-236). Springer International Publishing.

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*, 746-761.

Earhart, B., La Rooy, D., & Lamb, M. (2016). Assessing the quality of forensic interviews with child witnesses. In W. T. O'Donohue & M. Fanetti (Eds.), *Forensic Interviews Regarding Child Sexual Abuse* (pp. 317-335). Springer International Publishing.

Edwards, A., Shipman, K., & Brown, A. (2005). The socialization of emotional understanding: A comparison of neglectful and nonneglectful mothers and their children. *Child Maltreatment*, *10*, 293-304.

Egeland, B., & Sroufe, L. A. (1981). Attachment and early maltreatment. *Child Development*, *52*, 44-52.

Eigsti, I. M., & Cicchetti, D. (2004). The impact of child maltreatment on expressive syntax at 60 months. *Developmental Science*, *7*, 88-102.

Eisen, M. L., Goodman, G. S., Qin, J., Davis, S., & Crayton, J. (2007). Maltreated children's memory: accuracy, suggestibility, and psychopathology. *Developmental Psychology*, *43*, 1275-1294.

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*, 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*, 334-345.

Farrant, K., & Reese, E. (2000). Maternal style and children's participation in reminiscing: Stepping stones in children's autobiographical memory development. *Journal of Cognition and Development*, *1*, 193-225.

Farrar, M. J., & Goodman, G. S. (1990). Developmental differences in the relation between scripts and episodic memory: Do they exist? In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 30-64). New York: Cambridge University Press.

Feiring, C., Taska, L., & Lewis, M. (1998). The role of shame and attributional style in children's and adolescents' adaptation to sexual abuse. *Child Maltreatment*, *3*, 129-142.

Feltis, B. B., Powell, M. B., Snow, P. C., & Hughes-Scholes, C. H. (2010). An examination of the association between interviewer question type and story-grammar detail in child witness interviews about abuse. *Child Abuse & Neglect*, *34*, 407-413.

Finkelhor, D. (1984). Child sexual abuse: New theory and research. In E. Ullmann & W. Hilweg (Eds.), Childhood and trauma: Separation, abuse, war (pp. 101-117). England: Ashgate Publishing Ltd.

Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry*, *55*, 530-541.

Finkelhor, D., Ormrod, R. K., & Turner, H. A. (2007). Poly-victimization: A neglected component in child victimization. *Child Abuse & Neglect*, *31*, 7-26.

Finzi, R., Cohen, O., Sapir, Y., & Weizman, A. (2000). Attachment styles in maltreated children: A comparative study. *Child Psychiatry and Human Development*, *31*, 113-128.

Fishbein, D., Warner, T., Krebs, C., Trevarthen, N., Flannery, B., & Hammond, J. (2009). Differential relationships between personal and community stressors and children's neurocognitive functioning. *Child Maltreatment*, *14*, 299-315.

Fivush, R. (2014). Emotional content of parent-child conversations about the past. In C. A. Nelson (Ed.), *Memory and affect in development: The Minnesota symposia on child psychology (Vol. 26)*, (pp. 39-79). New York: Psychology Press.

Fivush, R., & Fromhoff, F. A. (1988). Style and structure in mother-child conversations about the past. *Discourse Processes*, *11*, 337-355.

Fivush, R., Gray, J. T., & Fromhoff, F. A. (1987). Two-year-old talk about the past. *Cognitive Development*, *2*, 393-409.

Fivush, R., Haden, C., & Adam, S. (1995). Structure and coherence of preschoolers′ personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology*, *60*, 32-56.

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

Fivush, R., & Schwarzmueller, A. (1995). Say it once again: Effects of repeated questions on children's event recall. *Journal of Traumatic Stress*, *8*, 555-580.

Fivush, R., & Schwarzmueller, A. (1998). Children remember childhood: Implications for childhood amnesia. *Applied Cognitive Psychology*, *12*, 455-473.

Fivush, R., & Shukat, J. R. (1995). Content, consistency, and coherence of early autobiographical recall. In M.S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, & Y. S. Ben-Porath (Eds.), *Memory and testimony in the child witness* (pp. 5-23). Thousand Oaks, CA: Sage.

Freeman, K. A., & Morris, T. L. (1999). Investigative interviewing with children: Evaluation of the effectiveness of a training program for child protective service workers. *Child Abuse & Neglect*, *23*, 701-713.

Frías, M. T., Brassard, A., & Shaver, P. R. (2014). Childhood sexual abuse and attachment insecurities as predictors of women's own and perceived-partner extradyadic involvement. *Child Abuse & Neglect*, *38*, 1450-1458.

Friedrich, W. N., Einbender, A. J., & Leucke, W. J. (1983). Cognitive and behavioral characteristics of physically abused children. *Journal of Consulting and Clinical Psychology*, *51*, 313-314.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development*, *56*, 1145-1155.

Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, *54*, 51-60.

Foley, M. A., Johnson, M. K., & Raye, C. L. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development*, 51-60.

Foley, M. A., Passalacqua, C., & Ratner, H. H. (1993). Appropriating the actions of another: Implications for children's memory and learning. *Cognitive Development*, *8*, 373-401.

Foley, M. A., & Ratner, H. H. (1998). Children's recoding in memory for collaboration: A way of learning from others. *Cognitive Development*, *13*, 91-108.

Foley, M. A., Ratner, H. H., & House, A. T. (2002). Anticipation and source-monitoring errors: Children's memory for collaborative activities. *Journal of Cognition and Development*, *3*, 385-414.

Foley, M. A., Santini, C., & Sopasakis, M. (1989). Discriminating between memories: Evidence for children's spontaneous elaborations. *Journal of Experimental Child Psychology*, *48*, 146-169.

Fugazza, C., Pogány, Á., & Miklósi, Á. (2016). Do as I… did! Long-term memory of imitative actions in dogs (Canis familiaris). *Animal cognition*, *19*, 263-269.

Gagnon, K., & Cyr, M. (2017). Sexual abuse and preschoolers: Forensic details in regard of question types. *Child Abuse & Neglect*, *67*, 109-118.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, *3*, 208-214.

Garry, M., & Polaschek, D. L. (2000). Imagination and memory. *Current Directions in Psychological Science*, *9*, 6-10.

Garry, M., Sharman, S. J., Wade, K. A., Hunt, M. J., & Smith, P. J. (2001). Imagination inflation is a fact, not an artifact. *Memory & Cognition*, *29*, 719- 729.

Garven, S., Wood, J. M., & Malpass R. S. (2000). Allegations of wrongdoing: The effects of reinforcement on children's mundane and fantastic claims. *Journal of Applied Psychology, 85*, 38-49.

Gaudin, J. M. (1999). Child neglect: Short-term and long-term outcomes. In H. Dubowitz (Ed.), *Neglected children: Research, practice, and policy* (pp. 89-108). Thousand Oaks, CA: Sage.

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*, 59-68.

George, C., & Main, M. (1979). Social interactions of young abused children: Approach, avoidance, and aggression. *Child Development*, *50*, 306-318.

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*, 381-395.

Giles, J. W., Gopnik, A., & Heyman, G. D. (2002). Source monitoring reduces the suggestibility of preschool children. *Psychological Science*, *13*, 288-291.

