The effect of question repetition on young children's eyewitness testimony

Sarah Joanne Krähenbühl

PhD

Department of Psychology

September, 2007

Acknowledgements

I have many people to thank personally both for their inspiration, support, encouragement and practical assistance but also nameless others without whom the research described in this thesis would not have taken place.

I would like to like to acknowledge my enormous admiration for the Open University, an organisation that enabled me to follow and develop my interests in taking a Psychology degree whilst continuing to work as a teacher. By providing this opportunity to study in a manner that was in accordance with my family life I was not only able to indulge my interests but also regained the confidence and tools required for academic success.

My thanks go to the Psychology department at Sheffield University who did not object when I had the temerity to tell them what I wanted to do, how, and why it was important that they allowed me to do it. Indeed, they were very supportive and helped to make my ideas become reality by gaining funding for my research from the ESRC.

I would also like to acknowledge the generosity of Dr. Helen Westcott and colleagues for allowing my access to their collection of police transcripts, which formed the basis of my research.

There are individual people involved in my research who have been vital for its completion. I would like to thank Chris and Peter Dodd for their professionalism and flexibility in their roles as confederates. I would also like to thank the children, teachers and head teachers of the following schools for their interest and welcome despite the disruption caused by my research activities: Leek First, All Saints' First, Westwood First, Beresford Memorial First,

Woodcroft First, St. Edward's First, St. Anne's Primary, Endon Hall Primary, and Wetley Rocks Primary.

I would like to thank my supervisors Dr. Mark Blades and Prof. Christine Eiser for their assistance. I would especially like to thank Mark for his constant support, his advice and willingness to allow me to follow my own ideas whilst keeping my enthusiasm within practical boundaries. In this time as my supervisor he has become a friend as well as a colleague, a relationship that I hope will continue in the years to come.

Finally, and most importantly, I would like to thank my parents, my children and my husband. My parents, Maggie and Ken Hunt have been unwavering in their absolute confidence and enthusiasm in my activities. My children, Rebecca, Marcus and Miriam have kept my feet firmly on the ground and ensured that I keep a sense of what really matters. As to my ever tolerant husband, fearsome reviewer, and excel guru Christoph; without his kindness, generosity, incredible patience and delight in playing 'devil's advocate' I don't think this thesis would have been started, let alone completed.

Abstract

Children who have been the victims of crime will usually be interviewed by police officers. Current interviewing guidelines warn against repeating questions, because children may interpret the repetition to mean that their first response was incorrect and therefore change their response. Previous researchers have not investigated the ways police interviewers use repeated questions. Given the guidelines we expected repeated questions to be rare. In Study 1 we analysed 95 police interviews with children aged 4-11 alleging abuse. Almost all contained repetition, and on average repeated questions accounted for a quarter of all questions asked. Repetitions led to changes in 75% of children's responses (55% were novel responses, 20% extended the original information elicited). We identified four principal question repetition styles used in police interviews: verbatim, gist, open questions repeated as closed, and closed questions repeated as open.

In Studies 2, 3, 4 and 5 we interviewed children aged 4-5, 6-7 and 8-9 about a staged event they had witnessed earlier (Studies 2, 3 and 4), or about an activity in which they had participated (Study 5). In these studies we varied the type and number of repetitions. We also varied the delay between repetitions and between the event and the interview. The children's responses were assessed for accuracy and consistency.

The number of accurate responses increased with age but decreased with repetition. Repetitions led to changes in approximately 25% of responses. The number of changed responses decreased with age and differed depending on whether the question was answerable or unanswerable. Most changes in responses led to a further inaccurate response (after an original inaccurate

response), or resulted in accurate responses becoming inaccurate. We did not find any pattern of repetition, or type of repeated question that consistently enhanced accuracy. The implications of these results for interviewing practices are discussed.

Contents

Chapter 1

A single question can have many implications and meanings depending on the context and the expectations of the parties involved. In such exchanges questions are asked in many different forms and situations; the context and manner of the presentation of the question will, in addition to the actual words used, define the purpose of (and to some extent the response to) that question.

This thesis will examine the process of asking questions and eliciting responses. A review of related research literature, examine associated theories, and discuss real-life interviewing practices will be included. General background information on conversational mechanisms, associated cognitive developments in language, communication and memory will be introduced. This will be followed by a systematic focus on the key issues surrounding the use and effect of question repetition on event recall performance: question format, position, type (answerable or unanswerable) and quantity, the effect of age, interview timing, interview instructions and experience of interviewee as observer or participant.

The information presented, whilst often applicable to adults, will be limited to children (defined by the United Nations, 1989 as any person under the age of 18). There will be an emphasis on information relating to children between 4 and 11 years of age.

The process of asking a question and the expectation of a response as an example of communication involves interaction between people with an underlying pragmatic purpose. Grice (1989) suggested that this interaction was characterised by the cooperative principle whereby effective communication was

promoted by both speakers and receivers. The cooperative principle assumes that the utterances presented are true, informative, relevant and clear (the four Gricean maxims). However, in addition to the overt explicit meaning of an utterance, the speaker can be assumed to have covertly proposed or implied other meanings (implicatures).

Communicative interactions involving the use of questions therefore exemplify or involve more than a simple exchange of information. The person asking the question makes a demand on the receiver to provide information according to their requirement. This requirement may overtly have provided information but the implicature of the request is also involved and may have affected the response given.

Asking a child questions has been one of the most common methods used to establish, analyse, challenge, query, enquire and gather information. For example, questions were used by Piaget to assess cognitive development (1956, 1952). Piaget was aware of difficulties associated with asking children questions and stated that "suggestion may thus be avoided by means of patience and analysis" but also acknowledged that there was difficulty in "the point to be regarded as the child's original contribution and that due to previous adult influences" (1929, pp. 27-28). The responses given by children to questions asked by adults have informed the designation of developmental attributes. For example, responses to questions led to Piaget's description of children under approximately 7 years of age as being 'non-conservers' (1952). However, Siegal (1991) and Donaldson (1978) argued that this label and these descriptions of developmental abilities resulted from the repeated questioning methodology used

in assessment rather than from a manifestation of children's limited understanding.

Another aspect of cognitive development, which affected children's responses to questions, was the shared understanding of conversational rules. Siegal suggested "many questions that require children to make judgements about their world inadvertently violate the rules of ordinary conversation" (1991, p. 122). An imbalance of power and knowledge between the child and the adult may be recognised by both parties (Freeman, Sinha, & Condliffe, 1981) and have an effect both on the way in which questions are asked or repeated and on the responses given to those questions.

The implicature of a question (Grice, 1989) may not necessarily represent a shared understanding of what is being asked. One instance where this may occur is when a question is repeated, especially when the reasons for repetition are not made explicitly clear. The effect of the repetition in this instance may have a differential effect depending on the level of language development and communication experience of the parties involved, as is the case when adults interview children. The changes in responses made by children as a result of question repetition may occur for one or more of the following reasons suggested by Siegal (1991): that the child believed their initial response was unacceptable and needed to be altered, that the child was concerned that their actual response would result in an unfavourable perception, that the child believed an age- or gender-typical response was required, or that the language used was misunderstood.

With the repetition of questions, there may be uncertainty or disparity as to the underlying motives and expectations of all parties involved through the

use of a particular conversational strategy. Such discrepancies could account for an apparently poor level of performance (Goodnow, 1984), or the attempt of children aged between 5- and 8-years-old to answer completely bizarre or nonsensical questions (Hughes & Grieve, 1980; Pratt, 1990; Waterman, Blades, & Spencer, 2000).

Children's cognitive development has been determined through interplay between both conversational and conceptual processes (Siegal, 1999). In order to respond to questions about an experienced event a child not only has to take part in the conversational structure but also has to utilise his or her memories and recall processes. Researchers have suggested that memory is not a unitary process but consists of multiple systems that interact in the encoding, storage and retrieval of information (Cordón, Pipe, Sayfan, Melinder, & Goodman, 2004). These systems may not be equally developed in all children but depend on their age, maturity and experience.

The form of memory typically required in recall of an experimental event, or in eyewitness testimony concerning a specific event is 'autobiographical episodic' memory (Nadel, 1994). This form of memory incorporates facts and events that are linked to a particular occurrence with significance for the person involved (Gordon, Baker-Ward, & Ornstein, 2001). Tulving (cited by Goodman & Melinder, 2007) suggested that children did not develop episodic memory until approximately 4-years-old, and that therefore younger children could not consciously think back to specific past events before this age. However, other researchers have established that children as young as 2-years-old are capable of giving accurate responses to questions requiring recall of specific information (Fivush, Haden, & Salimah, 1995) about experienced

events (for reviews of early development see Fivush, 1993; Gathercole, 1998). Dent (1992) found that accuracy and completeness of recall of children aged 8-12-years-old was at a level comparable to that of adults. Between the ages of 3 and 10 years accuracy of recall in response to questions increases with age both in experimental contexts (Beuscher & Roebers, 2005; Cassel, Roebers, & Bjorklund, 1996; Geddie, Fradin, & Beer, 2000; Gobbo, Mega, & Pipe, 2002; Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Poole & White, 1991; Powell & Thomson, 1996), and after naturally occurring events (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Goodman, Quas, Batterman-Faunce, Riddlesberger, & Kuhn, 1994).

Although trends generally showed an increase in accuracy with age there are also some anomalies. For example, Cassel et al. (1996, example questions p. 121) found that the accuracy of free recall questions or invitations ("Tell me everything you can remember about what happened in the video you saw last week") for 9-10-year-olds was approximately half that of adults. With misleading ("The mother owned the bike, didn't she" – when she did not) and multiple choice questions ("Did the bicycle belong to: the mother, the boy, or the girl?") however, the 9-10-year-olds performed almost at the level of adults. For 5-6-year-olds accuracy to misleading questions was one fifth of that for adults and half adult accuracy in free recall. In contrast adults accuracy for 'positive-leading' questions ("You remember that the girl owned the bike, don't you?" – asked when the initial response to this question was incorrect or had elicited "don't know" type response) was less than that of all of the children (regardless of age) tested. This discrepancy in relative accuracy between age groups was also found in other results of variable interviewing practices. Poole and White

(1991) showed that accuracy increased with age although 8-year-olds' performance in responses to diverse questions was less than half the accuracy of adults.

Research on accuracy of recall is not limited to experimental situations; non-experimental based studies have also confirmed children's abilities to provide accurate information in response to questions. Bahrick, Parker, Fivush and Levitt (1998) found that children aged 3-4-years-old's accounts of a natural disaster were extensive and included (according to parental assessment) less than a tenth inaccuracies. Furthermore, in a case study concerning an allegation of sexual abuse, comparison of a tape recording and a subsequent interview enabled corroboration of over half of the information provided by the child (aged 13 at the time of disclosure) involved (Orbach & Lamb, 1999).

An evaluation of the accuracy of a response may not be limited to a single occasion but be attributed to whether a particular response is maintained over time when questions are repeated or whether the responses change. Changed responses in interviews, regardless of inherent accuracy, represent inconsistency (Ghetti, Goodman, Eisen, Qin, & Davis, 2002; Quas, Davis, Goodman, & Myers, 2007; Siegal, Waters, & Dinwiddy, 1988). The term 'consistency' refers to whether, over time, the same information is reported in semantic (in meaning), if not verbatim form (expressed in the same words). Inconsistency does not necessarily imply contradiction; information elicited may be novel but related to previous information or may be totally unrelated. For example, if, on two separate occasions, a child was asked "And what happened then?" to which the child replied the first time "I was crying", and the second time "I ran home", this would denote an inconsistent response, but not

necessarily contradictory or conflicting, information. Thus inconsistent responses can be accurate, may both (or all) be inaccurate, or else there may be only one accurate response with the other remaining responses being inaccurate.

Researchers have shown that consistency in responses increases between the 3- and 16-years-old (Ghetti et al., 2002; Gobbo et al., 2002; Poole & White, 1991, 1993; Powell & Thomson, 1996; Quas & Schaaf, 2002; Roebers & Schneider, 2002). Poole and White (1991) found that amongst children in single and repeated interviews 4-year-olds were less consistent than either 6- or 8-year-olds. In contrast with age related improvements in consistency, Poole and White also found that adult consistency was lower than that of 6- and 8-year-olds. This result may, however, have been obtained through the response to a question about possible occupation, about which adults, given their experience of the work place, may have been more willing to speculate. Both Roebers and Schneider (2002) and Powell and Thomson (1996) found that consistency doubled between 4-5 and 6-8-year-olds but Roebers and Schneider did not find a difference in consistency between 8- and 10-year-olds. Consistency increasing with age was also found in reports of sexual abuse (Ghetti et al., 2002) where an effect of age in consistency of reports between 3-5, 6-10, and 11-16-year-olds almost reached significance ($p < .07$).

The level of consistency has been generally measured according to the number of changes made in accuracy totals of responses to repetitions. Howie, Sheehan, Mojarrad and Wrzesinska (2004) introduced an examination of the directions, in terms of accuracy and inaccuracy, of changes made in responses. "Desirable" shifts were defined as those changes where an inaccurate response became accurate after repetition of the question, and "Undesirable" shifts were

those accurate responses that became inaccurate after repetition. In response to repeated questions, children aged 4-5 changed their responses more than twice as frequently than children aged 7-8. The majority of changes were undesirable, although provision of a rationale to explain the reason for the repetition reduced the probability of undesirable changes in both age groups.

The studies mentioned above have indicated an increase in accuracy and consistency with age; they also reflected the diversity in questioning strategies that involved different interviewing timescales, repetition of questions between and within interviews, different quantities, formats and patterns of repetition. All of these factors, which may have influenced children's accuracy and consistency, will be discussed in detail.

Before examining the effect of question formats on children's responses it is necessary to define these terms as their usage is not always consistent in research literature. The form of question asked may have an effect on the amount of information elicited in the response. The main distinction in question formats is between open-ended questions where a response is not limited and closed questions where the response is limited by a range of available options (for examples see Fritzley & Lee, 2001; Home Office, 1992, 2001; Memon & Vartoukian, 1996; Peterson & Biggs, 1997; Peterson, Dowden, & Tobin, 1999; Pipe, Sutherland, Webster, Jones, & La Rooy, 2004; Waterman, Blades, & Spencer, 2001, 2004) including forced-choice options (Gee, Gregory, & Pipe, 1999; Powell & Roberts, 2002; Roebers, Gelhaar, & Schneider, 2004).

Within open-ended questions there are gradations of 'open-ness'. A question that asks the interviewee to "tell me everything" would be described as an 'open invitation' or 'free recall' as there is no direction given even as to the

required subject matter (for example Beuscher & Roebers, 2005; Cordón, Saetermoe, & Goodman, 2005; Gobbo et al., 2002; Hershkowitz, 2001; Howie et al., 2004; Memon & Vartoukian, 1996; Orbach & Lamb, 2000; Warren, Hulse-Trotter, & Tubbs, 1991). Other open-ended questions, those that start with "wh-" ("who, what, when, why or which") or "how" are often defined as 'specific questions' as they direct the interviewee towards the subject matter that the interviewer wishes to discuss without specifying the response required (for example Beuscher & Roebers, 2005; Gee et al., 1999; Holliday, 2003b; Howie et al., 2004; Orbach & Lamb, 2000; Peterson et al., 1999; Pipe et al., 2004; Warren et al., 1991).