Gleeson, G., & Fitzgerald, A. (2014). Exploring the association between adult attachment styles in romantic relationships, perceptions of parents from childhood and relationship satisfaction. *Health*, *6,* 1643-1661.

Goetzold, S. (2015). An open and shut case of closed questions: An exploration of Joint Investigative Interview Training in Scotland. *Child Abuse Review*, *26*, 116-129.

Gómez, R. L., & Edgin, J. O. (2016). The extended trajectory of hippocampal development: Implications for early memory development and disorder. *Developmental Cognitive Neuroscience*, *18*, 57-69.

Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child development*, *61*, 1859-1871.

Goodman, G. S., Batterman-Faunce, J. M., Schaaf, J. M., & Kenney, R. (2002). Nearly 4 years after an event: Children's eyewitness memory and adults' perceptions of children's accuracy. *Child Abuse & Neglect*, *26*, 849-884.

Goodman, G. S., Bottoms, B. L., Rudy, L., Davis, S. L., & Schwartz-Kenney, B. M. (2001). Effects of past abuse experiences on children's eyewitness memory. *Law and Human Behavior*, *25*, 269-298.

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*, 69-99.

Goodman, G. S., Jones, O., & McLeod, C. (2017). Is There Consensus About Children's Memory and Suggestibility? *Journal of Interpersonal Violence*, *32*, 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*, 54-75.

Goodman, G. S., Taub, E. P., Jones, D. P., England, P., Port, L. K., Rudy, L., & Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development*, 57, i-159.

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*, 525-540.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 501-518.

Greenberg, M. T., Siegel, J. M., & Leitch, C. J. (1983). The nature and importance of attachment relationships to parents and peers during adolescence. *Journal of Youth and Adolescence*, *12*, 373-386.

Guadagno, B. L., Powell, M. B., & Wright, R. (2006). Police officers' and legal professionals' perceptions regarding how children are, and should be, questioned about repeated abuse. *Psychiatry, Psychology and Law*, *13*, 251-260.

Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Personality and Individual Differences*, *5*, 303-314.

Haden, C. A., Ornstein, P. A., Eckerman, C. O., & Didow, S. M. (2001). Mother–child conversational interactions as events unfold: Linkages to subsequent remembering. *Child Development*, *72*, 1016-1031.

Hala, S., Brown, A. M., McKay, L. A., & San Juan, V. (2013). Two-and-a-half-year-olds' memory for sources of actions: Contextual support facilitates recall. *Journal of Cognition and Development*, *14*, 343-358.

Hamilton, G., Brubacher, S. P., & Powell, M. B. (2016). Expressions of shame in investigative interviews with Australian Aboriginal children. *Child Abuse & Neglect*, *51*, 64-71.

Hamond, N. R., & Fivush, R. (1991). Memories of Mickey Mouse: Young children recount their trip to Disneyworld. *Cognitive Development*, *6*, 433-448.

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences*, *17*, 111-120.

Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review*, *24*, 33-73.

Hayne, H., & Herbert, J. (2004). Verbal cues facilitate memory retrieval during infancy. *Journal of Experimental Child Psychology*, *89*, 127-139.

Hayne, H., & Imuta, K. (2011). Episodic memory in 3-and 4-year-old children. *Developmental Psychobiology*, *53*, 317-322.

Hayne, H., & Jack, F. (2011). Childhood amnesia. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 136-145.

Hayne, H., & Rovee-Collier, C. (1995). The organization of reactivated memory in infancy. *Child Development*, *66*, 893-906.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511-524.

Hébert, M. Langevin, R., & Bernier, M. J. (2013). Self-reported symptoms and parents' evaluation of behaviour problems in preschoolers disclosing sexual abuse. *International Journal of Child, Youth and Family Studies*, *4*, 467-483.

Hébert, M., Langevin, R., & Daigneault, I. (2016). The association between peer victimization, PTSD, and dissociation in child victims of sexual abuse. *Journal of Affective Disorders*, *193*, 227-232.

Hershkowitz, I. (2009). Socioemotional factors in child abuse investigations. *Child Maltreatment*, *14*, 172-181.

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). Chichester, UK: Wiley-Blackwell.

Hershkowitz, I., Fisher, S., Lamb, M. E., & Horowitz, D. (2007). Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. *Child Abuse & Neglect, 31*, 99-110.

Hershkowitz, I., Horowitz, D., & Lamb, M. E. (2005). Trends in children's disclosure of abuse in Israel: A national study. *Child Abuse & Neglect, 29*, 1203-1214.

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*, 336-344.

Hershkowitz, I., Lamb, M. E., & Horowitz, D. (2007). Victimization of children with disabilities. *American Journal of Orthopsychiatry, 77*, 629-635.

Hershkowitz, I., Lamb, M. E., Katz, C., & Malloy, L. C. (2015). Does enhanced rapport-building alter the dynamics of investigative interviews with suspected victims of intra-familial abuse? *Journal of Police and Criminal Psychology, 30*, 6-14.

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*, 753-769.

Hirst, W., & Manier, D. (2008). Towards a psychology of collective memory. *Memory, 16*, 183-200.

Home Office. (1992). Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings. London, England: Author, with Department of Health.

Hostinar, C. E., Stellern, S. A., Schaefer, C., Carlson, S. M., & Gunnar, M. R. (2012). Associations between early life adversity and executive function in children adopted internationally from orphanages. *Proceedings of the National Academy of Sciences, 109*, 17208-17212.

Howe, M. L., Cicchetti, D., Toth, S. L., & Cerrito, B. M. (2004). True and false memories in maltreated children. *Child Development, 75*, 1402-1417.

Hudson, J. A., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology, 6*, 483-505.

Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. *First Language, 1*, 149-160.

Hutcheson, G. D., Baxter, J. S., Telfer, K., & Warden, D. (1995). Child witness statement quality: Question type and errors of omission. *Law and Human Behavior, 19*, 631-648.

Jack, F., MacDonald, S., Reese, E., & Hayne, H. (2009). Maternal reminiscing style during early childhood predicts the age of adolescents' earliest memories. *Child Development, 80*, 496-505.

Jacoby, L. L. (1983). Perceptual enhancement: persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 21-38.

Joëls, M. (2008). Functional actions of corticosteroids in the hippocampus. *European Journal of Pharmacology*, *583*, 312-321.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 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*, 661-668.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85.

Kandel, E. R., & Squire, L. R. (2001). Neuroscience. *Annals of the New York Academy of Sciences*, *935*, 118-135.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*, 637-671.

Karen, R. (1994). *Becoming attached: Unfolding the mystery of the infant–mother bond and its impact on later life*. Warner Books.

Kellogg, N. D. (2017). "Why didn't you tell?" Helping families and children weather the process following a sexual abuse disclosure. In D. M. Teti (Ed.), *Parenting and Family Processes in Child Maltreatment and Intervention* (pp. 35-43). Springer International Publishing.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, *113*, 164-180.

Kim, J., & Cicchetti, D. (2006). Longitudinal trajectories of self-system processes and depressive symptoms among maltreated and nonmaltreated children. *Child Development*, *77*, 624-639.

Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry*, *51*, 706-716.

Kim, I. K., Kwon, E. S., & Ceci, S. J. (2017). Developmental reversals in report conformity: Psycho-legal implications. *Applied Cognitive Psychology, 31*, 128-138.

Kinard, E. M. (1999). Psychosocial resources and academic performance in abused children. *Children and Youth Services Review*, *21*, 351-376.

Kinard, E. M. (2001). Perceived and actual academic competence in maltreated children. *Child Abuse & Neglect*, *25*, 33-45.

King, M. A., & Yuille, J. C. (1987). Suggestibility and the child witness. In S. J. Ceci, M. E. Toglia & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 24-35). US: Springer.

Kitzinger, J. (2000). Media templates: Patterns of association and the (re)construction of meaning over time, *Media Culture and Society*, *22*, 61–84.

Klimes-Dougan, B., & Kistner, J. (1990). Physically abused preschoolers' responses to peers' distress. *Developmental Psychology*, *26*, 599-602.

Korkman, J., Santtila, P., Drzewiecki, T., & Kenneth Sandnabba, N. (2008). Failing to keep it simple: Language use in child sexual abuse interviews with 3–8-year-old children. *Psychology, Crime & Law*, *14*, 41-60.

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*, 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*, 868-884.

Krause, N., Pompedda, F., Antfolk, J., Zappalá, A., & Santtila, P. (2017). The effects of feedback and reflection on the questioning style of untrained interviewers in simulated child sexual abuse interviews. *Applied Cognitive Psychology*, *31*, 187-198.

Lamb, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse & Neglect*, *18*, 1021-1028.

Lamb, M. E., & Brown, D. A. (2006). Conversational apprentices: Helping children become competent informants about their own experiences. *British Journal of Developmental Psychology*, *24*, 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*, 483–491.

Lamb, M. E., Hershkowitz, I., & Lyon, T. D. (2013). Interviewing victims and suspected victims who are reluctant to talk. *APSAC (American Professional Society on the Abuse of Children) Advisor*, *25*, 16-19.

Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). *Tell me what happened*. Chichester, UK and Hoboken, NJ: Wiley.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Boat, B., & Everson, M. D. (1996). Investigative interviews of alleged sexual abuse victims with and without anatomical dolls. *Child Abuse & Neglect*, *20*, 1251-1259.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., & Yudilevitch, L. (1996) Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development*, *19*, 627-637.

Lamb, M. E., Orbach, Y., Hershkowitz, I., Horowitz, D., & Abbott, C. B. (2007). Does the type of prompt affect the accuracy of information provided by alleged victims of abuse in forensic interviews? *Applied Cognitive Psychology*, *21*, 1117-1130.

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*, 1201-1231.

Lamb, M. E., Orbach, Y., Hershkowitz, I., Horowitz, D., & Abbott, C. B. (2007). Does the type of prompt affect the accuracy of information provided by alleged victims of abuse in forensic interviews? *Applied Cognitive Psychology*, *21*, 1117-1130.

Lamb, M. E., Orbach, Y., Sternberg, K. J., Aldridge, J., Pearson, S., Stewart, H. L., Esplin, P.W., & Bowler, L. (2009). Use of a structured investigative protocol enhances the quality of investigative interviews with alleged victims of child sexual abuse in Britain. *Applied Cognitive Psychology*, *23*, 449-467.

Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators' verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior*, *24*, 699.

Lamb, M. E., Orbach, Y., Warren, A. R., Esplin, P. W., & Hershkowitz, I. (2007). Enhancing performance: Factors affecting the informativeness of young witnesses. In M. P. Toglia, J. D. Read, D. F. Ross & R. C. L. Lindsay (Eds), *The handbook of eyewitness psychology, Vol I: Memory for events* (pp. 429-451). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment. *Journal of Applied Developmental Psychology*, *15*, 255-280.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1995). Making children into competent witnesses: Reactions to the amicus brief In re Michaels. *Psychology, Public Policy, and Law*, *1*, 438-449.

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*, 1586-1596.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002a). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science*, *6*, 35-41.

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*, 926-934.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2002b). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science*, *6*, 114-125.

351

La Rooy, D., Brubacher, S. P., Aromäki-Stratos, A., Cyr, M., Hershkowitz, I., Korkman, J., Myklebust, T., Naka, M., Peixoto, C. E., Roberts, K. P., & Stewart, H. (2015). The NICHD protocol: A review of an internationally-used evidence-based tool for training child forensic interviewers. *Journal of Criminological Research, Policy and Practice*, *1*, 76-89.

La Rooy, D. J., Earhart, B., & Nicol, A. (2013). Joint Investigative Interviews (JIIs) conducted with children in Scotland: a comparison of the quality of interviews conducted before and after the introduction of the Scottish Executive (2011) guidelines. *Scots Law Times, 31,* 217-219.

La Rooy, D., Lamb, M. E., & Memon, A. (2011). Forensic interviews with children in Scotland: A survey of interview practices among police. *Journal of Police and Criminal Psychology*, *26*, 26-34.

La Rooy, D., Nicol, A., Halley, J., & Lamb, M. E. (2012). Joint investigative interviews with children in Scotland. *Scots Law Times*, *18*, 163-172.

Lane, K. (1988). Using actors as "clients" for an interviewing simulation in an undergraduate clinical psychology course. *Teaching of Psychology*, *15*, 162-164.

Langevin, R., Cossette, L., & Hébert, M. (2016). Emotion regulation in sexually abused preschoolers. *Child Psychiatry and Human Development*, *47*, 1-12.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, *31*, 568-578.

Lev-Wiesel, R., & Sternberg, R. (2012). Victimized at home revictimized by peers: Domestic child abuse a risk factor for social rejection. *Child and Adolescent Social Work Journal*, *29*, 203-220.

Levitov, J. E., Fall, K. A., & Jennings, M. C. (1999). Counselor clinical training with client-actors. *Counselor Education and Supervision*, *38*, 249-260.

Levy, S. R., Ayduk, O., & Downey, G. (2001). The role of rejection sensitivity in people's relationships with significant others and valued social groups. In M. R. Leary (Ed.), *Interpersonal Rejection* (pp. 251-290). New York: Oxford University Press.

Lexton, A., Smith, M., Olufemi, D., & Poole, G. (2005). Taking a risk and playing it safe: The use of actors in interagency child protection training. *Child Abuse Review*, *14*, 195-206.

Liao, C. S., Kao, S. P., Liang, S. Y., & Hsieh, M. C. (2015). Training actors as standardized patients. *Tzu Chi Medical Journal*, *27*, 96-97.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, *52*, 297-318.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*, 361-366.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585-589.

London, K., Bruck, M., Wright, D. B., & Ceci, S. J. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological issues, and implications for forensic interviewers. *Memory*, *16*, 29-47.

London, B., Downey, G., Bonica, C., & Paltin, I. (2007). Social causes and consequences of rejection sensitivity. *Journal of Research on Adolescence*, *17*, 481-506.

Loos, M. E., & Alexander, P. C. (1997). Differential effects associated with self-reported histories of abuse and neglect in a college sample. *Journal of Interpersonal Violence*, *12*, 340-360.