Closed questions have a limited range of alternative responses; in effect, they provide the choices of response available from which the interviewee can select. One form of closed question requires a "Yes" or "No" response. There are also forced choice questions asking "Was it ... or ...". This form of question is not necessarily deemed suggestive if it contains all possibilities, or a number of probable or equally likely alternatives (Home Office, 2001). In relation to gathering information about a suspect's appearance the question "Tell me everything that you remember" would be considered an open-ended invitation or free recall. "What was he wearing?" provided an open-ended specific version, and "Was his shirt red or green?" would be defined as closed as the choice of response was provided by the question.

Experimental research has provided evidence for the superiority of open invitations and specific questions above closed questions in eliciting accurate and more detailed information in responses (Dent & Stephenson, 1979; Fivush, 1993; Hutcheson, Baxter, Telfer, & Warden, 1995; Peterson et al., 1999;

Steward, Bussey, Goodman, & Saywitz, 1993) including interviews with children with developmental handicaps (Perlman, Ericson, Esses, & Isaacs, 1994). A caveat however, is that young children aged between 3- and 6-years-old do not spontaneously provide a great deal of information in free recall or in response to specific questions (Ceci & Bruck, 1995; Dent & Stephenson, 1979; Fivush, 1993; Fivush, Peterson, & Schwarzmueller, 2002). Nevertheless, if accuracy was assessed as a proportion of what was actually said then the accuracy of young children age between 3- and 6-years-old was found to be on a par with other age groups (Cole & Loftus, 1987; unless coached, see Poole & Lindsay, 1995).

Research on forensic transcripts of investigative interviews and post-medical interviews has also shown the positive effects of using free recall and specific questions in eliciting details and the problems associated with closed questions (Cronch, Viljoen, & Hansen, 2006; Garven, Wood, Malpass, & Shaw, 1998; Lamb, Sternberg, & Esplin, 1998; Peterson & Biggs, 1997; Peterson, Pardy, Tizzard-Drover, & Warren, 2005; Shrimpton, Oates, & Hayes, 1998). Free recall and specific questions used in investigative interviews with children aged between 4- and 15-years-old elicited most narrative details (Orbach & Lamb, 2000) regardless of the age of the child interviewed or the stage of the interview (Hershkowitz, 2001). Age differences were also absent from the proportion of details (approximately half of the total) elicited from the use of free recall although younger children (aged 4-years-old) provided fewer total details than older children aged 8-years-old (Lamb et al., 2003a).

The application of different question formats had an effect on the consistency of children's (aged between 4- and 9-years-old) reports (Lamb &

Fauchier, 2001; Orbach & Lamb, 2001). Lamb and Fauchier (2001) examined 24 forensic interviews given by seven children (average age 6-years-old). All of the children's allegations led to convictions, which suggested that the details they provided were sufficiently credible and were consistent with additional or alternative evidence. Lamb and Fauchier found that, unlike other questioning formats, contradictions in information were not found in responses to open-ended invitations. In contrast to the reliability of open-ended specific questions (Lamb and Fauchier described these as 'directive') used in experimental studies, a third of the responses elicited by specific questions represented contradictions, which accounted for four tenths of the questions asked in the interviews.

The presence or absence of free recall questions also has an effect on children's disclosure. Hershkowitz, Horowitz and Lamb (2005) stated that almost one third of victims (aged between 3- and 15-years-old) interviewed in Israel did not disclose abuse during forensic interviews despite the suspicion that abuse had occurred. In a following study of fifty forensic interviews of children aged 4-13-years-old who did not disclose and fifty interviews of children who did disclose, it was found that fewer open-ended invitations and prompts (described as 'free recall prompts') were used in interviews with children who did not disclose than were used in interviews with children who did disclose (Hershkowitz, Orbach, Lamb, Sternberg, & Horowitz, 2006).

Despite the evidence for the superiority of free recall and specific questions, closed questions continued to be the prevalent questioning format used in forensic interviews of children aged between 3- and 15-years-old in Israel, Norway, Sweden, England and Wales and US (Cederborg, Orbach, Sternberg, & Lamb, 2000; Davies & Westcott, 1999; Davies, Westcott, & Horan,

2000; Korkman, Santtila, & Sandnabba, 2006; Lamb et al., 1996; Sternberg, Lamb, Davies, & Westcott, 2001; Sternberg et al., 1996; Thoresen, Lønnum, Melinder, Stridbeck, & Magnussen, 2006). For example, in these studies between a third and a half of questions in interviews were option-posing or suggestive and were responsible for between two and three fifths of the details elicited (Cederborg et al., 2000; Korkman et al., 2006; Sternberg et al., 2001).

Another aspect of question format relates not to the actual syntactic form of the question but its type depending on the knowledge of the interviewee. A question may be either answerable or unanswerable. The interviewer may already know the accurate response to the question asked and expect the interviewee to know the correct response. Hence the answerable question is asked in order that the interviewee can be heard to present his or her own account of the issue under discussion in his or her own words. The accurate response to a question may also be unknown by the interviewee, and would, in effect, be 'unanswerable'. An accurate response to 'unanswerable' questions is possible if the interviewee response is, "I don't know" (or equivalent). Such behaviour is encouraged in the 'ground rules' presented at the beginning of an investigative interview conducted under Achieving Best Evidence ('ABE') interviewing protocol guidelines (Home Office, 2001)[1] currently in use in England and Wales.

Accuracy of responses to unanswerable questions has been shown to be lower than accuracy for answerable questions for both children and adults

---

[1] The ABE (Home Office, 2001) protocol will be discussed in more detail in Chapter 2.

(Waterman et al., 2001, 2004) although direct comparison may be spurious as these questions cannot be counterbalanced to provide equitable difficulty. Amongst adult participants Pezdek, Sperry and Owens (2007) found that those participants who had voluntarily self-generated information (these participants were told they could say "I don't know" if appropriate) were more likely to remember and repeat that generated information than participants who had been 'forced' to generate information to form responses to questions that were unanswerable. Over time, the participants' confidence in their responses increased. However, the encouragement of children to say "I don't know" (if appropriate) as advocated in the ABE (Home Office, 2001) protocol has not been shown to necessarily have a positive effect on the accuracy of responses given by children aged between 2- and 10-years-old (Fritzley & Lee, 2003; Moston, 1987; Peterson & Grant, 2001).

The reluctance to accurately say "I don't know" to unanswerable questions varies according to the format of the question posed. Children aged between 4- and 9-years-old were more likely to indicate that they did not know the answer to an unanswerable open specific (wh-) questions than to an unanswerable closed questions where response options (i.e. "yes" or "no") were readily available and an alternative response did not need to be generated (Krähenbühl & Blades, 2006; Peterson & Grant, 2001; Waterman et al., 2001, 2004).

As mentioned above, researchers have shown that children are reluctant to admit their ignorance by spontaneously using an "I don't know" response (for exceptions see Moston, 1987; Mulder & Vrij, 1996). Children aged between 4- and 10-years-old maintained their generated responses to questions that they

could not answer even when they had been explicitly encouraged to say "I don't know", or warned that questions may "fool you or include wrong suggestions" (Beuscher & Roebers, 2005, p. 4; Krähenbühl & Blades, 2006; Peterson et al., 1999; Price & Connolly, 2004; Waterman et al., 2000, 2001, 2004). This reluctance to say "I don't know" was found especially with children aged 7- to 10-years-old rated as having low self-esteem (Howie & Dowd, 1996) and occurred in children aged 6- to 8-years-old even when the instruction was repeated (Roebers & Schneider, 2005).

In real-life contexts interviews concerning a single issue containing both answerable and unanswerable questions of differing formats may take place on more than one occasion. In police interviews, children who are witnesses or victims are likely to describe their experiences (both formally and informally) on many occasions to different people. Whitcomb (as cited in Ceci & Bruck, 1995) reported that children are estimated to have undergone an average of 12 'formal' interviews during the course of an investigation.

The purpose of repetition within and between these interviews is not necessarily clear. According to Chambers dictionary, there are 12 different definitions of 'repeat' including: 'to say, do, perform, go over, again: to iterate: to quote from memory: to recount' (A.M.Macdonald, 1981, p. 1144). Repeated interviews may merely be a practical necessity involved in the judicial process. During the time between the first discussion of the relevant event and subsequent retelling of that event it is possible that changes will be made to the related information; new information (either accurate or inaccurate) may be incorporated and previously mentioned information may be forgotten or altered. Through

these repeated interviews, some questions are likely to be repeated by different interviewers regardless of whether this was intentional or a specific objective.

Repeated interviews have been included in experimental research (Cassel & Bjorklund, 1995; Jones & Pipe, 2002; Jones & Powell, 2005; Melnyk & Bruck, 2004; Poole & White, 1993; Quas & Schaaf, 2002; Roebers & Schneider, 2002), and in real-life contexts (Burgwyn-Bailes, Baker-Ward, Gordon, & Ornstein, 2001; Ornstein et al., 2006; Peterson, Moores, & White, 2001; Salmon & Pipe, 2000; Tizzard-Drover & Peterson, 2004). Children aged 4-5-years old are capable of maintaining accurate reports about true events when questioned about the same event on multiple occasions (Fivush & Shukat, 1995) and may even improve their accuracy even when misleading questions have been asked (Goodman et al., 1991, cited in Quas, Goodman, Ghetti, & Redlich, 2000).

Accuracy may also be adversely affected across repeated interviews. For example, in repeated interviews following a routine medical examination accuracy of children aged 2- to 13-years-old decreased with time although accuracy for details relating to the central information of an actual injury was higher than for the peripheral information concerning hospital details (Peterson et al., 2001; Tizzard-Drover & Peterson, 2004). Individual differences, age of children aged 3- to 7-years-old (Burgwyn-Bailes et al., 2001) and the use of props, rather than reliance on verbal reporting, also affected the accuracy of recall in children aged 5-years-old (Salmon & Pipe, 2000).

Researchers found that the introduction of misinformation had an adverse effect on children's accuracy when interviewed on multiple occasions. With the exception of elderly adults, children aged between 7- and 10-years-old were less consistent in their responses than adults (Coxon & Valentine, 1997; Poole &

White, 1993); consistency followed a developmental change with older children aged between 5- and 9-years-old being more consistent and less prone to suggestibility than children aged two years younger (Cassel & Bjorklund, 1995; Cassel et al., 1996; Gobbo, 2000; Quas & Schaaf, 2002; Roebers & Schneider, 2002). The form of interviewing and misinformation introduced also effected the level of subsequent suggestibility; early introduction (within two days) of misinformation with children aged 5-6-years-old (Melnyk & Bruck, 2004; Pipe et al., 2004), the inclusion of stereotypical information with children aged 3-4-years-old (Leichtman & Ceci, 1995), or the use of closed questions with children aged 4-8-years-old (Poole & White, 1991) exacerbated the detrimental effects of suggestibility on children's responses. After introduction of misinformation during repeated interviews a final interview conducted by an unfamiliar interviewer in a non-suggestive manner, who even challenged the child's inclusion of misinformation, did not result in a decrease in the reports of false events even after debriefing for some children aged 6-8-years-old (Erdmann, Volbert, & Böhm, 2004). These results suggested that if a question was unintentionally suggestible because it directed a child to a response (when the correct response should be "I don't know") then the response 'created' by the child could become part of their recollection of that event.

The effects of suggestibility were also evident when children's parents were involved in the interview either directly or indirectly. When a child had a close relationship with a parent the child aged between 3- and 13-years-old was more able to resist the inclusion of misinformation in their subsequent interview (Clarke-Stewart, Malloy, & Allhusen, 2004; Quas, Goodman, Bidrose, Pipe, & Craw, 1999), and children aged 4-years-old were more likely to resist false

suggestions of abuse made by 'others' when later interviewed by a parent (Goodman, Sharma, Thomas, & Considine, 1995).

Regardless of the number or quality of informal or formal interviews the procedure in England and Wales judicial process is that a single video-recorded interview (that provides the child's 'evidence in chief') is carried forward to be presented at the ensuing trial (Home Office, 2001). Previous interviews may have an impact on the information elicited but little is known about the effect of question repetition on accuracy and consistency within that single ABE interview.

The possible adverse effects of repetition cannot be addressed by avoiding repetition altogether as responses given by interviewees to questions are not always adequate or appropriate. When young children aged between 4- and 8-years-old are asked to describe a personal experience, for example, "What did you do at the weekend?" their responses are often brief, uninformative or the question is ignored altogether (Lamb et al., 2003a). As a result, an interviewer who has the responsibility of eliciting a child's responses to questions may need to repeat those questions in order to gain the appropriate level of detail or breadth of description.

Concern regarding a detrimental effect of question repetition on children's accuracy is provided in both ABE and its precursor the Memorandum of Good Practice ('MOGP') interviewing protocols (Home Office, 1992, 2001). These protocols suggest that repetition of questions should not be performed soon after the initial question, nor persistently because, children may assume that their original response was incorrect or inadequate and change their response to one that they believed was required by the interviewer.

The term 'repetition', its specific form, frequency or implementation has not been clearly defined in either interviewing protocols or by researchers but there is a general consensus that repetition may affect responses. Poole and White stated that "Although young children can interpret a repeated question as an implicit request for a response change, there is no reason to believe that requests to repeat testimony will always prompt children to become less accurate" (1991, p. 976). In other words, Poole and White suggested that children may become less accurate some of the time, but also implied that children could also become more accurate. A possible change in accuracy in either direction would need verification in order to ascertain whether the use of repetition was beneficial or detrimental to children's accuracy.

In Poole and White's study (1991) children (aged 4-, 6- and 8-years-old) and adults observed a staged live event concerning a benign dispute between two adults over a pen. The participants were interviewed either immediately or a week later (there was another condition with two interviews which is not described in this review) with a series of seven questions (of specific and closed formats) that were repeated twice (i.e. three blocks of questions). Poole and White found that accuracy in response to repetition in both the immediate interview and delayed interview declined for all children with the exception of a rise in accuracy amongst 4-year-olds in the delayed interview condition. In comparison to original accuracy accurate novel responses were more frequent in the delayed interview. Inaccurate novel responses to repeated questions were more frequently provided by 4-year-olds than other age groups in the immediate interview with fewer in the delayed interview. Poole and White suggested that

"Although there was no explicit pressure to lie... the children may have felt indirectly pressured by the repeated questions..." (1991, p. 984).

An exception to age related changes in responses to repetition was found by Finnilä, Mahlberg, Santilla, Sandnabba and Niemi (2003) who showed that 4-5-year-olds changed their responses to misleading questions less frequently than 7-8-year-olds. Finnilä et al. suggested that the older children were more aware of conversational rules that "dictate that a repeated question means that the previous answer was wrong" (2003, p. 39).