Lukowski, A. F., Garcia, M. T. L., & Bauer, P. J. (2011). Memory for events and locations obtained in the context of elicited imitation: Evidence for differential retention in the second year of life. *Infant Behavior and Development*, *34*, 55-62.

Lynch, M., & Cicchetti, D. (1998). Trauma, mental representation, and the organization of memory for mother-referent material. *Development and Psychopathology*, *10*, 739-759.

Lyon, T. D. (1999). The new wave of suggestibility research: A critique. *Cornell Law Review*, *84*, 1004–1087.

Lyon, T. D., & Dorado, J. S. (2008). Truth induction in young maltreated children: The effects of oath-taking and reassurance on true and false disclosures. *Child Abuse & Neglect*, *32*, 738-748.

Macfie, J., Toth, S. L., Rogosch, F. A., Robinson, J., Emde, R. N., & Cicchetti, D. (1999). Effect of maltreatment on preschoolers' narrative representations of responses to relieve distress and of role reversal. *Developmental Psychology*, *35*, 460-465.

Maguire, S. A., Williams, B., Naughton, A. M., Cowley, L. E., Tempest, V., Mann, M. K., Teague, M. & Kemp, A. M. (2015). A systematic review of the emotional, behavioural and cognitive features exhibited by school-aged children experiencing neglect or emotional abuse. *Child: Care, Health and Development*, *41*, 641-653.

Malloy, L. C., Brubacher, S. P., & Lamb, M. E. (2011). Expected consequences of disclosure revealed in investigative interviews with suspected victims of child sexual abuse. *Applied Developmental Science*, *15*, 8-19.

Manly, J. T., Cicchetti, D., & Barnett, D. (1994). The impact of subtype, frequency, chronicity, and severity of child maltreatment on social competence and behavior problems. *Development and Psychopathology*, *6*, 121-143.

Mandler, G. (2002). Organisation: What levels of processing are levels of. *Memory*, *10*, 333-338.

Mannarino, A. P., & Cohen, J. A. (1986). A clinical-demographic study of sexually abused children. *Child Abuse & Neglect*, *10*, 17-23.

Masten, C. L., Guyer, A. E., Hodgdon, H. B., McClure, E. B., Charney, D. S., Ernst, M., Kaufman, J., Pine, D.S., & Monk, C. S. (2008). Recognition of facial emotions among maltreated children with high rates of post-traumatic stress disorder. *Child Abuse & Neglect*, *32*, 139-153.

Maughan, A., & Cicchetti, D. (2002). Impact of child maltreatment and interadult violence on children's emotion regulation abilities and socioemotional adjustment. *Child Development*, *73*, 1525-1542.

McCabe, A., & Peterson, C. (1991). Getting the story: A longitudinal study of parental styles in eliciting narratives and developing narrative skill. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 217-255). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, *114*, 1-16.

McDermott, K. B., & Roediger, H. L. (2016). Memory (Encoding, storage, retrieval). *Noba textbook series: Psychology. Champaign, IL: DEF publishers.*

McDonough, L., Mandler, J. M., McKee, R. D., & Squire, L. R. (1995). The deferred imitation task as a nonverbal measure of declarative memory. *Proceedings of the National Academy of Sciences*, *92*, 7580-7584.

McElvaney, R. (2015). Disclosure of child sexual abuse: Delays, non-disclosure and partial disclosure. What the research tells us and implications for practice. *Child Abuse Review*, *24*, 159-169.

McLeer, S. V., Callaghan, M., Henry, D., & Wallen, J. (1994). Psychiatric disorders in sexually abused children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*, 313-319.

McWilliams, K., Harris, L. S., & Goodman, G. S. (2014). Child maltreatment, trauma-related psychopathology, and eyewitness memory in children and adolescents. *Behavioral Sciences & the Law*, *32*, 702-717.

Melinder, A., Endestad, T. O. R., & Magnussen, S. (2006). Relations between episodic memory, suggestibility, theory of mind, and cognitive inhibition in the preschool child. *Scandinavian Journal of Psychology*, *47*, 485-495.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Memory and Language*, *2*, 1-21.

Meltzoff, A. N. (1985). Immediate and deferred imitation in fourteen-and twenty-four-month-old infants. *Child Development*, 62-72.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, *24*, 470.

Meltzoff, A. N. (1995). What infant memory tells us about infantile amnesia: Long-term recall and deferred imitation. *Journal of Experimental Child Psychology*, *59*, 497-515.

Memon, A., Bull, R., & Smith, M. (1995). Improving the quality of the police interview: Can training in the use of cognitive techniques help? *Policing and Society: An International Journal*, *5*, 53-68.

Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Psychology*, *87*, 403-415.

Mills, R., Alati, R., O'Callaghan, M., Najman, J. M., Williams, G. M., Bor, W., & Strathearn, L. (2011). Child abuse and neglect and cognitive function at 14 years of age: Findings from a birth cohort. *Pediatrics*, *127*, 4-10.

Ministry of Justice. (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance for Vulnerable or Intimidated Witnesses, including Children*. London: Ministry of Justice.

Morris, G., Baker-Ward, L., & Bauer, P. J. (2010). What remains of that day: The survival of children's autobiographical memories across time. *Applied Cognitive Psychology*, *24*, 527-544.

Moscovitch, M. (1985). Memory from infancy to old age: Implications for theories of normal and pathological memorya. *Annals of the New York Academy of Sciences*, *444*, 78-96.

Moston, S., & Engelberg, T. (1993). Police questioning techniques in tape recorded interviews with criminal suspects. *Policing and Society: An International Journal*, *3*, 223-237.

Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect*, *20*, 623-631.

Mullen, P. E., Martin, J. L., Anderson, J. C., Romans, S. E., & Herbison, G. P. (1996). The long-term impact of the physical, emotional, and sexual abuse of children: A community study. *Child Abuse & Neglect*, *20*, 7-21.

Murachver, T., Pipe, M. E., Gordon, R., Owens, J. L., & Fivush, R. (1996). Do, show, and tell: Children's event memories acquired through direct experience, observation, and stories. *Child Development*, *67*, 3029-3044.

Myklebust, T., & Alison, L. (2000). The current state of police interviews with children in Norway: How discrepant are they from models based on current issues in memory and communication? *Psychology, Crime and Law*, *6*, 331-351.

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science*, *4*, 7-14.

Nelson, S., Faust, J., Doyle, K., & Delucia, C. (2015). Beyond Depression: Correlates of suicidal behavior in abused children. *Journal of Child & Adolescent Trauma*, *8*, 93-101.

355

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review*, *111*, 486-511.

Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M., & Koski, J. (2014).Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Developmental Science*, *17*, 743-756.

Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. *Advances in Child Development and Behaviour*, *35*, 40-89.

Oates, R. K., Lynch, D. L., Stern, A. E., O'Toole, B. I., & Cooney, G. (1995). The criminal justice system and the sexually abused child. Help or hindrance? *The Medical Journal of Australia*, *162*, 126-130.

Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and non-stressful events. *Medicine, Science and the Law*, *31*, 4-10.

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, *24*, 733-752.

Orbach, Y., & Lamb, M. E. (1999). Assessing the accuracy of a child's account of sexual abuse: A case study. *Child Abuse & Neglect*, *23*, 91‑98.

Orbach, Y., & Lamb, M. E. (2000). Enhancing children's narratives in investigative interviews. *Child Abuse & Neglect*, *24*, 1631-1648.

Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect*, *25*, 323-333.

Orbach, Y., Lamb, M. E., La Rooy, D., & Pipe, M. E. (2012). A case study of witness consistency and memory recovery across multiple investigative interviews. *Applied Cognitive Psychology*, *26*, 118-129.

Ornstein, P. A., Baker-Ward, L., Gordon, B. N., Pelphrey, K. A., Tyler, C. S., & Gramzow, E. (2006). The influence of prior knowledge and repeated questioning on children's long-term retention of the details of a pediatric examination. *Developmental Psychology*, *42*, 332-344.

Ornstein, P. A., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology*, *6*, 49-60.

Osborn, C. J., Dean, E. P., & Petruzzi, M. L. (2004). Use of simulated multidisciplinary treatment teams and client actors to teach case conceptualization and treatment planning skills. *Counselor Education and Supervision*, *44*, 121-134.

Oshri, A., Sutton, T. E., Clay-Warner, J., & Miller, J. D. (2015). Child maltreatment types and risk behaviors: Associations with attachment style and emotion regulation dimensions. *Personality and Individual Differences*, *73*, 127-133.

Otgaar, H., Howe, M. L., Peters, M., Sauerland, M., & Raymaekers, L. (2013). Developmental trends in different types of spontaneous false memories: Implications for the legal field. *Behavioral Sciences & the Law*, *31*, 666-682.

Ozbaran, B., Erermis, S., Bukusoglu, N., Bildik, T., Tamar, M., Ercan, E. S., Aydin, C., & Cetin, S. K. (2009). Social and emotional outcomes of child sexual abuse: A clinical sample in Turkey. *Journal of Interpersonal Violence*, *24*, 1478-1493.

Paolucci, E. O., Genuis, M. L., & Violato, C. (2001). A meta-analysis of the published research on the effects of child sexual abuse. *The Journal of Psychology*, *135*, 17-36.

Papaleontiou Louca, E., & Kyriakides, S. (2016). Attachment, trust, love and personality development. In A. Zamosteanu (Ed.), *Theoretical and Applied in Psychology SICAP23* (pp. 135-139). Bologna, Italy: Medimond.

Parke, R. D., Cassidy, J., Burks, V. M., Carson, J. L., & Boyum, L. (1992). Family contributions to peer relationships among young children. In R.D. Parke & G. Ladd (Eds.), *Family-peer relations: Modes of linkage* (pp.107-134). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Pathman, T., & Bauer, P. J. (2013). Beyond initial encoding: Measures of the post-encoding status of memory traces predict long-term recall during infancy. *Journal of Experimental Child Psychology*, *114*, 321-338.

Paz-Alonso, P. M., & Goodman, G. S. (2016). Developmental differences across middle childhood in memory and suggestibility for negative and positive events. *Behavioral Sciences & the Law*, *34*, 30-54.

Perner, J., Kloo, D., & Gornik, E. (2007). Episodic memory development: Theory of mind is part of re-experiencing experienced events. *Infant and Child Development*, *16*, 471-490.

Perner, J., & Ruffman, T. (1995). Episodic memory and autonoeticconciousness: developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology*, *59*, 516-548.

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*, 609-629.

Peters, W. W., & Nunez, N. (1999). Complex language and comprehension monitoring: Teaching child witnesses to recognize linguistic confusion. *Journal of Applied Psychology*, *84*, 661-669.

Peterson, C. (1996). The preschool child witness: Errors in accounts of traumatic injury. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *28*, 36-42.

Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology*, *35*, 1493-1506.

357

Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review*, *22*, 370-402.

Peterson, C. (2012). Children's autobiographical memories across the years: Forensic implications of childhood amnesia and eyewitness memory for stressful events. *Developmental Review*, *32*, 287-306.

Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child development*, *67*, 3045-3070.

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and wh-questions. *Law and Human Behavior*, *23*, 539-555.

Peterson, C., & Grant, M. (2001). Forced-choice: Are forensic interviewers asking the right questions? *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *33*, 118-127.

Peterson, C., Grant, V., & Boland, L. (2005). Childhood amnesia in children and adolescents: Their earliest memories. *Memory*, *13*, 622-637.

Peterson, C., Morris, G., Baker-Ward, L., & Flynn, S. (2014). Predicting which childhood memories persist: Contributions of memory characteristics. *Developmental Psychology*, *50*, 439-488.

Peterson, C., Noel, M., Kippenhuck, L., Harmundal, L., & Vincent, C. (2009). Early memories of children and adults: Implications for infantile amnesia. *Cognitive Sciences*, *4*, 65-89.

Peterson, C., Pardy, L., Tizzard-Drover, T., & Warren, K. (2005). When initial interviews are delayed a year: Effect on children's 2-year recall. *Law & Human Behavior*, *29*, 527–541

Peterson, C., & Rideout, R. (1998). Memory for medical emergencies experienced by 1-and 2-year-olds. *Developmental Psychology*, *34*, 1059-1072.

Peterson, C., Wang, Q., & Hou, Y. (2009). "When I was little": Childhood recollections in Chinese and European Canadian grade school children. *Child Development*, *80*, 506-518.

Peterson, C., Warren, K. L., & Short, M. M. (2011). Infantile amnesia across the years: A 2-year follow-up of children's earliest memories. *Child Development*, *82*, 1092-1105.

Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology*, *15*, S7-S24.

Pillemer, D. B., Picariello, M. L., & Pruett, J. C. (1994). Very long-term memories of a salient preschool event. *Applied Cognitive Psychology*, *8*, 95-106.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology*, *19*, 291-312.

358

Pipe, M. E., Lamb, M. E., Orbach, Y., Stewart, H., Sternberg, K. J., & Esplin, P. (2007). Factors associated with nondisclosure of suspected abuse during forensic interviews. In M. E. Pipe, M. E. Lamb, Y. Orbach & A. C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay and denial* (pp. 77-96). N. Y: Routledge.

Pipe, M. E., Thierry, K. S., & Lamb, M. E. (2006). The development of event memory: Implications for child witness testimony. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), Handbook of eyewitness psychology, Vol 1: Memory for events (pp. 447–472). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.

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,* 179-190.

Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: Developmental effects of child abuse and neglect. *Developmental Psychology, 36,* 679-688.

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, 99,* 9072-9076.

Pollak, S. D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology, 38,* 784-791.

Pollio, H. R., & Foote, R. (1971). Memory as a reconstructive process. *British Journal of Psychology, 62,* 53-58.

Pompedda, F., Zappalà, A., & Santtila, P. (2015). Simulations of child sexual abuse interviews using avatars paired with feedback improves interview quality. *Psychology, Crime & Law, 21,* 28-52.

Poole, D. A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review, 32,* 165-180.

Poole, D. A., Bruck, M., & Pipe, M. E. (2011). Forensic interviewing aids: Do props help children answer questions about touching? *Current directions in psychological science, 20,* 11-15.

Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals.* Washington DC: American Psychological Association.

Poole, D. A., & White, L. T. (1993). Two years later: Effect of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology, 29,* 844-853.

Poole, D. A., & White, L. T. (1995). Tell me again and again: Stability and change in the repeated testimonies of children and adults. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, Y. S. Ben-Porath, *Memory and testimony in the child witness. Applied psychology: Individual, social, and community issues, Vol. 1,* (pp. 24-43). Thousand Oaks, CA: Sage Publications.

Porter, C., Lawson, J. S., & Bigler, E. D. (2005). Neurobehavioral sequelae of child sexual abuse. *Child Neuropsychology*, *11*, 203-220.

Powell, M.B. (2000). Interviewing of Aboriginal people. *Australian Police Journal*, *54*, 209–212.

Powell, M. B. (2002). Specialist training in investigative and evidential interviewing: Is it having any effect on the behaviour of professionals in the field? *Psychiatry, Psychology and Law*, *9*, 44-55.

Powell, M. B. (2008). Designing effective training programs for investigative interviewers of children. *Current Issues in Criminal Justice*, *20*, 189-208.

Powell, M. B., Cavezza, C., Hughes-Scholes, C., & Stoové, M. (2010). Examination of the consistency of interviewer performance across three distinct interview contexts. *Psychology, Crime & Law*, *16*, 585-600.

Powell, M.B., Fisher, R.P., & Hughes-Scholes, C.H. (2008). The effect of using trained versus untrained adult respondents in simulated practice interviews about child abuse. *Child Abuse & Neglect*, *32*, 1007–1016.

Powell, M.B., Guadagno, B., & Benson, M. (2016). Improving child investigative interviewer performance through computer-based learning activities. *Policing and Society, 26,* 365-374.

Powell, M. B., & McMeeken, L. (1998). "Tell me about the time when. . .": 9 Golden rules for interviewing a child about a multiple offence. *Australian Police Journal*, *52*, 104–108.

Powell, M. B., Roberts, K. P., Ceci, S. J., & Hembrooke, H. H. (1999).The effects of repeated experience on children's suggestibility. *Developmental Psychology*, *35*, 1462–1477.

Powell, M. B., & Snow, P. C. (2007). Guide to questioning children during the free-narrative phase of an investigative interview. *Australian psychologist*, *42*, 57-65.

Powell, M.B., & Wright, R. (2008). Investigative interviewers' perceptions of the value of different training tasks on their adherence to open-ended questions with children. *Psychiatry, Psychology and Law*, *15*, 272-283.

Powell, M.B., Wright, R., & Clark, S. (2010). Improving the competency of police officers in conducting investigative interviews with children. *Police Practice and Research: An International Journal*, *11*, 211-226.

Pratt, C. (1990). On asking children-and adults-bizarre questions. *First Language, 10*, 167-175.

Price, E. A., Ahern, E. C., & Lamb, M. E. (2016). Rapport-building in investigative interviews of alleged child sexual abuse victims. *Applied Cognitive Psychology*, *30*, 743-749.

Price, H. L., Connolly, D. A., & Gordon, H. M. (2016). Children who experienced a repeated event only appear less accurate in a second interview than those who experienced a unique event. *Law and Human Behavior*, *40*, 362-373.

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/Revue canadienne des sciences du comportement*, *43*, 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*, 1-6.

Purdie, V., & Downey, G. (2000). Rejection sensitivity and adolescent girls' vulnerability to relationship-centered difficulties. *Child Maltreatment*, *5*, 338-349.

Radford, L., Corral, S., Bradley, C., & Fisher, H. L. (2013). The prevalence and impact of child maltreatment and other types of victimization in the UK: Findings from a population survey of caregivers, children and young people and young adults. *Child Abuse & Neglect*, *37*, 801-813.

Ratner, H. H., Foley, M. A., & Gimpert, N. (2000). Person perspectives on children's memory and learning: What do source-monitoring failures reveal? In K. P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 85-114). New York: Taylor & Francis.

Reese, E. (2002). Social factors in the development of autobiographical memory: The state of the art. *Social Development*, *11*, 124-142.

Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive Development*, *8*, 403-430.

Reese, E., Jack, F., & White, N. (2010). Origins of adolescents' autobiographical memories. *Cognitive Development*, *25*, 352-367.

Riggins, T., Cheatham, C. L., Stark, E., & Bauer, P. J. (2013). Elicited imitation performance at 20 months predicts memory abilities in school-aged children. *Journal of Cognition and Development*, *14*, 593-606.

Riggs, S. A., & Kaminski, P. (2010). Childhood emotional abuse, adult attachment, and depression as predictors of relational adjustment and psychological aggression. *Journal of Aggression, Maltreatment & Trauma*, *19*, 75-104.

Roberts, K. P. (2002). Children's ability to distinguish between memories from multiple sources: Implications for the quality and accuracy of eyewitness statements. *Developmental Review*, *22*, 403-435.

Roberts, K. P., & Blades, M. (1998).The effects of interacting with events on children's eyewitness memory and source monitoring. *Applied Cognitive Psychology*, *12*, 489–503.

Roberts, K. P., & Blades, M. (1999).Children's memory and source monitoring for real-life and television events. *Journal of Applied Developmental Psychology*, *20*, 575–596.

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). Chichester, UK: Wiley-Blackwell.

Roberts, K. P., & Lamb, M. E. (1999). Children's responses when interviewers distort details during investigative interviews. *Legal and Criminological Psychology*, *4*, 23-31.

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,* 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/Revue canadienne des sciences du comportement*, *45*, 1-11.

Rodriguez-Srednicki, O. (2002). Childhood sexual abuse, dissociation, and adult self-destructive behavior. *Journal of Child Sexual Abuse*, *10*, 75-89.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803-814.

Rogosch, F. A., Cicchetti, D., & Aber, J. L. (1995). The role of child maltreatment in early deviations in cognitive and affective processing abilities and later peer relationship problems. *Development and Psychopathology*, *7*, 591-609.

Rosenbaum, M. E., & Kreiter, C. (2002). Teaching delivery of bad news using experiential sessions with standardized patients. *Teaching and Learning in Medicine*, *14*, 144-149.

Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, *104*, 467-498.

Rubin, D. C., & Umanath, S. (2015). Event memory: A theory of memory for laboratory, autobiographical, and fictional events. *Psychological Review*, *122*, 1-23.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734-760.

Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology*, *27*, 527-538.

Russell, J., Cheke, L. G., Clayton, N. S., & Meltzoff, A. N. (2011). What can What–When–Where (WWW) binding tasks tell us about young children's episodic foresight? Theory and two experiments. *Cognitive Development*, *26*, 356-370.

Russell, J., & Hanna, R. (2012). A minimalist approach to the development of episodic memory. *Mind & Language*, *27*, 29-54.

Samra, J., & Yuille, J. C. (1996). Anatomically-neutral dolls: Their effects on the memory and suggestibility of 4-to 6-year-old eyewitnesses. *Child Abuse & Neglect*, *20*, 1261-1272.

Sattler, J.M. (1992). *Assessment of Childen: WISC-III and WPPSI-R supplement, 3rd ed.* San Diego: Jerome M. Sattler, Inc.