The format of the repeated question also affected both accuracy and consistency of children's responses. The term repetition has been generally used to refer to verbatim repetition, when the same question (which can be made in either open or closed formats) is repeated word for word (for example Gilstrap, 2004; Howie et al., 2004; Memon & Vartoukian, 1996; Poole & White, 1991). For example Poole and White (1991) found that the decline in accuracy of responses to repeated questions was primarily in responses to closed questions. This distinction in accuracy according to question format was also shown by Memon and Vartoukian (1996) who found increased accuracy in responses to repeated open questions but decreased accuracy in responses to closed questions (the results were close to significance). However, as mentioned by Memon and Vartoukian it was not possible to assume that the responses were due to the question format or the difficulty of questions as the open and closed formats were not counterbalanced; the results merely demonstrated how children's responses to question forms varied as a function of repetition.

With the exceptions of Poole and White (1991) and Finnilä et al. (2003) question repetition within single interviews has been limited to a single repetition

(Howie et al., 2004; Memon & Vartoukian, 1996; Moston, 1987; Powell & Thomson, 1996; Warren et al., 1991). Finnilä et al. repeated questions up to three times in a 'high pressure' interview condition although only the last response the child gave was coded. Poole and White repeated their questions twice in all interviews (thus providing three instances of the same question) and found that although overall accuracy declined with repetition the accuracy of 4-5- and 8-9-year-olds improved (relative to the first repetition) with the second repetition in the immediate interview condition. In the single delayed interview condition the accuracy of responses to repeated questions for 4-5-year-olds was highest in the second repetition. It was possible, that through repetition of questions, the younger children were, in effect, 'reminded' of the incident and it was this positive effect that provided cues or outweighed concern with conversational rules (as suggested by Finnilä et al.) which adversely affected the older children's and adult's accuracy. To date, further research using multiple repetitions of questions has not been conducted.

Overt suggestion that an original response to a repeated question was incorrect, or asking the interviewee "are you sure?" resulted in changes in responses in children aged between 4- and 12-years-old when questions were repeated (Candel, Merckelbach, & Muris, 2000; Endres, Poggenpohl, & Erben, 1999; Warren et al., 1991). However, as Howie et al. (2004) showed it was not necessary to imply directly that responses were inadequate, inaccurate or uncertain for a child to change their responses as a result of repetition.

Howie et al. (2004) examined whether a rationale for repetition influenced children's tendency to change ('shift') their responses and the direction that those changes took. After having watched a video showing

interaction between a teacher and children during which a fire drill practice took place children aged 4-5 and 7-8-years-old were interviewed. Half of the children had questions repeated by the same interviewer, half by a different interviewer, and half of the children were given a rationale as to the reason for repetition (that the interviewer had recording difficulties which were overcome later in the interview) and half did not. The results showed that older children changed their responses less and were more resistant to misleading questions than younger children. The provision of a rationale did not affect overall changing of responses although the rationale reduced undesirable shifts (changes from accurate to inaccurate) in younger children and increased desirable shifts (changes from inaccurate to accurate) in both age groups. With younger children the rationale for the use of repetition reduced shifts when the same interviewer asked all of the questions, whereas the reverse applied to older children who reduced shifts with a second interviewer. "I don't know" responses to repetitions were unaffected by rationale.

A change in response may have either a desirable or an undesirable effect on accuracy (Howie et al., 2004), but regardless of that effect a change in response would denote inconsistency. However, the common assumption that consistency equated with accuracy (as found between interviews by Peterson et al., 2001) was shown to be erroneous when children encountered repeated questions (Quas et al., 2007). In Quas et al.'s study children aged between 4 and 7-years experienced an event in which one third of the children were touched innocuously. In one condition children who had not been touched were instructed to lie during a subsequent interview. This interview included repeated questions concerning touch location in addition to other details about which the

children had not been instructed to lie. Children instructed to lie were consistent in maintaining that information in response to repeated questions. The children who had been touched and told the truth were the most inconsistent.

Researchers have established that another factor that influenced the accuracy and completeness of recall of the event was the timing of an initial interview. Experimenters who utilised an immediate interview (within approximately 15 minutes of the event) showed age-related differences in correct recall when accuracy increased with age in children aged between 3- and 12-years-old (Cassel & Bjorklund, 1995; Gobbo et al., 2002; Ornstein et al., 2006; Poole & Lindsay, 1995; Poole & White, 1991; Sutherland & Hayne, 2001).

Sutherland and Hayne (2001) also found large age-related differences in accuracy of recall in children aged 5-12-years-old following an interview one day after an event. These age differences reduced substantially in an initial interview after an interval of six weeks, however, there were no significant difference in amounts of recall (in either free recall or specific questions) between 11-12-year-olds and adults. The 5-6-year-olds provided fewer correct responses regardless of the interval before the interview but their accuracy in free-recall remained relatively constant unlike the reduction in accuracy for the other age groups.

Children do not, however, generally disclose information about an incident of sexual abuse immediately but more commonly disclose more than 48 hours later (Goodman-Brown, Edelstein, Goodman, Jones, & Gordon, 2003). The decision and opportunity to disclose (or not) will vary with each child according to their particular circumstances (London, Bruck, Ceci, & Shuman, 2005), and the way in which the interview was conducted (Hershkowitz et al.,

2006; Korkman et al., 2006). In Goodman-Brown et al.'s study of children aged 4-16-years alleging abuse, two fifths disclosed to police or social services within 48 hours of the latest assault and almost a further fifth disclosed between 48 hours and two weeks later. Children aged between 6- and 18-years-old who were younger at the onset of abuse delayed their disclosure to a greater extent than older children (Sjöberg & Lindblad, 2002). Furthermore, the specific nature of the abuse also had an influence on disclosure delay with a stranger as perpetrator in rape cases associated with earlier disclosure than for inter-familial rape (Smith et al., 2000).

Tizzard-Drover and Peterson (2004) suggested that an interview conducted relatively quickly (within one week of an Accident and Emergency attendance) consolidated the memory, increased accessibility of the original memory by serving as a partial re-exposure or reinstatement of the original event and provided, through systematic questioning, an organisational structure which attenuated forgetting and facilitated recall. An early first interview before a subsequent interview after a delay of one year had a positive effect on accuracy and completeness for 3-4-year-olds but not with older children. In contrast a longer delay of three weeks was not found to have an effect on memory performance of 3-, 5- and 7-year-old children who either experienced or did not experience an immediate interview (Baker-Ward et al., 1993).

The issue of timing has been recognised by governmental authorities and is discussed in the interviewing protocol guidelines (Home Office, 1992, 2001). Reference to a preference for short delays before interviewing is made in the ABE interviewing protocol which states that an interview "take place as soon

after an allegation or referral emerges as is practicable" (Home Office, 2001, p. 23).

The Pigot Report in England and Wales recommended that investigative proceedings involving a child witness should be dealt with "as rapidly as is consonant with the interests of justice" (Home Office, 1989, para. 2.14). The legal profession assumed that, with increased time, children's memories became vulnerable, particularly with children aged 5-6-years-old in comparison with children aged 9-10-years-old (Flin, Boon, Knox, & Bull, 1992a). However, regardless of the delay before interviewing, the Pigot Report recommendation and governmental 'speedy trial' policy to give priority to child abuse or child witness cases, children often wait longer than 12 months before their case comes to trial (Davis, Hoyano, Keenan, Maitland, & Morgan, 1999; Plotnikoff & Woolfson, 2004). There is an average delay before trials in the Crown Court of 11.6 months, 9.9 months for magistrates' court, and 8.6 months for youth courts (Plotnikoff & Woolfson, 2004).

Researchers have not addressed the effect of the timing or presentation pattern of question repetitions within single interviews. In most experimental research the majority of repetitions made are presented as a complete block, where the initial series of questions is subsequently repeated as a block of questions (Howie et al., 2004; Poole & White, 1991; Powell & Thomson, 1996; Warren et al., 1991), or with a single question or comment between repetitions (Finnilä et al., 2003; Moston, 1987).

Different delays between repetitions (without systematic variance) have been used (Goodman, Batterman-Faunce, Schaff, & Kenney, 2002; Memon & Vartoukian, 1996) although each participant received the same pattern of

questions. Memon and Vartoukian (1996) did not examine the effect of repetition order on accuracy or consistency. Goodman et al. (2002) interviewed 7- and 10-year-old children about an event (an unexceptional social interaction with an unfamiliar male confederate) that had occurred four years previously (the original event was described by Goodman & Reed, 1986). A specific question asking for details of the event was repeated six times (in gist form) with varying lengths of delays between the repetitions (each child received the same order of questions) in order to ascertain whether other non-repeated intervening questions, the use of puppets and/or comments would induce suggestibility which would be reflected in the responses to the repeated specific questions. An effect of question repetition was not found but as the children generally failed to remember the initial event (which had taken place four years prior to the current interview) it was not possible to make any conclusions as to the effect of such a repetition strategy on children's accuracy.

The event itself may also have an effect on the accuracy and consistency of children's recall. Events that children are required to recall may concern witnessing or being the victim of neglect, maltreatment, physical, or sexual abuse. Such personally experienced abusive events are likely to be highly stressful and distressing for the child. Research on children's memory recall for stressful events (for a review of the effect of stress on memory see Toth & Cicchetti, 1998) is limited due to ethical constraints. Therefore studies have utilised access to interviews concerning actual allegations of abuse to establish children's memory recall of stressful events (Ghetti et al., 2002), and also recall of medical procedures such as post emergency hospital treatment (Peterson & Bell, 1996; Peterson & Biggs, 1997; Peterson et al., 2005; Peterson & Whalen,

2001), medical treatment (Burgwyn-Bailes et al., 2001; Chen, Zeltzer, Craske, & Katz, 2000; Goodman et al., 1994; Merritt, Ornstein, & Spicker, 1994), inoculations (Alexander et al., 2002), paediatric examinations (Greenhoot, Ornstein, Gordon, & Baker-Ward, 1999; Ornstein et al., 2006; Salmon & Pipe, 2000), and discussions of stressful events (Shrimpton et al., 1998). Ghetti et al. (2002) examined the consistency of 222 children's (aged 3- to 16-years) reports of sexual and physical abuse across two interviews (a psychological consultation and a forensic interview). The interviews were not conducted with the specific intention to repeat each other although both focussed on the same incident. Ghetti et al. found that cognitive abilities (measured using Wechsler Intelligence Scales) did not predict consistency in abuse reports but that in children aged between 3- and 16-years-old had an effect with greater consistency in older children. Memory ability predicted consistency in sexual but not physical abuse reports.

The distinction between observer and participant reflects the position of the child as witness or victim. Flin, Bull, Boon and Knox (1992b) conducted a survey in 1992 which found that over half of the children who testified in court did so as a witness rather than as a victim. Researchers have established that children between 3- and 11-years-old remembered better those events in which they had participated than those they had witnessed (Baker-Ward, Hess, & Flannagan, 1990; Gobbo et al., 2002), and that it was possible to distinguish from children's interviews whether the child participated in, or observed, the event (Akehurst, Köhnken, & Hofer, 2001), and whether the events reported were imagined or real (Bruck, Ceci, & Hembrooke, 2002; Larsson & Granhag, 2005; Ochsner, Zaragoza, & Mitchell, 1999; Sussman, 2001). Additionally,

Ochsner et al. (1999) found that children aged 6-7-years-old who witnessed an actual theft were more accurate but not necessarily less prone to suggestive questions than those who observed a neutral event. These results suggested that studies of eyewitness recall of neutral eyewitness events underestimated the accuracy of children's recall performance in criminal witness situations.

Children have been put into the category of 'witness' or 'victim' by their position as observer or participator in experimental studies. The majority of experimenters using an interviewing methodology implemented one event modality as the stimulus event. For example, in papers published between 2005 and 2006 the following events were used for children to observe: video recordings (Allwood, Granhag, & Jonsson, 2006; Beuscher & Roebers, 2005; Roebers & Schneider, 2005; Schwarz & Roebers, 2006; Shapiro, Blackford, & Chen, 2005; Shapiro & Purdy, 2005), live staged events that were primarily informative (Krähenbühl & Blades, 2006; Thomsen & Berntsen, 2005), entertaining (Gilstrap & Ceci, 2005), or mildly disturbing (Chae & Ceci, 2005), and events in which the children participated: a craft activity (Dunsmore, Halberstadt, Eaton, & Robinson, 2005), play session (Connolly & Price, 2006; Cordón et al., 2005), visit to a prepared site (La Rooy, Pipe, & Murray, 2005; Zajac & Hayne, 2006) or a variety of ordinary (doing a jigsaw, listening to a story) and less ordinary (meeting a koala bear, receiving a surprise) classroom activities (Agnew, Powell, & Snow, 2006).

Direct comparison of accuracy and consistency in interviews following a stressful event with children participating in an event ('victims') and observing an 'equivalent' staged event ('witnesses') are seldom made. In a study on the effect of stress on children's recall Shrimpton et al. (1998) compared groups of

children aged between 4 and 12 years who had undergone venipuncture ('victim') and a control group who observed ('witnessed') a demonstration of the venipuncture procedure. The results showed significant effects of age and of experience as the 'victim' children in all age groups provided fewer incorrect details in their free recall, and more correct to specific and misleading questions than the 'witness' children. Age effects were found (with the exception of the number of incorrect details given in free recall) in both 'witness' and 'victim' groups; accuracy increased with age although post hoc tests showed no significant difference between the 8-9 and 10-12-year-old children.

To demonstrate the effect of modality in respect to the influence of the level of participation, researchers have implemented more than one modality of an event into their experiments (Roberts & Blades, 1998; Roebers et al., 2004; Thierry & Spence, 2004; Tobey & Goodman, 1992). With the exception of Roberts and Blades these researchers found that participation (of children with an age range of between 3- and 7-years-old) in an event increased recall accuracy and lowered susceptibility to misleading questions in comparison to those who observed the event. Roberts and Blades (1998) found that 4-year-olds, unlike the older participants, were more confused about details concerning the interactive event than those who participated in the observation only event. This discrepancy might be attributed to the modalities of the events used. Roberts and Blades used live staged events whereas the other researchers contrasted participation with recorded material.

Associated theories

In addition to an examination of certain factors that may effect accuracy and consistency of recall of events, the theories of source monitoring, scripts and

schemas and fuzzy-trace provide explanations for distortions in responses to repeated questions.

Source monitoring refers to 'hypothetical cognitive processes' (Lindsay, 2002, p. 83) involved in distinguishing the sources of memories of events. A source is attributed rather than recalled and offers an explanation for when children (and adults) make memory errors that are not based on inaccurate memories for content (Roberts, 2000). The quality of encoding at the time of the event, subsequent experience and the quality of decision processes at the time of remembering will determine the accuracy of source monitoring. The sources being monitored are virtually always temporally separate experiences but the use of repeated questions within an interview may have an effect on which source is being recalled. Children who experience repeated questions may be unable to distinguish between the memory of the actual event or a suggestion provided in the interview (inadvertently or otherwise) by a question.

During childhood source monitoring skills develop (for a review see Roberts, 2002) and may be manifest in certain types of judgements, for example, recalling whether an action was performed or imagined, participatory or observed (Foley & Johnson, 1985; Roberts & Blades, 1999; Welch-Ross, 1995). Capability in source monitoring may depend on the delay length between event and recall (Roberts & Blades, 1995), or whether the information experienced between the event and interview was event-consistent or inconsistent (Roberts & Powell, 2006). In children aged between 4- and 10-years-old younger children have been shown to have the capability to source monitor although they tend to experience more difficulties than older children (and adults) if there are multiple source cues (Johnson, Hashtroudi, & Lindsay, 1993), if the stimulus was

auditory rather than visual (Markham, Howie, & Hlavacek, 1999), or depending on the interviewing strategy used (Roberts & Blades, 1995, 1998, 1999). Children who cannot source monitor effectively may not be able to identify which pieces of information came from their own experience of the original event and which came from other sources, which has implications for their accounts of information (Ackil & Zaragoza, 1995; Allwood et al., 2006; Ceci, Huffman, & Smith, 1994; Poole & Lindsay, 1995).

Within an interview source monitoring takes place when a child attempts to distinguish between a response to a question based on the actual event and the response to a previous question in the interviewing process. If the child recognises a question as a repetition then the child may not refer to the actual original event, but to their response to a previous version of the question. Even if the earlier answer had been inaccurate it is possible, under source monitoring, that this response, rather than the actual original event, may now be perceived as the 'accurate' memory of the event because of the source monitoring error.

Another explanation for distortions in recall is provided by schema theory where recall is based on a representation formed through prior experience rather than specifically related to the details of the actual event. A schema is a structured cluster of mental representations of generic knowledge acquired from past experience and may be used to represent events, sequences of events, perceptions, relationships and so forth (Bartlett, 1932). Within these schemas individuals build up an event representation of similar events to form a script, in effect an 'encoding' of a stereotypical or generalised sequence of actions which enables recall about such events (Roberts, 2002).

Children may form a script that represents the way in which conversations with unfamiliar adults proceed. These scripts may include evasive responses or complying with false statements in order to 'be polite' and not contradict, both of which could create some discrepancies between interviewer expectations when confronting a child in an interview situation. Similarly, adults will also have developed a script concerning the ways in which to talk to children, the implications of children's responses and behaviour. In respect to repeated questions there is recognised concern that children have an existing script for conversational rules that includes the assumption that repetition of the question implies inadequacy of the initial response and therefore encourages a subsequent change in response (Home Office, 2001; Moston, 1987; Siegal, 1991; Warren et al., 1991).

Schema acquisition develops with age (Farrar & Goodman, 1990; Fivush, Kuebli, & Clubb, 1992) and may be used to enable memory recall. This would suggest that older children would be more consistent in their responses to repetitions as they would have a more established script on which to base their recall. However, the formation of pre-event stereotypes and the context in which recall takes place may distort accurate recall (Marsh, 2007; Memon, Holliday, & Hill, 2006).

Fuzzy-trace theory is used to describe how recall of an event may change with time. Fuzzy-trace theory differentiates between independently operating verbatim and gist representations which become encoded in memory (Brainerd & Reyna, 1995, 1998; Reyna & Brainerd, 1995; Reyna & Kiernan, 1994). Verbatim memory is a representation of the surface content of a memory, the

experience as it occurred rather than its semantic content. The associated senses, patterns or meanings of an experience become represented in gist form.

Fuzzy-trace theorists (Brainerd & Reyna, 1995; Cowan, 1998; Davies, 1995; Miller & Bjorklund, 1998; Reyna & Brainerd, 1995, 1998; Reyna & Kiernan, 1994) suggest that, after a delay, a gist representation would be encoded and retained. Therefore responses to questions posed after a delay may not be inaccurate in terms of that gist memory even if they are inaccurate about the specific details required.

There are developmental differences in the ability to recognise gist and inhibit verbatim details (Marx & Henderson, 1996; Reyna & Brainerd, 1995, 1998). According to fuzzy-trace theory, in children under 10 years of age there is an initial bias in the younger children towards verbatim encoding which develops into a bias towards gist encoding in the older children. Each method of encoding may be advantageous depending on the type of information to be recalled (Miller & Bjorklund, 1998) and the delay between encoding and recall (Reyna & Kiernan, 1994).

Reyna and Kiernan (1994) found that initial verbatim memory representations were highly accurate, but this accuracy diminished quickly with delay. Gist representation, however, remained stable after a delay. In questions that accessed gist representations, accuracy of response was not clearly definable because a question requiring specific details may access the accurate gist representation of that memory but not provide the details required. For example, the question "What was the colour of his jacket?" may have resulted in the access of a gist representation of dark, rough textured and dirty clothing resulting in the response that the clothing was black (i.e. "dark") whereas in actuality it

was another dark colour. In police or other interviews such a response would not be assessed as totally accurate, but in view of the gist trace accessed it was correct. This is not to imply that gist representation equates with imprecise processing, for as Acredolo stated it "organizes, explains and predicts specifics" (1995, p. 85). "Linear" sentences or questions that require verbatim memory (for example "Did he say that you needed to go behind the bushes before or after he shouted at you?") relied on verbatim representations, which decayed more quickly than gist representations over a one week delay (Reyna & Kiernan, 1994). If the verbatim representation was no longer accessible then the gist representation would be accessed, which may lead to problems in accuracy in providing the actual words used but would nevertheless provide relatively accurate information.

Interviewing practices

Investigative interviewing protocols reflect the aim to elicit complete and accurate responses and often incorporate similar methodological procedures. In comparison to informal conversations the implementation of interviewing protocols can have the potential to increase accuracy, completeness and consistency of children's testimonies. The processes and strategies included in interviewing protocols have been informed by psychological research into questioning strategies. Researchers have shown that free recall and specific type questions elicited the most accurate information (Holliday, 2003a; Lamb et al., 1996; Perlman et al., 1994; Seidler & Howie, 1999), the most complete or the greatest quantity of detail (Hershkowitz, 2001; Orbach et al., 2000), and the most consistent reports both within single interviews (Lamb & Fauchier, 2001; Orbach & Lamb, 2001) and between interviews (Jones & Pipe, 2002) in comparison to

other question forms. Specific guidance to promote the use of such questions has been incorporated both into interviewing strategies such as the cognitive interview, structured investigative interview, and narrative elaboration technique. This guidance has also been utilised directly in the production of interviewing protocols of which ABE (Home Office, 2001) is currently in use in England and Wales.

The Cognitive interview

The cognitive interview (CI) was developed by Fisher and Geiselman (Fisher, Geiselman, & Amador, 1989; Geiselman, Fisher, Mackinnon, & Holland, 1985) to provide a method by which to increase the quality and quantity of information elicited from victims, witnesses and suspects. CI has also been used by social workers (Aldridge, 1999) and in medical staff investigations (McIndoe & Walsh, 2000). CI was later adapted for use with children and to include an interviewing structure (Fisher & Geiselman, 1992; Geiselman, Saywitz, & Bornstein, 1990). An assumption of CI is that a memory is composed of several features, which may be recalled by various retrieval paths that overlap at least part of one or more of the memory features (Geiselman, Saywitz, & Bornstein, 1993). In a CI the interviewee is encouraged, through a structured progression to: (i) report everything, (ii) mentally reconstruct the context of the event, (iii) recall in different temporal orders, (iv) report events from alternative perspectives, and recall specific details through the use of 'memory jog' techniques (Milne & Bull, 1999).

These CI techniques may use question repetition in either verbatim form, for example by asking "and what happened next" on more than one occasion, or gist repetition, for example asking "What was he wearing?" and later "What

clothes did he have on?" to facilitate the examination of the same material from all angles. The CI has been found to be highly successful (Kebbell, Milne, & Wagstaff, 1999) not only in eliciting more details but also in enabling children aged between 4- and 11-years-old to resist suggestive questions (Milne & Bull, 2003), and in reducing the incorporation of misinformation (Holliday, 2003b; Holliday & Albon, 2004) even after delay (Akehurst, Milne, & Köhnken, 2003; Larsson, Granhag, & Spjut, 2003).

However, Memon (1999) expressed a note of caution as to whether CI was an appropriate tool for interviewing children of all ages. Memon suggested that children under 7 years of age might not understand what is being asked of them or may be adversely affected by demand characteristics. Additionally, if children were asked to describe an event in response to question repetition, they may feel under pressure to change their responses, which was generally detrimental to accuracy.

Structured investigative interviews

Researchers at the US National Institute of Child Health and Human Development (NICHD) developed a semi-structured interview protocol to guide forensic interviewers on conducting interviews and to "translate research-based recommendations into operational guidelines in order to enhance the retrieval of informative, complete, and accurate accounts of alleged incidents of abuse..." (Orbach et al., 2000 p. 738). The protocol described how interviewers should: introduce themselves, outline the purpose of the interview, and establish the 'ground rules' such as the need to indicate when an answer to a question was unknown, or interrupt if the interviewer stated something that was incorrect. The protocol does not explicitly promote or caution against the use of repeated

questions, but the examples of questioning strategies used incorporated question repetition (in gist form) without any rationale for its use presented to the child being interviewed.

Research into the effectiveness of the NICHD interviewing techniques across age groups has demonstrated positive results with its use associated with an increase in the amount of details elicited (Hershkowitz, 2001; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2001; Orbach & Lamb, 2000) and a reduction of inconsistencies (Lamb & Fauchier, 2001). The use of the structured protocol resulted in over four fifths of 4-12-year-olds making an allegation in comparison to a third of children who experienced a standard interview (Lamb et al., 2001). However, making an allegation or providing more details does not necessarily equate to accuracy. Also, it was not established through this research whether it was the use of the questions advocated by the protocol or other elements such as the use of repetition, the rapport building, the pre-questioning stage practice and so forth which uniquely enhanced or detracted from the contents of the children's responses.

Other interviewing techniques such as Narrative Elaboration Technique (NET) and Statement Validity Assessment (SVA) provide strategies for conducting interviews and assessing the contents but do not directly mention the use or effect or repetition. The purpose of NET is to enable a child, through training of retrieval strategies, to narrate past events without being influenced by others and thereby possibly compromising accuracy (Saywitz, 1995; Saywitz & Snyder, 1993). NET has been shown to be very successful at improving the number of details and quantity of spontaneous recall of children aged between 4- and 11-years-old without compromising accuracy or increasing errors (Bowen &

Howie, 2002; Brown & Pipe, 2003b; Saywitz & Snyder, 1996) and even attenuated recall of those with lower IQ (Brown & Pipe, 2003a). However, it is unclear whether these positive benefits would be of practical use, as the questioning methods of these studies did not reflect actual interviewing practice in forensic settings.

A characteristic of SVA (for details see Steller, 1989; Steller & Boychuk, 1992) is the system of content-based criteria analysis (CBCA), that rates the child's statement for the occurrence of content criteria (for example: reproduction of speech or description of interactions) the presence of which is equated with higher quality statements and adds confidence that the statement was based on genuine personal experience rather than partial or total fabrication (Vrij, 2002). The presence (or absence) of criteria is combined with a validity rating (based on assessment of realism, contradictory alternative evidence or suggestion of coaching) to determine the overall credibility of the statement. The child is interviewed using a semi-structured phased interview (known as "Step-wise") as described by the NICHD (Orbach et al., 2000) and the statement elicited assessed.

Repetition of questions and the possible effect of this practice are not directly mentioned by CBCA. Under the validity checklist the characteristics of the interview, whether the interview included suggestive, leading or coercive questioning, is considered. It is possible that repeated questions could be considered coercive, but the emphasis of CBCA evaluation is on the child's statement rather than how it was elicited.

Innovations and Interviewer training

Procedural reforms have been introduced in order to "(a) ...decrease potential harm to children because of their involvement in criminal cases and (b) to facilitate prosecution of adults who commit crimes, especially sexual abuse, against children." (Goodman, Quas, Bulkley, & Shapiro, 1999, p. 256). In England and Wales these reforms have had a twofold effect: the provision of special measures designed to create a system and environment that was responsive to children, and the provision of the ABE (Home Office, 2001) interviewing protocol for use by trained professionals in investigative interviews. However, apart from the ABE advice to exert caution with the use of repeated questions, a recognition or discussion of the problems associated with repetition does not appear to form part of interviewer training (see Appendix 2) or be of immediate concern to qualified police interviewers (see Appendix 3).

Instruction issues

While the main responsibility for ensuring that an interview elicits complete and accurate testimony rests with the interviewer, the veracity of the report also depends on the child understanding what is required of him or her and acting accordingly. In NICHD, MOGP and ABE interviewing protocols (Home Office, 1992, 2001; Orbach et al., 2000) there are instructions given to the children in the form of 'ground rules'. These include explaining that the interviewer cannot possibly know the correct details because the interviewer was not there at the time of the incident, explaining the appropriate use of the "I don't know" response, and encouraging the child to indicate when he or she did not understand a question or if the interviewer made a mistake. However, the administration of these instructions was found to be inconsistent (Warren,

Woodall, Hunt, & Perry, 1996). Also the efficacy of the impact of instructions on children's behaviour in responses is doubtful as researchers have established that children aged between 5- and 8-years-old even responded to nonsensical or bizarre questions particularly when the question was in a closed format (Waterman et al., 2000). Although children would almost invariably reply to nonsensical or bizarre questions older children (aged 7-years) tried to make sense of the questions and attempted to qualify their responses (Hughes & Grieve, 1980). Similarly, children (aged 5-7-years) gave responses to bizarre questions despite being able to accurately state whether or not the questions made any sense (Pratt, 1990).

Interviewer knowledge

The social context of an interview has been shown to affect the accuracy of information elicited in interviews with children aged between 3- and 9-years-old following staged events in experiments (Bjorklund et al., 2000; Goodman et al., 1995; Ricci, Beal, & Dekle, 1996; Tobey & Goodman, 1992; Waterman et al., 2004; Welch-Ross, 1999) and may also contribute to the disclosure or non-disclosure or amount of information elicited in interviews with children aged 4- to 14-years-old in interviews relating to suspected sexual or physical abuse (Hershkowitz et al., 2006; Lamb & Garretson, 2003). In all of these studies it is the relationship between the child and the interviewer and/or the questioning regime that results that affects the information elicited.

Inaccurate information suggested in conversations prior to interviews can compromise the accuracy of eyewitness memory reports (Ceci & Bruck, 1995). For example, Tobey and Goodman (1992) interviewed 4-year-old children about a staged event involving play activities with a 'babysitter'. In one condition the

children had interacted with a police officer who mentioned that the babysitter had previously behaved inappropriately, in the second condition the children did not interact with the police officer. In subsequent interviews the children in the 'police' condition made more errors in their free recall and were more likely to make incorrect responses to misleading questions than were the children in the 'not-police' condition.

However, when the interviewer was the child's own mother the children's free recall was not affected by differences in interviewer condition (Goodman et al., 1995). In Goodman et al.'s study the interviewers were either the children's mothers or unknown by the children. Before conducting the interviews half of the interviewers were provided with biased prior information relating to the children's play activity. The free recall elicited from the children in interviews with their own mother's taking the interviewing role elicited similar levels of accuracy in free recall regardless of whether the mothers had been given misinformation or not. When 'stranger' interviewers interviewed the children the introduction of misinformation to these interviewers had an adverse effect on the children's free recall accuracy.