Saywitz, K., & Camparo, L. (1998). Interviewing child witnesses: A developmental perspective. *Child Abuse & Neglect*, *22*, 825-843.

Saywitz, K., & Camparo, L. B. (2014). Using the Developmental Narrative Elaboration interview to elicit children's statements. In L. B. Camparo & K. Saywitz, *Evidence-based child forensic interviewing: The Developmental Narrative Elaboration interview* (pp. 99-147). Oxford University Press.

Saywitz, K., Camparo, L. B., & Romanoff, A. (2010). Interviewing children in custody cases: Implications of research and policy for practice. *Behavioral Sciences & the Law, 28,* 542-562.

Saywitz, K. J., & Geiselman, R. E. (1998). Interviewing the child witness: Maximizing completeness and minimizing error. In S. Lynn & K. McConkey (Eds.), *Truth in memory* (pp. 190-223). New York: Guilford press.

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*, 682-691.

Saywitz, K. J., & Snyder, L. (1996). Narrative elaboration: Test of a new procedure for interviewing children. *Journal of Consulting and Clinical Psychology*, *64*, 1347-1357.

Saywitz, K. J., Snyder, L., & Nathanson, R. (1999). Facilitating the communicative competence of the child witness. *Applied Developmental Science*, *3*, 58-68.

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory in 3-and 4-year-old children. *Developmental Psychobiology*, *55*, 125-132.

Scarf, D., Smith, C., & Stuart, M. (2014). A spoon full of studies helps the comparison go down: A comparative analysis of Tulving's spoon test. *Frontiers in psychology*, *5, 893-898.*

Schacter, D. L. (1987). Memory, amnesia, and frontal lobe dysfunction. *Psychobiology*, *15*, 21-36.

Schaeffer, P., Leventhal, J. M., & Asnes, A. G. (2011). Children's disclosures of sexual abuse: Learning from direct inquiry. *Child Abuse & Neglect*, *35*, 343-352.

Schneider, L., Price, H. L., Roberts, K. P., & Hedrick, A. (2011). Children's episodic and generic reports of alleged abuse. *Applied Cognitive Psychology*, *25*, 862–870.

Schreiber, N., Bellah, L. D., Martinez, Y., McLaurin, K. A., Strok, R., Garven, S., & Wood, J. M. (2006). Suggestive interviewing in the McMartin Preschool and Kelly Michaels daycare abuse cases: A case study. *Social Influence*, *1*, 16-47.

Scottish Executive. (2003). *Guidance Interviewing Child Witnesses and Victims in Scotland*. Edinburgh: Author.

Scottish Executive. (2011). *Guidance on Joint Investigative Interviewing of Child Witnesses in Scotland*. Edinburgh: Author.

Scottish Executive. (2014). *National Guidance for Child Protection in Scotland*. Edinburgh: Author.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, 11-21.

Sedlak, A. J., & Broadhurst, D. D. (1996).The national incidence study of child abuse and neglect. *Washington DC: US Department of Health and Human Services*, 8730763.

Shields, A., & Cicchetti, D. (1998). Reactive aggression among maltreated children: The contributions of attention and emotion dysregulation. *Journal of Clinical Child Psychology*, *27*, 381-395.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience & Biobehavioral Reviews*, *34*, 1080-1091.

Shipman, K., Edwards, A., Brown, A., Swisher, L., & Jennings, E. (2005). Managing emotion in a maltreating context: A pilot study examining child neglect. *Child Abuse & Neglect*, *29*, 1015-1029.

Shipman, K. L., & Zeman, J. (2001). Socialization of children's emotion regulation in mother–child dyads: A developmental psychopathology perspective. *Development and Psychopathology*, *13*, 317-336.

Shipman, K., Zeman, J., Penza, S., & Champion, K. (2000). Emotion management skills in sexually maltreated and nonmaltreated girls: A developmental psychopathology perspective. *Development and Psychopathology*, *12*, 47-62.

Siegman, A. W., & Reynolds, M. (1984). The facilitating effects of interviewer rapport and the paralinguistics of intimate communications. *Journal of Social and Clinical Psychology*, *2*, 71-88.

Simcock, G., & Hayne, H. (2002). Breaking the barrier? Children fail to translate their preverbal memories into language. *Psychological Science*, *13*, 225-231.

Simcock, G., & Hayne, H. (2003). Age-related changes in verbal and nonverbal memory during early childhood. *Developmental Psychology*, *39*, 805-814.

Sluzenski, J., Newcombe, N., & Ottinger, W. (2004). Changes in reality monitoring and episodic memory in early childhood. *Developmental Science*, *7*, 225-245.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, *82*, 171-177.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, *93*, 13515-13522.

Sroufe, L. A., Egeland, B., Carlson, E. A., & Collins, W. A. (2009). *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*. New York: Guilford Press.

Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001a). The memorandum of good practice: Theory versus application. *Child Abuse & Neglect*, *25*, 669-681.

Sternberg, K. J., Lamb, M. E., Esplin, P. W., & Baradaran, L. P. (1999). Using a scripted protocol in investigative interviews: A pilot study. *Applied Developmental Science*, *3*, 70-76.

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*, 439-451.

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Yudilevitch, L., Orbach, Y., Esplin, P. W., &Hovav, M. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse & Neglect*, *21*, 1133-1146.

Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001b). 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*, 997-1005.

Stevenson, K. M., Leung, P., & Cheung, K. F. M. (1992). Competency-based evaluation of interviewing skills in child sexual abuse cases. *Social Work Research and Abstracts*, *28*, 11-16.

Styron, T., & Janoff-Bulman, R. (1997). Childhood attachment and abuse: Long-term effects on adult attachment, depression, and conflict resolution. *Child Abuse & Neglect*, *21*, 1015-1023.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*, 299-313.

Suddendorf, T., Nielsen, M., & Von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science*, *14*, 26-33.

Sullivan, L. A., Kirkpatrick, S. W., & MacDonald, P. M. (1995). Interpretations of facial expressions of emotion by sexually abused and non-abused girls. *Journal of Child Sexual Abuse*, *4*, 45-62.

Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2002). Children's conceptual knowledge of lying and its relation to their actual behaviors: Implications for court competence examinations. *Law and Human Behavior*, *26*, 395-415.

Tarter, R. E., Hegedus, A. M., Winsten, N. E., & Alterman, A. I. (1984). Neuropsychological, personality, and familial characteristics of physically abused delinquents. *Journal of the American Academy of Child Psychiatry*, *23*, 668-674.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P., & Kim, D. M. (2003). The neurobiological consequences of early stress and childhood maltreatment. *Neuroscience & Biobehavioral Reviews*, *27*, 33-44.

Teisl, M., & Cicchetti, D. (2008). Physical abuse, cognitive and emotional processes, and aggressive/disruptive behavior problems. *Social Development*, *17*, 1-23.

Teisl, M., Rogosch, F.A., Oshri, A. & Cicchetti, D. (2012) Differential expression of social dominance as a function of age and maltreatment experience. *Developmental Psychology*, *48*, 575-588.