In forensic interviews the interviewers commonly have prior knowledge of some of the details relating to the event being discussed (although were not present when the event occurred) and have previously met the children to be interviewed in order to establish rapport (Home Office, 1992, 2001). On such occasions the child assumes (rightly or wrongly) that the interviewer has prior knowledge of the event even if the interviewer was not present at the event. Research into the effect of this assumption on children's accuracy or quantity of information provided has been limited to experiments that included (in one

condition) the presence of the interviewer at the stimulus event. Welch-Ross (1999) examined the effect of awareness of interviewer knowledge on children's (aged 3-6-years-old) likelihood of being misled in a subsequent interview. Welch-Ross found that children were more likely to be misled by a knowledgeable interviewer than a naïve interviewer.

Even when misinformation is not introduced an effect of interviewer knowledge is found on the accuracy of responses to unanswerable questions following a staged event (Waterman et al., 2004). Waterman et al. examined the accuracy of children's (aged 5-9-years-old) responses to questions of different types posed by either knowledgeable or naïve interviewers. The children performed consistently well on responses to answerable questions regardless of interviewer knowledge. However, in responses to unanswerable questions, which required an "I don't know" type response, were more likely to respond correctly when the question was posed by a naïve interviewer.

To date there has been no research into the effect of different levels of interviewer knowledge; between actual involvement or presence at the event being discussed as in the studies by Waterman et al. (2004) and Welch-Ross (1999), the assumption that the interviewer has knowledge of the event as in the studies by (Goodman et al., 1995; Tobey & Goodman, 1992) and when an interviewer is naïve. In most experimental studies the interviewer is assumed to be naïve to the event as he or she was not present during the event. On the basis of research mentioned above such an interviewer can be expected to elicit less accurate responses to free recall if misinformation is introduced but more accurate responses to unanswerable questions without the inclusion of misinformation than would a knowledgeable interviewer.

In this chapter we have outlined the effects of question repetition on accuracy and consistency in both experimental and actual interviews with children. There is a consensus that repetition is likely to detract from accuracy and that it encourages changes in children's responses. However, what remains unclear is the extent to which question repetition is used in investigative interviews, what form that repetition may take and what effect repetition has on children's responses within a single interview. Our studies will explore the use of question repetition in investigative interviews, and examine the effect of question repetition on children's accuracy and consistency.

A discussion of interviewing protocols and the implementation of those protocols used in forensic interviews will follow in Chapter 2. Chapter 2 will examine in detail the quantity, frequency, and context of the use of question repetition in actual interviews. In Chapters 3, 4, 5 and 6 a series of experiments are reported that were conducted in order to ascertain which factors influence the impact of question repetition on children's accuracy and consistency in responses. The studies in Chapters 3, 4 and 5 use an observed event on which the interview is based. Chapter 3 examines the effect of different forms of question repetition. Chapter 4 assesses the effect of multiple question repetitions. Chapter 5 determines whether the timing of an interview and question repetition patterns affect children's responses. Chapter 6 uses an event in which the children are actively involved to establish whether participation affects responses to repeated questions.

Chapter 2

Study 1

In this chapter we will examine interviewing processes as they are actually performed by police in interviews with young children. We will make specific reference to interview protocols, structures, features and to the use and effect of question repetition.

Introduction

The UN convention of children's rights established unequivocally that children should be given the opportunity to have express their views in judicial proceedings according to the laws of that nation (Article 12, United Nations, 1989). Under legislation for England and Wales, competency to give evidence is not dependent on age ("Youth Justice and Criminal Evidence Act", 1999, section 53). Accordingly, children are taking part in judicial proceedings; according to a NSPCC press release (February 15th, 2007) nearly 30,000 children in the UK give evidence in court each year.

Ainsworth (1998) has suggested that the basic premise of the Criminal Justice system that witnesses tell 'the truth, the whole truth and nothing but the truth' may be "naive, simplistic and unrealistic" (p. 33). However, as children do partake in legal proceedings as witnesses it is the duty of those involved in legal proceedings to provide a system that enables them to do so. Of particular importance is to establish procedures that allow children's witness evidence to be elicited and represented fairly. To improve the quality of children's recall researchers have examined strategies and methodologies used in questioning, the results of which have influenced the development of, or have become

incorporated into investigative interviewing protocols (Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Goodman et al., 1994; Goodman et al., 1999; Goodman et al., 1992; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Lamb et al., 1996; Moston, 1987, 1990; Ornstein, Baker-Ward, Gordon, & Merritt, 1997; Poole & Lindsay, 1995; Ricci et al., 1996; Sternberg et al., 1996; Sternberg et al., 1997; Warren et al., 1991).

On the basis of research and governmental directives (Davies, Marshall, & Robertson, 1998; Home Office, 1989, 1998) investigative interviewing protocols have been developed and revised to facilitate the eliciting of appropriate eyewitness testimony from children involved in legal systems. Examples of interviewing protocols include 'Finding Words', 'Forensic Interviewing Protocol', NICHD, and ABE, which was revised from the earlier MOGP (American Prosecutors Research Institute, 2003; Home Office, 1992, 2001; Orbach et al., 2000; State of Michigan Governor's Task Force on Children's Justice & Department of Human Services, 1998).

In England and Wales the Crown Prosecution Service's (CPS) decision to prosecute on not depends on many factors of which the quality of the child's testimony plays a major part particularly in conjunction with the quantity and quality of further evidence required. For as Davis, Hoyano, Keenan, Maitland and Morgan stated (in reference to England and Wales),

> "... the strength of the additional evidence required was dependent on the clarity and consistency of the child's account. Where a child's testimony was considered to be exceptionally clear and detailed, evidence of opportunity might be considered sufficient. When the child's account was vague or inconsistent, a case would only be prosecuted where there was other strong evidence supporting the child's account, such as clear medical signs or testimony from other children who were making similar allegations." (1999, p. 46).

Attrition rates for child abuse and neglect cases are high, which suggests that the quality of children's testimony, as described above by Davis et al. (1999) appears to be insufficient. Gallagher and Pease (2000) found that in England and Wales less than one fifth of cases (in 1997) proceeded to prosecution although over four fifths of these resulted in conviction. Factors concerning the witness's statement, credibility, ability to give evidence, reaction to the court case itself or reluctance to proceed accounted for just under half of the reasons the police gave for taking no further action (Gallagher & Pease, 2000). All of these factors related to the recorded statements or the process of eliciting a statement, which demonstrates the significance of the interviewing process to the continuation of a case.

As described in Chapter 1 (pp. 12-13) researchers in forensic testimony have established that children are capable of providing accurate testimony about events that they have witnessed or experienced (Lamb, Sternberg, Orbach, Hershkowitz, & Horowitz, 2003b). There are, however, differences in the quality of testimony given. Generally, younger children provided fewer details (Lamb et al., 2003b), were less consistent both between interviews (Ghetti et al., 2002), and within single interviews (Lamb & Fauchier, 2001) often as a consequence of suggestive questioning (Lamb & Fauchier, 2001; Orbach & Lamb, 2001).

The NICHD, MOGP and ABE interviewing protocols (Home Office, 1992, 2001; Orbach et al., 2000) are all semi-structured, and emphasise the positive use of free recall and open-ended prompts, and the use of specific rather than closed questions. The beneficial effect of these prompts and questioning styles have been demonstrated in studies of forensic transcripts (Hershkowitz, 2001; Korkman et al., 2006; Lamb & Fauchier, 2001; Lamb et al., 1996; Orbach

et al., 2000; Orbach & Lamb, 2000) and in experimental studies (Dent & Stephenson, 1979; Fivush, 1993; Steward et al., 1993). Such questions not only elicited the most accurate details (Akehurst et al., 2003; Lamb et al., 1996; Memon & Vartoukian, 1996; Orbach et al., 2000; Peterson et al., 1999; Quas & Schaaf, 2002; Sternberg et al., 1996) across lengthy delays in time (Jones & Pipe, 2002; Peterson, 2002), but also resulted in lower susceptibility to suggestibility and misinformation (Gee et al., 1999; Gobbo, 2000; Holliday, 2003a).

The MOGP (Home Office, 1992) was developed to meet the requirements and recommendations of the Criminal Justice Acts (1988, 1991), the Home Office Advisory Group on Video Evidence (Home Office, 1989), and Government inter-agency guide ("Working together", 1988; revised edition 1991) which followed the Butler-Sloss Report on child abuse in Cleveland (1987). Adherence to MOGP protocol was not obligatory but was to be followed, if at all possible, to ensure that interviews with children under 17 years of age alleging sexual offences (and children under 14 years of age alleging violent offences) would be performed appropriately. The aim of the interview was to elicit an accurate and complete account in a manner that was fair, acceptable to the courts and consonant with the best interests of the child (Bull, 1992; Home Office, 1992). The MOGP provided guidance and advice on (i) satisfying the legal conditions controlling when and where to make a video recording, (ii) pre-interview planning, (iii) conducting the interview, and (iv) arrangements for the video recording after completion.

Following the Youth Justice and Criminal Evidence Act (1999) the MOGP (Home Office, 1992) protocol was revised and replaced by ABE (Home

Office, 2001) in order to incorporate improved protection for vulnerable or intimidated witnesses (including adults) as discussed in the Speaking Up for Justice report (Home Office, 1998). The interviewing structures of the ABE and MOGP interviews mirror certain aspects of the NICHD interview protocol with four interviewing stages described below: Rapport, Free Narrative Account, Questioning, and Closing the Interview.

The Rapport stage is used, regardless of previous contact with the interviewer, to: establish a relationship between the child and the interviewer, explain the aims and conventions (through "ground rules") of the interview, and enhance the interviewer's knowledge of the child's social, cognitive and emotional development. The ground rules include explaining that it is acceptable for the child to say, "I don't know" or "I don't understand", if appropriate.

During the Free Narrative Account stage the interviewer takes on the role of facilitator rather than interrogator and asks only open-ended questions or 'invitations' to speak. During this stage the information obtained from the child should not be subject to interviewer direction or influence (Home Office, 1992).

Questions in the Questioning stage are initially open-ended but may move to specific questions (those that begin with 'wh-') and then to closed (where the question posed has limited alternatives for responses) and, if required, to leading questions (although these are generally to be avoided). Questions should be simply constructed, use age appropriate vocabulary and be used to examine inconsistencies in the child's account. Repetition of questions should be used with caution, as

> "Repeating a question soon after a child has answered ... may be interpreted by children as a criticism of their original response... persistent repetition of a question may lead a child to give an answer he or she believes the interviewer wants to hear." (Home Office, 1992, p. 18).

The function of the final closure stage is to put the child in a positive state of mind at the end of the interview. The interviewer may also wish to check some details to ensure that they have understood correctly and then return to neutral topics of conversation. The child is given the opportunity to ask questions, and is provided with contact details before the interview is terminated.

Evidence has indicated that interviewers do not always follow the methods and procedures prescribed in the interviewing protocols (Powell & Snow, 2007; Westcott & Kynan, 2006). A common finding is that there is an inappropriate reliance on closed and suggestive questions or prompts (Sternberg et al., 2001; Warren et al., 1996; Westcott & Kynan, 2006; Westcott, Kynan, & Few, 2006). Closed or suggestive questions such as "He told you to do that, didn't he?", "Then you went upstairs?", "Did he take off your clothes?" provided up to half of the information gained in some interviews (Cederborg et al., 2000; Sternberg et al., 2001). Although improved police interviewing practices have led to a reduction in the use of suggestive and closed questions, the frequency of open-ended questions has not increased (Thoresen et al., 2006). In addition, some police interviewers provided children with few opportunities for free recall or narrative (Davies et al., 2000; Sternberg et al., 2001; Warren et al., 1996) thus reducing the opportunity presented in the interviewing protocol for children to provide uninfluenced or undirected testimony.

During early implementation of the MOGP (Home Office, 1992) Davies, Wilson, Mitchell and Milsom (1995) found that the structure and question guidance of the interviewing protocol was not being consistently followed. In a sample of 40 video taped interviews, a quarter lacked a free narrative phase although the rapport, questioning and closure stages were clearly performed.

Only a third of interviewers began the questioning stage with open-ended questions. Nevertheless, three quarters of the interviews were judged to provide clear accounts of the alleged incidents and leading questions were rare. These question and stage related problems could have been due to unfamiliarity or a lack of experience in the MOGP interviewing principles but later evaluations provided similar findings (Davies & Westcott, 1999; Davies et al., 2000; Sternberg et al., 2001; Westcott & Kynan, 2006; Westcott & Jones, 1997). Specifically, open-ended questions were seldom asked (Davies et al., 2000), and option-posing or suggestive prompts were used to elicit two fifths of all the information provided by the child in the interview (Sternberg et al., 2001). The free narrative stage was found to be absent or brief, particularly in interviews with children under 7 years of age (Westcott & Kynan, 2006).

Not all questions asked by interviewers will be answerable; some questions posed by interviewers may be unanswerable because children lack the relevant knowledge to respond. The MOGP and ABE (Home Office, 1992, 2001) interviewing protocols instruct interviewers to inform the child that they should tell the interviewer if they do not know the answer to a question. However children are reluctant to admit ignorance, would rather speculate than say that they do not know an answer (Peterson et al., 1999; Waterman et al., 2000, 2004), and have been shown to attempt to answer questions they do not understand or which make no sense (Hughes & Grieve, 1980; Pratt, 1990; Waterman et al., 2000).

Problems with interviewing practices are not limited to England and Wales; similar concerns regarding questioning strategies used by interviewers are expressed by researchers in the US (Warren et al., 1996), Sweden

(Cederborg, 2004; Cederborg et al., 2000), Israel (Lamb et al., 1996) and Norway (Myklebust & Alison, 2000).

In addition to procedural reforms interviewer training has been introduced to protect children from potential harm connected to their involvement in criminal cases and to enable the prosecution of those responsible for committing crimes (Goodman et al., 1999). However, despite training guidance in the MOGP and ABE interviewing protocols (Home Office, 1992, 2001), it has been found that children are still asked inappropriate questions (Lamb et al., 1996; Powell, Fisher, & Wright, 2005), and encounter confusing language, terminology or question structures (Perry et al., 1995; Plotnikoff & Woolfson, 2004; Saywitz, Jaenicke, & Camparo, 1990). This continuing issue formed the stimulus for research on specific aspects of interview and training development. For example, the inability of professionals to maintain the use of open-ended questions was addressed by Powell and Snow (2007) through provision of practical suggestions for formulating open-ended questions in the free narrative stage. The use of a review programme for trained interviewers was found to have positive outcomes with improvements in quality (Westcott et al., 2006).

The MOGP and ABE protocols (1992, 2001) provide guidance on the recommended form of interview procedure with concrete examples of certain techniques, for example, on how to establish whether a child understood the concept of truth and lies. The guidance is not so clear in respect to question repetition; use of the same form of question is cautioned against but without further suggestions of alternative strategies to elicit the required information.

Gilstrap (2004) analysed the questions used by police interviewers in unstructured interviews with 3-7-year-olds who had witnessed a staged event. Gilstrap suggested that the interviewer's behaviour, through questioning, indicated or suggested to the child what responses or areas of topic were relevant (see also Garven et al., 1998). Gilstrap found that virtually all of the interviewers used question repetition and that over a tenth of all questions were repeated, which resulted in a fifth of the inaccurate responses.

The child's perception of what the interviewer's behaviour suggested is an example of the implementation of a schema or script (see Chapter 1, pp. 36-37) formed through prior experience (Bartlett, 1932; Roberts, 2002). The child may have formed a script that represented conversational rules, adult and child interaction, acquiescence to the demands of those in authority and so forth. Although the interviewer may be endeavouring to elicit greater detail by use of question repetition, the child's previous experience of question repetition may have generated a script in which repetition is used when an initial response is not acceptable.

The use of repetition is inevitable if the response to the interviewer's initial question is inadequate or inappropriate (repetitions 'with motive') or if the question is, for example, misheard (questions that are repeated for practical reasons, and are in essence 'neutral'). Warren, Hulse-Trotter and Tubbs (1991) found that children who were given a direct suggestion that an original response was incorrect changed their responses more than children who were given no explanation for the question repetition. However, changes in responses may also occur when the rationale for repetition did not imply inaccuracy (Howie et al., 2004) or when no rationale was given (Krähenbühl & Blades, 2006).

The use of question repetition has been shown to lead to changes in children's responses, suggesting inconsistency, especially among younger children (Krähenbühl & Blades, 2006; Moston, 1987; Poole & White, 1991). The level of consistency in responses (as a result of question repetition, or otherwise) was rated as one of the top three criteria factor used by professionals in their assessment of the accuracy of a child's testimony (Steward et al., 1996). Adults (in the role of legal representatives, jurors and so forth) may already have stereotypes concerning child witnesses and criminal behaviour (Gilstrap, Fritz, Torres, & Melinder, 2005). The preconceptions of these adults and their assessment of the quality of the child's testimony in addition to the general conception that consistency equates with accuracy (this has been shown to be erroneous by Quas et al., 2007) may have an adverse effect on perceptions of child credibility.

Therefore, in Study 1 we examined the structure of the interviews conducted by the police officers. We documented the quantity, type, position and pattern of question repetition used by the police interviewers in their investigative interviews and examined the effect that those repetitions had on children's responses.

The transcripts in Study 1 were recorded as having been conducted according to the MOGP interviewing protocol (Home Office, 1992). The questioning structure advocated in MOGP can also be found in similar forms in other currently used interviewing protocols such as Finding Words (American Prosecutors Research Institute, 2003), the Forensic Interviewing Protocol (State of Michigan Governor's Task Force on Children's Justice & Department of Human Services, 1998), and ABE in England and Wales (Home Office, 2001).

We examined the police interview transcripts to identify instances of repeated questions and made the following hypotheses.

First, that as the interviews were recorded as having been conducted using the MOGP protocol (Home Office, 1992) the following features would be present: four distinct interviewing stages (rapport, free narrative, questioning and closure), the use of closed questions only after open-ended questions had been exhausted (resulting in *open to closed* rather than *closed to open* repetitions), and the avoidance of leading questions. We expected question repetition would be infrequent because the MOGP guidelines state that repetition should be used with caution. We also expected that if repeated questions were used in the interviews, there would be long intervals between repetitions as MOGP recommended that short intervals between repetitions should be avoided.

Second, we categorised the repeated questions either as 'with motive' (where an alternative response of any form was considered necessary) or as 'neutral' (where repetition was used because of practical difficulties such as sound levels). We expected that 'with motive' repetitions would be more common than 'neutral' ones as it was expected that not all responses to initial questions would fulfil the interviewer's requirements.

Third, we expected that questions repeated from *closed to open* forms would lead to additional information being elicited, without contradiction of an earlier response (Lamb & Fauchier, 2001; Orbach et al., 2000; Orbach & Lamb, 2000, 2001). In comparison to other question forms closed questions lead to a higher number of changed responses (e.g. Quas & Schaaf, 2002; Waterman et al., 2001), therefore we expected *open to closed* repetitions would result in the most frequent changes in responses.

Method

Sample

The sample comprised 100 police transcripts of video recorded interviews conducted between 1994 and 1997 from 13 collaborating police forces in England and Wales. These transcripts had been selected on an 'opportunity sampling' basis and were recorded as having been conducted using the MOGP interviewing protocol (Home Office, 1992). Five transcripts were omitted because they were not first interviews or because pages were missing or illegible. The following information pertains to the remaining 95 transcripts.

Accompanying each transcript was a record sheet with information about the child being interviewed, the interviewer, the interview duration, the alleged offence and details of the alleged offender. These record sheets were not always fully completed.

To examine differences in interviewing practices according to the age of the child being interviewed the transcripts were divided into four age groups: 4-5-year-olds, n = 21 (22.2%), 6-7-year-olds, n =19 (20.0%), 8-9-year-olds, n = 20 (21.0%) and 10-11-year-olds, n = 35 (36.8%). There were 28 males (29.5%) and 67 (70.5%) females. Children's ages were recorded in years only therefore it was not possible to calculate mean ages or standard deviations.

The interviews ranged in length from 11 minutes to 136 minutes, M = 40.15 minutes, SD = 19.09 minutes, missing data n = 10 (10.5%). The mean interview length for both 4-5-year-olds and 6-7-year-olds was 36 minutes, for 8-9-year-olds M = 47 minutes, and for 10-11-year-olds M = 42 minutes.

The majority of interviewers were police officers, n = 82 (86.3%). There were also social workers, n = 5 (5.3%), and a psychologist, n = 1 (1.1%), missing

data, n = 7 (7.4%). Interviewer gender was female, n = 74 (77.9%), male n = 16 (16.8%), and missing data n = 5 (5.3%). All interviewers had been trained to conduct MOGP (Home Office, 1992) interviews although details of this training, where and when it took place, were not given. Eighty-four (88.4%) of the transcripts were recorded by interviewers as having been conducted in accordance with the MOGP protocol, n = 2 (2.1%) were recorded as not having been conducted in accordance (although no details were provided to describe the alternative procedure), n = 9 (9.5%) missing data.

The abusive experiences reported ranged from exposure, to rape and abduction. There was no indication as to whether the alleged incident(s) were single or multiple occurrences. In the recorded opinion of the interviewers n = 84 (88.4%) of the children had experienced the abuse they alleged, n = 5 (5.3%) had not experienced abuse, n = 1 (1.1%) unsure, and n = 5 (5.3%) missing data. Of the n = 85 where abuse was thought to have occurred n = 66 (77.6%) resulted in court cases, n = 15 (17.6%) did not proceed, n = 4 (4.7%) missing data. Information relating to the outcome of the court cases was not available.

Procedure

All of the questions asked in the transcripts were collated and coded as either open or closed. Closed questions were defined as those with a limited range of responses available, for instance those that required either 'yes' or 'no' responses or a response selected from the options posed in the question. Examples of closed questions included, "Did he wash you every day?", "So it hurt you?" and "Was it one time or lots of times?" Open questions were defined as those that required a child to produce a response not specified by the format of the question asked. For example the open questions "Oh right, so how many

times has he done that then?", "Why does he pay you money?", and "What did he say to you?" were used in the transcripts.

Questions involved in repetition were recorded and coded for position, quantity, style, purpose, and consequence of repetition (outlined below). The coding systems were developed specifically for this study. Our definition of repetition was: when an initial utterance concerning the same subject of response was stated again either in the same or different form. A repetition may only be recognisable in context. An example of repetition taken from the transcripts was: "So how did it all end - end – you know, finish up?" repeated as "How did it all stop?", and finally as "What happened at the end?" Another example was "Where did you get the money, A, when you're leaving or before?" repeated initially as "Before you're coming out?" and subsequently as "Before things or after things?", "Was it when you were coming home he gave you the money or did he give you the money when you first went in?", and finally as "Mm?"

Repetition position

The number of intervening phrases between the initial 'base' question and the following repetition was counted. A phrase was defined as a whole or part sentence (when a sentence was left incomplete), or when there was a stated action that provided an alternative focus, for example, if the child got up and went to look out of the window.

The number of intervening phrases was put into the following categories: 0, 1, 2-5, 6-20, and 20+. The first three categories were the equivalent of the 'immediate' repetition which has been used in experimental research (Howie et al., 2004; Poole & White, 1991), the latter two categories represented medium

and long intervals. When a child remained silent the number of intervening phrases was coded as 0.

Repetition quantity

In each interview the total number of repetitions and the number of repetition in each set (a set was defined as the initial 'base' question and all subsequent related repetitions) were also recorded.

Repetition style

The repetition styles were categorised as 'neutral' (for example when the interviewer needed to repeat because the child had misheard the original) or 'with motive' (when the interviewer appeared to intend to affect a child's response).

In the neutral category were the following styles:

1. Misheard - when the original question was misheard, for this coding to be applied it was necessary for the child to say, "pardon", "what did you say" or similar.
2. Linguistic error - for example asking, "What's your birthday?" repeated as "When's your birthday?"
3. Rephrasing – altering the question to more age appropriate language, this required a comment to this effect by the interviewer.
4. Sound - when a problem with recording or sound levels necessitated a repetition.

The 'with motive' repetition styles (which it must be emphasised did not necessarily have negative associations) were as follows:

1. Verbatim - repeated word for word.

2. Open to closed - from an open-ended question format to a closed question format. For example, "Who else was in the van?" repeated as "So … were you the only one in the van?"
3. Closed to open - from a closed question format to an open-ended question format. For example, "Do you have a name for the place that he touched you?" repeated as "What do you call that?"
4. Gist - repetition that was not verbatim but retained the same format. For example, "What did he say or do next?" repeated as "What did he do then?"
5. Answer suggested - when the question itself was not repeated directly, but through suggestion of the answer. For example, when an interviewer asked "Right and did you tell anybody what M had done?" to which the child replied "No". The next question "No?" was categorised as an answer suggested repetition.

Repetition purpose

We used the following categories to define the purpose of the repetition:

1. Summarise - to recapitulate or return to a previous section or topic.
2. Clarify or verify - this required an accompanying comment, for example "What I mean is did you…?", "you said", "is that what you said", or similar to be stated.
3. Misheard - when the interviewer or child claimed to have forgotten, misheard, or misunderstood the question; a comment to this effect would be required.
4. Elaborate - to encourage an expansion of the original response from a "mmm", "uhh", "um" or "er" type of answer.
5. Irrelevant - when the response given had no relevance to the question posed.

6. More detail - to elicit more specific information from the child.
7. Silent – when the child gave no response.
8. Leading – where the interviewer's question provided, initiated, or directed the child's response, for example, "He shouldn't do that to you should he?" Coding of a leading question would take precedence even if it followed a silent response from the child (see code 7).
9. Ask again - this was a default category when an alternative purpose could not be deduced.

Consequence of repetition

The response to the repeated question was coded to show whether, and in what way, a change had taken place.

1. Stayed the same as the original response.
2. Added to the original response.
3. Novel response in relation to the original response.
4. Stayed the same as the novel response (this could occur only with second or subsequent repetitions).
5. Added to the novel response (this could occur only with second or subsequent repetitions).

If a child was silent and then responded to a repetition this would be coded as a novel response, similarly, if a child gave an initial response and then remained silent to a repetition this would also be coded as a novel response.

Inter-rater reliability

10% of the transcripts were selected randomly from each age group and coded by a second rater who had been trained in the coding for the style, purpose and consequence of repetition but who was naïve to the hypotheses. The kappa

for repetition style, purpose and consequence were $\kappa = 0.93$, $\kappa = 0.78$, and $\kappa = 0.99$ respectively (all $p < .001$).

Results

The interviews contained many questions although there was considerable variation between interviews. Interviews with 4-5-year-olds contained $M = 216.75$, $SD = 170.27$ questions, with 6-7-year-olds $M = 159.30$, $SD = 82.51$, with 8-9-year-olds $M = 194.75$, $SD = 85.31$, and 10-11-year-olds $M = 183.15$, $SD = 82.55$. The percentages of open questions (closed question percentages are in parenthesis) were 42.0% (58.0%), 45.3% (54.7%), 40.3% (59.7%) and 47.0% (53.0%) for 4-5, 6-7, 8-9, and 10-11-year-olds respectively.

Repetition was found in 93 (98%) transcripts. The following information refers only to the 93 transcripts that contained repetition.

Repeated questions (either as the base question or as a repetition of this question) accounted for $M = 25.3\%$ of all the questions asked.

Position

To assess adherence to the MOGP interviewing protocol the percentage of interviews that included each of the recommended stages was calculated. The data for each age group are shown in Table 2:1. The four interviewing stages were not included in all interviews; omission of the free narrative and closure stages was particularly prevalent in interviews with the 4-5-year olds.

Table 2:1
The frequency of interview stages performed according to age group

	Stage of interview			
	Rapport	Free narrative	Questioning	Closure
Age				
4-5	95.2%	61.9%	100.0%	76.2%
6-7	100.0%	89.5%	100.0%	84.2%
8-9	100.0%	90.0%	100.0%	90.0%
10-11	94.3%	85.7%	100.0%	88.6%

Quantity of repetitions within interview stages (this analysis only uses data from interviews where all stages were performed)

A 4 Age group (4-5, 6-7, 8-9 and 10-11-year-olds) x 4 Interview stage (Rapport, Free narrative, Questioning, Closure) analysis of variance (ANOVA) with repeated measures was applied to the number of question repetitions to determine the frequency of repetition usage according to interview stage.

There was an effect of age group: $F(3,63) = 2.78$, $p < .05$, and an effect of interview stage: $F(3,61) = 46.25$, $p < .001$. There was a two-way interaction between age and interview stage: $F(9,160) = 148.6$, $p < .05$.

The effect of age showed that the 4-5-year-olds experienced repetition more frequently than other age groups $M = 60.71$, $SD = 44.62$, with 6-7-year-olds $M = 42.95$, $SD = 20.64$, 8-9-year-olds $M = 50.55$, $SD = 30.99$, and 10-11-year-olds $M = 35.48$, $SD = 27.83$. A Tukey's honestly significant difference (HSD) comparison showed a difference between the 4-5-year-olds and the 10-11-year-olds only ($p < .05$).

The effect of interview stage showed that the questioning stage contained most repetitions $M = 36.90$ (80.3%), $SD = 28.18$, followed by the rapport stage $M = 8.32$ (17.5%), $SD = 9.35$, the free narrative stage $M = 1.08$ (1.9%), $SD =$

1.73, and the closure stage $M = 0.13$ (0.2%), $SD = 0.56$. Pairwise comparisons showed differences between all interviewing stages ($p < .001$).

The interaction between age and stage (see Table 2:2) showed the highest number of repetitions in the questioning stage particularly in interviews with the 4-5-year-olds. This age group had twice as many repetitions in the rapport and free narrative stages in comparison to the 10-11-year-olds. The 10-11-year-olds consistently received fewest repetitions except for the closure stage where repetitions were nevertheless rare.

Table 2:2
The mean number of repetitions experienced by children in interviews according to age and stage of interview

Age	Interview stage			
	Rapport	Free Narrative	Questioning	Closure
4-5	10.25 (11.24)	1.15 (1.34)	50.24 (40.32)	0.00 (0.00)
6-7	6.89 (5.77)	1.53 (2.43)	34.68 (18.82)	0.00 (0.00)
8-9	12.50 (11.52)	1.44 (2.04)	36.65 (24.47)	0.11 (0.42)
10-11	5.26 (7.10)	0.54 (0.96)	29.85 (23.26)	0.28 (0.84)

Standard deviation in parenthesis

Tests of simple effects for interview stage showed differences in quantity of repetitions for all interview stages ($p < .001$).