Teoh, Y. S., & Lamb, M. E. (2010). Preparing children for investigative interviews: Rapport-building, instruction, and evaluation. *Applied Developmental Science*, *14*, 154-163.

Teoh, Y. S., Yang, P. J., Lamb, M. E., & Larsson, A. S. (2010). Do human figure diagrams help alleged victims of sexual abuse provide elaborate and clear accounts of physical contact with alleged perpetrators? *Applied Cognitive Psychology*, *24*, 287-300.

Tessler, M., & Nelson, K. (1994). Making memories: The influence of joint encoding on later recall by young children. *Consciousness and Cognition*, *3*, 307-326.

Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive interviewing and the accuracy of children's accounts. *Law and Human Behavior*, *21*, 405-426.

Thoresen, C., Lønnum, K., Melinder, A., & Magnussen, S. (2009). Forensic interviews with children in CSA cases: A large-sample study of Norwegian police interviews. *Applied Cognitive Psychology*, *23*, 999-1011.

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*, 3–34.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). London: Academic.

Tulving, E. (1985). *Elements of episodic memory*. Oxford University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual review of psychology*, *53*, 1-25.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human. In H. S. Terrace & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3-56). U.S.: Oxford University Press.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, *8*, 198-204.

Tustin, K., & Hayne, H. (2010). Defining the boundary: age-related changes in childhood amnesia. *Developmental Psychology*, *46*, 1049-1061.

Tulving, E. (2005). Episodic memory and autonoesis: Uniquely human. In. H. S. Terrace & J. Metcalfe (Eds.),*The missing link in cognition: Origins of self-reflective consciousness* (pp. 3-56). New York: Oxford University Press.

Unger, J. A. M., & De Luca, R. V. (2014). The relationship between childhood physical abuse and adult attachment styles. *Journal of Family Violence*, *29*, 223-234.

Usher, J. A., & Neisser, U. (1993). Childhood amnesia and the beginnings of memory for four early life events. *Journal of Experimental Psychology: General*, *122*, 155-165.

Utsunomiya, H., Takano K., Okazaki, M., & Mitsudome, A. (1999). Development of the temporal lobe in infants and children: Analysis by MR-based volumetry. *American Journal of Neuroradiology*, *20*, 717-723.

Valentino, K., Cicchetti, D., Rogosch, F. A., & Toth, S. L. (2008). True and false recall and dissociation among maltreated children: The role of self-schema. *Development and Psychopathology*, *20*, 213-232.

Valentino, K., Cicchetti, D., Toth, S. L., & Rogosch, F. A. (2011). Mother–child play and maltreatment: A longitudinal analysis of emerging social behavior from infancy to toddlerhood. *Developmental Psychology*, *47*, 1280-1294.

Van Abbema, D., & Bauer, P. (2005). Autobiographical memory in middle childhood: Recollections of the recent and distant past. *Memory*, *13*, 829-845.

Van Roestel, R. (2016). Relationship between Theory-of-Mind development and children's response to pre-interview instructions. . *Undergraduate honors theses*, *Paper 50*.

Volpini, L., Melis, M., Petralia, S., & Rosenberg, M. D. (2016). Measuring children's suggestibility in forensic interviews. *Journal of Forensic Sciences*, *61*, 104-108.

Wandrey, L., Lyon, T. D., Quas, J. A., & Friedman, W. J. (2012). Maltreated children's ability to estimate temporal location and numerosity of placement changes and court visits. *Psychology, Public Policy, and Law*, *18*, 79.

Warren, A. R., & Lane, P. (1995). Effects of timing and type of questioning on eyewitness accuracy and suggestibility. In M. S. Zaragoza, J. R. Graham, G. C. N. Hall, R. Hirschman, Y. S. Ben-Porath, (Eds.), *Memory and testimony in the child witness. Applied psychology: Individual, social, and community issues* (pp. 44-60). Thousand Oaks, CA: Sage Publications.

367

Warren, A. R., Woodall, C. E., Hunt, J. S., & Perry, N. W. (1996). "It sounds good in theory, but...": Do investigative interviewers follow guidelines based on memory research? *Child Maltreatment*, *1*, 231-245.

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*, 128-135.

Waterman, A. H., Blades, M., & Spencer, C. (2000). Do children try to answer nonsensical questions? *British Journal of Developmental Psychology*, *18*, 211-225.

Waterman, A. H., Blades, M., & Spencer, C. (2004). Indicating when you do not know the answer: The effect of question format and interviewer knowledge on children's 'don't know' responses. *British Journal of Developmental Psychology*, *22*, 335-348.

Weiss, B., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1992). Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. *Child Development*, *63*, 1321-1335.

Welch-Ross, M. K. (1999). Preschoolers' understanding of mind: Implications for suggestibility. *Cognitive Development*, *14*, 101-131.

Welch-Ross, M. K., Diecidue, K., & Miller, S. A. (1997). Young children's understanding of conflicting mental representation predicts suggestibility. *Developmental Psychology*, *33*, 43–53.

Westcott, H. L., & Kynan, S. (2006). Interviewer practice in investigative interviews for suspected child sexual abuse. *Psychology, Crime & Law*, *12*, 367-382.

Westcott, H. L., Kynan, S., & Few, C. (2006). Improving the quality of investigative interviews for suspected child abuse: A casestudy. *Psychology, Crime & Law*, *12*, 77-96.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*, 409-415.

Woodward, C. A. (1998). Standardized patients: a fixed-role therapy experience in normal individuals. *Journal of Constructivist Psychology*, *11*, 133-148.

Woon, F. L., & Hedges, D. W. (2008). Hippocampal and amygdala volumes in children and adults with childhood maltreatment-related posttraumatic stress disorder: A meta-analysis. *Hippocampus*, *18*, 729-736.

Woon, F. L., Sood, S., & Hedges, D. W. (2010). Hippocampal volume deficits associated with exposure to psychological trauma and posttraumatic stress disorder in adults: a meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *34*, 1181-1188.

Yi, M., Jo, E., & Lamb, M. E. (2016). Assessing the effectiveness of NICHD protocol training focused on episodic memory training and rapport-building: A study of Korean police officers. *Journal of Police and Criminal Psychology*, 1-10.

Yi, M., Lamb, M. E., & Jo, E. (2015). The quality of Korean police officers' investigative interviews with alleged sexual abuse victims as revealed by self-report and observation. *Journal of Police and Criminal Psychology*, *30*, 274-281.

Yii, 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*, 270-281.

Young, J. C., & Widom, C. S. (2014). Long-term effects of child abuse and neglect on emotion processing in adulthood. *Child Abuse & Neglect*, *38*, 1369-1381.

Yuille, J. C., Cooper, B. S., & Hervé, H. F. (2009). The step-wise guidelines for child interviews: The new generation. *Pedoparafile: Psychological perspectives, forensic psychiatric.* Italy: Franco Angeli.

Yuille, J. C., Hunter, R., Joffe, R., & Zaparniuk, J. (1993). Interviewing children in sexual abuse cases. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving children's testimony* (pp. 95-115). New York: Guilford Press.

Zola-Morgan, S. M., & Squire, L. R. (1990). The primate hippocampal formation: Evidence for a time-limited role in memory storage. *Science*, *250*, 288-29.