Intervals between repetitions

Analysis of repetition interval showed that the shortest repetition intervals predominated $M = 9.46$ (34.1%), $SD = 11.14$, and $M = 10.11$ (36.4%), $SD = 8.67$, for 0 and 1 intervals respectively. For 2-5 interval $M = 4.62$ (16.6%), $SD = 5.20$, 6-20 $M = 91.92$ (6.9%), $SD = 2.48$, and 21+ $M = 1.65$ (6.0%), $SD = 2.27$. Tests of simple effects showed a difference between all intervals ($p < .001$) with the exceptions of between 0 and 1, and between 6-20 and 21+ intervals.

The number of question repetitions according to age group and interval frequency are represented in Table 2:3. This shows that the younger children

experienced more repetitions with shorter intervals than other age groups although as this will be demonstrated later the younger children also remained silent more often than other age groups, which left the interviewer little option except to repeat the question immediately.

Table 2:3
Mean number of repetitions according to repetition interval

	Interval between repetitions				
Age	0	1	2-5	6-20	21+
4-5	14.43	11.91	8.19	3.57	2.29
	(15.60)	(13.14)	(6.96)	(3.70)	(2.80)
6-7	8.74	10.63	3.16	1.53	1.37
	(7.37)	(6.82)	(2.59)	(1.50)	(1.42)
8-9	11.39	11.11	4.72	1.83	1.83
	(13.28)	(5.53)	(3.72)	(2.53)	(2.07)
10-11	6.15	7.79	3.24	1.29	1.38
	(7.36)	(7.53)	(5.03)	(1.51)	(2.47)

Standard deviation in parenthesis

Repetition sets

The numbers of repetitions and repetition sets (the number of repetitions in addition to the initial question) found differed according to the stage of the interview. The maximum number of repetitions in the rapport stage was 46 (4-5-year-old), 7 in the free narrative stage (8-9-year-old), 170 in the questioning stage (4-5-year-old), and 3 in the closure stage (10-11-year-old). The minimum number of repetitions was 0 in all stages except for the questioning stage where the minimum was 1.92 (10-11-year-old). In the questioning stage (which accounted for 80.3% of all repetitions) the 4-5-year-old children experienced the highest mean repetition sets $M = 16.10$, $SD = 10.25$ with $M = 2.99$ repetitions per set of repetitions compared with $M = 14.21$, $SD = 7.79$ with $M = 2.46$ repetitions per set for 6-7-year-olds, $M = 14.35$, $SD = 9.31$ with $M = 2.58$ repetitions per set for 8-9-year-olds, and $M = 12.15$, $SD = 9.00$ with $M = 2.41$ repetitions per set for 10-11-year-olds.

There was a wide range of the number of repetitions within the questioning stage (represented in Table 2:4). The widest range of repetitions was in interviews with the 4-5-year-olds who also experienced the highest mean number of repetitions.

Table 2:4
Repetitions used in the questioning stage of interviews according to age group

	Questioning stage	
Age	Mean	Range
4-5	50.24 (40.32)	2-170
6-7	34.68 (18.82)	6-70
8-9	36.65 (24.47)	10-94
10-11	29.85 (23.26)	4-111

Standard deviation in parenthesis

Style

The 'neutral' categories of repetition accounted for 7.0% of repetitions across all age groups (*misheard* = 4.1%, *linguistic error* = 0.3%, *rephrase* = 2.2%, *sound* = 0.4%).

The 'with motive' categories accounted for 93.0% of repetitions (*gist* repetition = 53.4%, *verbatim* = 9.2%, *closed to open* =13.5%, *open to closed* = 12.3% and *answer suggested* = 4.6%). Table 2:5 shows the styles of repetitions experienced by each age group; the means represent the repetitions within each age group (and not as a mean of the total repetitions). Table 2:5 also shows that 4-5-year-olds experienced most *verbatim* repetitions (16.2%) and least *closed to open* (11.7%), *open to closed* (10.7%) and *answer suggested* (3.8%). The highest amount of *answer suggested* repetitions were found for 10-11-year-olds (6.1%) closely followed by 8-9-year-olds (5.9%).

Table 2:5
Mean number of repetitions according to style of repetition and age group across the entire interview

Repetition style	Age group 4-5	6-7	8-9	10-11
Misheard	0.76 (0.89)	1.32 (2.79)	1.30 (1.53)	1.24 (2.33)
Linguistic	0.05 (0.22)	0.05 (0.23)	0.15 (0.37)	0.06 (0.24)
Rephrase	0.81 (1.83)	1.32 (1.49)	0.35 (0.59)	0.33 (0.65)
Sound	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.33 (0.65)
Verbatim	6.57 (6.27)	1.63 (1.64)	1.50 (1.73)	1.36 (1.75)
Open to closed	4.33 (4.15)	3.53 (3.32)	3.65 (2.72)	2.82 (3.24)
Closed to open	4.76 (4.61)	4.53 (4.48)	3.90 (3.29)	2.79 (3.33)
Gist	21.76 (19.62)	12.58 (6.22)	17.85 (13.21)	10.82 (9.65)
Answer suggested	1.52 (2.86)	0.53 (0.90)	1.80 (2.50)	1.27 (1.82)

Standard deviation in parenthesis

Purpose

Across age groups the most common purposes for repetition were: as a result of *silent response* $M = 10.38$ (36.6%), $SD = 13.01$, to *elicit more detail* $M = 6.98$ (24.6%), $SD = 6.56$, or to *clarify or verify* $M = 3.70$ (13.0%), $SD = 3.80$. The lowest number of repetitions were *leading* questions $M = 0.55$ (1.9%), $SD = 1.03$. The other purposes were: to *ask again* $M = 2.33$ (8.2%), $SD = 3.18$, *irrelevant* $M = 1.52$ (5.3%), $SD = 2.74$, *misheard* $M = 1.38$ (4.9%), $SD = 2.40$, to *summarise* $M = 0.91$ (3.2%), $SD = 1.31$, and to *elaborate* $M = 0.62$ (2.2%), $SD = 1.23$.

Gist repetition was the most common form of repetition in the transcripts and also a style recommended in interviewing protocols (Home Office, 1992,

2001). The 4-5-year-olds gave the most *silent* or *irrelevant* responses to *gist* repetition and were not as often asked repeated questions to *clarify* or provide *more detail* as older age groups as shown in Table 2:6.

Table 2:6
The purposes of *gist* repetitions only according to age group

	Age group			
Purpose	4-5	6-7	8-9	10-11
Silent	44.2%	37.8%	33.6%	30.0%
More detail	21.2%	24.8%	27.3%	24.9%
Clarify	9.4%	14.7%	18.0%	21.4%
Irrelevant	10.7%	2.9%	3.3%	2.9%
Leading	0.7%	0.0%	1.2%	0.0%
Other *	13.8%	19.8%	16.5%	20.9%

**Combines: summarise, misheard, elaborate, and ask again*

There was a relationship between the purpose and the style of repetition. Approximately half of the *answer suggested* repetitions were not leading but repeated to *clarify* a previous response. Repetition from *closed to open* was most commonly to *elicit more details*. The most frequent reason for *open to closed* repetitions was as a result of an initial *silent response.*

Consequence

To establish whether question repetition affected children's responses, changes in responses were examined. Across all age groups repetitions led to a change in $M = 6.19$ (75.2%), $SD = 4.29$ of responses comprising $M = 1.69$ (20.5%), $SD = 1.12$ additional information, and $M = 4.51$ (54.7%), $SD = 3.16$ novel information. Table 2:7 shows that the pattern of consequences of repetition for the 4-5-year-olds involved them remaining with their original response more often than other age groups.

Table 2:7
The consequence of responses to repetitions according to age group

	Age group			
Consequence	4-5	6-7	8-9	10-11
Same as original	32.3%	24.9%	27.7%	26.0%
Adds to original	9.0%	19.6%	19.5%	19.5%
Novel	44.9%	47.5%	44.1%	46.8%
Same as novel	9.4%	3.9%	5.2%	4.1%
Adds to novel	4.4%	4.2%	3.5%	3.6%

The most common purpose for repetition was due to an original *silent* response; 4-5-year-olds maintained this response (i.e. silence) in 46.9% of their subsequent responses. The equivalent for the other age groups was lower with 21.1% for 6-7-year-olds, 29.1% for 8-9-year-olds and 15.1% for 10-11-year-olds maintaining silence.

Discussion

Our analysis showed that interviewing protocols were not being adhered to in respect to interview structure and recommended question format usage (Home Office, 1992). We also found that the interviewing style, the apparent purpose for, and the consequence of repetition differed according to the age of the child.

The failure to perform all of the interviewing stages (particularly the free narrative stage), and the predominant use of closed questions, supported concerns that protocols were not fully implemented (Cederborg et al., 2000; Davies et al., 2000; Thoresen et al., 2006; Warren et al., 1996; Westcott & Kynan, 2006). The omission of the free narrative stage (Davies et al., 1995), particularly with 4-5-year-olds (see Table 2:1) meant that the stage which usually elicited the most accurate information was not included in many interviews with the youngest age group (Hershkowitz, 2001; Holliday, 2003a; Jones & Pipe,

2002; Korkman et al., 2006; Lamb & Fauchier, 2001; Lamb et al., 2001; Lamb et al., 2003a; Orbach & Lamb, 2000; Peterson, 2002).

Previous researchers have established that repetition of questions may lead to changes in responses, which were generally detrimental to children's accuracy and consistency (Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Moston, 1987; Powell & Thomson, 1996) and yet repetition was used in all but two of the interviews and involved over a quarter of all questions asked (including the initial question).

It is possible that a single question or little used question type could have a disproportionate effect (detrimental or beneficial) on the child's response and subsequent credibility; however, we examined the consequences of all repetitions regardless of content. The effect of that repetition, consistent with previous research, was age related; changes in responses decreased with an increase in age (Howie et al., 2004; Poole & White, 1991).

Our results also showed that the 4-5-year-olds experienced most repetitions (see Table 2:2), which were generally immediate or close together (see Table 2:3). Whilst most of the repetitions were found for all age groups in the questioning stage (see Table 2:4) both the mean and range number of repetitions were highest for 4-5-year-olds. These children were least consistent, and provided the least additional information. The interviewing style used with 4-5-year-olds was therefore different from that used with the older age groups (see Table 2:5). This difference may have been a consequence of the large number of times that the 4-5-year-olds remained silent in response to questions, and showed that repetition was not a successful way to encourage reluctant 4-5-year-olds to answer a question.

The motivations for using repetition and the types of repetition were generally appropriate. In line with the recommended guidelines (Home Office, 1992) there were low numbers of leading questions and a predominance of *gist* repetitions. The use of repetition was mainly because a child did not give a response to the original question, or because further information was required (see Table 2:6). Changes in responses did not always result in children contradicting their original response as further information was elicited in almost one fifth of responses to repetitions.

Over three quarters of responses to repetitions represented a change in response. The frequent changes in responses to repetitions supported those researchers who have suggested that children understood repetition to mean that a change in response was necessary (Garven et al., 1998; Gilstrap, 2004; Howie et al., 2004; Krähenbühl & Blades, 2006). This suggested that children may have been affected by their schema (Bartlett, 1932; Roberts, 2002) representing how discussions function with adults (particularly relatively unfamiliar adults) who repeat questions. The differences in interviewing styles experienced by the different age groups could also be attributed to the interviewers' schema concerning the conduct of age-appropriate conversations with children. The veracity of the responses given by the children in the transcripts could not be assessed. However, the children changed their responses to repeated questions, which adversely affected their consistency and possibly their credibility (Davis et al., 1999). Although children often changed their responses to repetitions, this was often in response to legitimate requests for more detail or when a response to an initial question had not been forthcoming.

Children's maintenance of their original response may be beneficial because experimental research has shown the first response is generally the most accurate (Moston, 1987; Poole & Lindsay, 1995; Warren et al., 1991). Original information was maintained in a quarter of the responses to repetition. The majority of 4-5-year-old children who maintained their original responses were those who remained silent. This may not have been useful to interviewers, but may have reflected children's resistance to answer questions when they did not know the answer. Changes in responses were more frequent than maintaining or adding to an original response in all repetition types (see Table 2:7). *Open to closed* repetitions, the direction of questioning advocated by the interviewing MOGP protocol (Home Office, 1992), resulted in the highest number of novel (unrelated to the original) responses. If the original response was likely to be the most accurate this indicated a problem with the movement from an *open to closed* questioning procedure.

Novel responses, regardless of accuracy, demonstrated a lack of consistency which could adversely affect the likelihood that a court case would proceed (Davis et al., 1999; Gallagher & Pease, 2000). To minimise this risk interviewing protocols caution against repetition and encourage the use of open-ended questions (Home Office, 1992; Orbach et al., 2000). Nevertheless, our results showed that open-ended repetitions (from *closed to open)* resulted in high frequency of changes in responses. Even if these changes resulted in improved accuracy (which we were unable to verify) the problem with a lack of consistency persisted.

The age differences both in interviewing procedures experienced by the children and responses to repetition styles, particularly between the 4-5-year-olds

and the other age groups suggested that current interviewing protocols such as the ABE (Home Office, 2001) using similar structures to the MOGP (Home Office, 1992) may not be equally appropriate for all age groups.

In Study 1 we were unable to verify the effect of different question repetition styles on children's accuracy. Therefore, in Study 2 (Chapter 3) an experiment was conducted, to examine the effects of the four repetition forms (*verbatim, gist, open to closed* and *closed to open*) that were most frequently used in the police interview transcripts of Study 1. We examined the effect those forms of repetition had on accuracy and consistency in responses when we questioned children about a staged event. Additionally, we examined the direction of changes in responses in reference to accuracy; did changes enhance accuracy, have no effect, or were they detrimental.

Chapter 3

Study 2

In Study 2 we implemented the repetition styles used by police interviewers in their interviews with young children in Study 1 to assess the effect of each repetition style on children's accuracy and consistency in responses. In addition we assumed that not all of the information required to answer questions accurately would be available to interviewees. Therefore, we included questions in our recall interview that were, in effect, 'unanswerable' in respect to the information provided by the event.

Introduction

During everyday conversations it is not unusual for individual questions to be repeated on more than one occasion. Additionally, as discussed in Chapter 1, the aim of obtaining an accurate and complete account of an experienced event is generally realised through asking questions, which may be repeated.

Researchers have established that the amount and accuracy of details elicited by interviewers is related to the type of question asked. Open invitations, where the interviewee is asked to, for example "tell me everything about …", and specific questions, such as those starting with "who, what, how, when, where, why" elicit more accurate and detailed responses than other types of closed questions where the range of response was limited (Dent & Stephenson, 1979; Hershkowitz, 2001; Lamb et al., 1996; Orbach & Lamb, 2000; Peterson et al., 1999).

Previous researchers have shown that both the number and the accuracy of details elicited in interviews increased with age (Beuscher & Roebers, 2005; Goodman et al., 1994; Krähenbühl & Blades, 2006; Poole & White, 1991).

Younger children did not always provide sufficient information or as much information as older children in response to open-ended question forms (Ceci & Bruck, 1995; Fivush et al., 2002).

Children are sometimes reluctant to acknowledge that they do not know the accurate response to a question. Researchers have shown that children provided responses to nonsensical or bizarre questions (Hughes & Grieve, 1980; Waterman et al., 2000). Children also provided inaccurate speculative responses to unanswerable questions even when they had been instructed that to say "I don't know" was acceptable (Beuscher & Roebers, 2005; Krähenbühl & Blades, 2006; Waterman et al., 2001, 2004).

When a child's response to a question is inadequate in any way (through a lack of informative detail, a contradiction to an earlier response, is totally irrelevant etc) then the question may be repeated. However, question repetition may not result in the expected or desired response, because it has been shown that the repetition of a question within a single interview can have a detrimental effect on accuracy (Memon & Vartoukian, 1996; Poole & White, 1991; Warren et al., 1991).

According to fuzzy-trace theory accuracy of a response to a question would depend on the form of memory representation accessed. Fuzzy-trace theory (see Chapter 1, pp. 37-39) associates the encoding of different memory representations depending on the delay between the event and the recollection of that event (Reyna & Kiernan, 1994). Reyna and Kiernan (1994) suggested that verbatim memory would diminish with time but that gist representations would remain stable after a delay. Gist encoding could be advantageous in increasing

accuracy if the repeated questions (especially those in non-verbatim style) gave access to this encoding (Miller & Bjorklund, 1998).

Repetition of questions is a strategy available to interviewers to elicit the required information. Interviewing protocols provide guidance on interviewing procedures (see Chapter 2, pp. 51-54). Concern over the possible detrimental effects of question repetitions is specifically recognised in the ABE interviewing protocol (Home Office, 2001). ABE not only cautions against the use question repetition but also suggested that, in respect to specific questions, if repetition was necessary the question should "not be repeated in the same form... Children may interpret this as a criticism of the earlier response and sometimes change their response as a consequence." (Home Office, 2001, p. 45).

Not all changes in responses to repeated questions will necessarily be detrimental to accuracy because some may also include a change from an inaccurate to an accurate response. Howie, Sheehan, Mojarrad and Wrzesinska (2004) defined such changes as "desirable" shifts. However, Howie et al. (2004) found that the majority of changes in responses were "undesirable" shifts; when an accurate response became inaccurate following question repetition. In Howie et al. the number of changes made in responses to repeated questions decreased with age, these results being consistent with other research (Ghetti et al., 2002; Poole & White, 1991; Roebers & Schneider, 2002).

It is possible that the form of repeated question may have an effect on the accuracy of a response to that repetition. Therefore, in this experiment we conducted a detailed examination of the effect of different question repetition forms. We implemented the following question repetition forms: *gist*, *verbatim*,

closed to open-ended, and *open-ended to closed* which were the most commonly used forms in Study 1.

We expected, in accordance with Poole and White (1991), that accuracy and consistency in responses to repetitions would increase with age, and that repetition would lead to a decline in accuracy for each age group (Memon & Vartoukian, 1996; Warren et al., 1991). The ABE (Home Office, 2001) protocol implied that verbatim repetition encouraged children to change their responses so we expected that there would be more changes in responses to verbatim repetitions than in responses to other forms of repetition.

In accordance with Howie et al. (2004) we expected that undesirable changes in responses to repetitions (accurate responses becoming inaccurate) would exceed desirable responses (when inaccurate responses became accurate).

We also expected that if a child changed their initial response as a result of question repetition, this novel response would be maintained in response to a subsequent repetition (Krähenbühl & Blades, 2006).

Method

All procedures were in accordance with the British Psychological Society ethical guidelines. Ethical approval for the experiment was obtained through the University of Sheffield's Department of Psychology ethics committee.

The interviewer and collaborators were all teacher-trained and had been previously employed as teachers in England. Recruitment of participants took place through schools. School and parental approval was obtained prior to the experimental event, and child's assent was obtained immediately before the interview commenced.

At the beginning of the individual interviews the interviewer told each child that they could withdraw at any time. If a child did not reply to a question the interviewer would not encourage a different response by a further repetition of the question or by asking the child to "please answer the question I asked you". A child who gave no response to five successive questions was assumed to wish to withdraw and the interview was terminated (although this situation did not occur).

Interviewer

A single interviewer completed all of the interviews. This interviewer was aware of the event used in the experiment but was not present during the event presentation. It was possible that the children's responses could be attributed to the specific interview style and personal characteristics of an individual interviewer. It was acknowledged that the children's responses could be affected by the interviewer's practice and characteristics and that another interviewer could possibly elicit different responses from the children through implementation of an alternative interviewing style. However, it was considered preferable to maintain consistency of interviewing practice by using the same interviewer for all of the experimental interviews.

Participants

The children came from lower to middle-class families and lived in or near a small market town in England. All children had English as their first language.

One hundred and sixty children from three different age groups completed the study: fifty-six 8-9-year-olds (n = 33 boys and n = 23 girls, $M = 8$ years 10 months, $SD = 3.74$ months), fifty-four 6-7-year-olds (n = 30 boys and n

= 24 girls, *M* = 6 years 10 months, *SD* = 3.49 months) and fifty 4-5-year-olds (n = 22 boys and n = 28 girls, *M* = 4 years 10 months, *SD* = 3.38 months). A further three children were interviewed but the data for these children were not included as one child's interview was interrupted and two children withdrew by stating that they did not want to answer any more questions but wanted to return to their classes.

Procedure

Children watched a staged live presentation (see Appendix 1 for a copy of the script) in their schools at the beginning of the week. Each group of children was told that a woman called Chris was going to give a talk about dog care. The class teachers had been informed as to the subject matter, the duration and structure of the presentation, and that another adult would enter the room at as specific point during the talk. The entrance of another adult was intended to reflect the event used by Poole and White (1991) and also to provide an unusual element to the event. The adult would enter, make a brief statement, look in the confederate's handbag and leave the room with the confederate's car keys.

The 15-minute presentation contained two sections presented concurrently: First, a talk about the presenter's dog, dog characteristics, care and associated equipment. Second, a reading of an unfamiliar illustrated story about a dog and its fleas (Wyllie & Brown, 2003). The props used during the event included a poster of dog breeds, a tin and a bag of dog food, a purple dog lead, a 'pooper scooper' (a plastic device for collecting dog excrement) and the storybook. The props used were kept in a bag until required and then remained visible for the remainder of the presentation. At the end of the presentation the

presenter prevented any further interaction with the children by tidying the equipment away.

The presentation was fully scripted, but was also audio taped so that any deviations from the script could be noted in case they affected the child's information and opportunity to make a response.

Interview procedure

Approximately one week later (7, 8 or 9 days) an unfamiliar adult interviewed the children individually. The teachers informed the children that the interviewer wanted to talk to them about the presenter's visit.

Each child was interviewed in an otherwise unoccupied room within the child's school. The interview commenced with instructions and explanations. If the child did not recall the to be remembered event a prompt was provided once only. The children were explicitly told "If you can't answer the question then don't worry, it doesn't matter, you should just say that you don't know or that you've forgotten, or something like that."

The interviews lasted approximately five minutes, were audio taped and subsequently transcribed. The children were all praised for their responses (regardless of attainment) and thanked for their participation; they did not otherwise receive any incentive for taking part.

Questions asked

The children were asked a total of 48 questions. Although the children were unaware of the question organisation the questions were, in effect, in three blocks: the base block with the initial questions (16 questions), the first repetition block and the second repetition block where the initial questions were repeated (16 questions in each repetition block). Half of the 16 questions in each block

were answerable (the information for an accurate response was provided in the presentation) and half were related to the presentation but were unanswerable (the information to answer these questions was not provided in the presentation). The repetitions of the base block questions were in either *verbatim, gist, open to closed* or *closed to open* form, giving rise to four base questions in each repetition form. Below is a full set of questions:

The repetitions have had their repetition code added to demonstrate the repetition pattern of this set of questions.

V = verbatim, C = closed to open, O = open to closed, G = gist

Base block

1. In the story which animal wanted to eat the ducks?
2. Did the chickens in the story get eaten?
3. Where did Chris keep her tissues?
4. What colour was Chris's scarf?
5. Did Chris wear a hat?
6. Was the dog lead purple?
7. In the story why did the dog keep on scratching?
8. At the end of the story did the fox get fleas?
9. Where does Chris take Susie for her walks?
10. Did Susie have clean teeth?
11. In the story how many chickens were there?
12. What was the name of the man who came into the classroom?
13. In the story did the fox have a sore patch on its leg?
14. Did Chris have a pen in her handbag?
15. In the story how old was the fox?

16. In the story did the dog get tired when he carried the sack of chickens?

Repetition block 1

1. V - In the story which animal wanted to eat the ducks?
2. C - What happened to the chickens in the story?
3. O - Were the tissues in Chris's handbag?
4. G - What did Chris's scarf look like?
5. V - Did Chris wear a hat?
6. C - What colour was the dog lead?
7. O - Did the dog stop scratching when he had fleas?
8. G - Did the fleas live on the fox at the end of the story?
9. V - Where does Chris take Susie for her walks?
10. C - What were Susie's teeth like?
11. O - Were there 10 chickens in all?
12. G - What was the man was called?
13. V - In the story did the fox have a sore patch on its leg?
14. C - What else did Chris have in her handbag?
15. O - Was the fox five years old?
16. G - Was it hard work for the dog when he carried the sack of chickens?

Repetition block 2

1. V - In the story which animal wanted to eat the ducks?
2. C - Where did Chris get her tissues out from?
3. O - Did the chickens escape from the fox?
4. G - How would you describe Chris's scarf?
5. V - Did Chris wear a hat?
6. C - What was making the dog itch?

7. O - Was the dog lead yellow?
8. G - Did the fox get the fleas on him at the end?
9. V - Where does Chris take Susie for her walks?
10. C - How many chickens were in the story?
11. O - Did Susie's teeth have bits of food on them?
12. G - The man's name, what was it?
13. V - In the story did the fox have a sore patch on its leg?
14. C - What was the fox's age?
15. O - Did the handbag have a pen in it?
16. G - Did carrying the sack of chickens wear the dog out?

The order of the form of question repetition was counterbalanced giving rise to four different sets of questions. Question order was manipulated so that an *open to closed* repetition was not applied to the same question twice. The order of the repetition forms was replicated within the two repetition blocks so that the interval (provided by intervening questions) between equivalent repetition forms was uniform.

The event developed was intended to be engaging, slightly unusual and of a length that would be appropriate to all age groups. The observation event and interviewing timescale used in the experiment replicated the structure (Poole & White, 1991) and details (Krähenbühl & Blades, 2006) of events used in previous research.

The age of participants ranged from 4 years to 9 years. This age range replicated the range of children's ages used by Poole and White (1991).

Accuracy and Change coding

The children's responses were coded for accuracy and for change in response in comparison to the previous response elicited. Changes in response did not necessarily represent a shift in accuracy. A change in response could represent a change from inaccuracy to accuracy, from accuracy to inaccuracy, or a novel inaccurate response. A novel inaccurate response was different to the previous response and thereby represented inconsistency but did not result in a change in accuracy.

Responses to answerable questions were scored as correct if they included appropriate information taken from the event. Unanswerable questions could not be answered from the information presented in the event and so a correct response to an unanswerable question was 'don't know', 'I can't remember' or a 'She [the presenter, Chris] didn't say' type of response. Any attempt to provide a subjective response such as 'I think so' to an unanswerable question was scored as incorrect.

A change in response was defined by the provision of a contradictory or novel response in comparison to a previous response to the same question. For example, if in answer to questions concerning the number of chickens a child initially said there were seven chickens and then said "Yes" to a question asking whether there were 10 chickens this would be coded as a change. If no semantic change was made the response was defined as having stayed the same even if the wording had changed. For example, in respect to the number of chickens, "It was seven" and "The chickens had lots of feathers and there were seven" would not be coded as a change, because the number required by the question, remained the same. If a child provided a change in response to a first repetition and then

maintained this response to a second repetition then this was coded as one change having taken place. If, after making a change in response to a first repetition a child returned to his or her initial response after a second repetition this would represent two changes having taken place.

There were 5 possible patterns of change: when no change was made (XXX), a change made and then maintained (XYY – one change), a change made but then returned to the initial response (XYX – two changes), a change made to the second repetition only (XXY one change) or changes made to both repetitions (XYZ – two changes).

Direction of change referred to a shift from an accurate to an inaccurate response ("undesirable") or vice versa ("desirable"), as defined by Howie et al. (2004). In addition, we included novel inaccurate shifts where an initial inaccurate response was changed to a different inaccurate response.

Any comments made by the children relating to repetition during the interview were recorded.

Inter-rater reliability

Assessment of inter-rater reliability for coding was based on a random sample of 10% of transcriptions. An independent judge trained in the coding procedure and familiar with the event, but naïve to the hypotheses coded the transcriptions.

The children's responses were re-coded by the inter-rater for accuracy, and changes in responses. The kappa for accuracy and for change codings were $\kappa = 0.98$ and $\kappa = 0.93$ respectively (both $p < .001$).

Results

Analyses between female and male participants, and between different lengths of delay between the event and the interview (7, 8 or 9 days) in each age group revealed no significant differences. Therefore, these factors were not considered any further.

Accuracy of all responses (including the initial base question)

A 3 Age group (4-5, 6-7 and 8-9-year-olds) x 3 Question order (base block, 1st repetition block, 2nd repetition block) analysis of variance (ANOVA) with repeated measures was applied to the children's responses.

There were main effects of age group: $F(2, 157) = 20.16$, $p < .001$, and of question order: $F(2, 156) = 16.92$, $p < .001$. There was an interaction between age and question order: $F(4, 312) = 2.61$, $p < .05$.

The mean accuracy for 4-5-year-olds was $M = 6.43$ (40.2%), $SD = 1.88$, for 6-7-year-olds $M = 8.01$ (50.1%), $SD = 2.03$, and for 8-9-year-olds $M = 8.56$ (53.5%), $SD = 2.15$. A Tukey HSD comparison showed a difference between the 4-5-year-olds and the other age groups, $p < .001$, but not between the 6-7 and the 8-9-year-olds.

There was a decline in accuracy from $M = 8.16$ (51.0%), $SD = 2.25$ in responses to the initial base questions, to $M = 7.41$ (46.3%), $SD = 2.17$ in the first repetition block, and to $M = 7.56$ (47.3%), $SD = 2.17$ in the second repetition block. Pairwise comparisons on response accuracy showed a difference between the base and both repetition blocks, $p < .01$, but not between the repetition blocks.

The interaction between age and order showed that there was a greater decline in accuracy with the younger children's (4-5 and 6-7-year-olds)

responses to questions in the first repetition block. Tests of simple effects showed a difference ($p < .001$) between the 4-5-year-olds and both other age groups in all question orders. These tests showed a difference between the 6-7 and the 8-9-year-olds in the first repetition block responses only ($p < .05$.)

Accuracy of responses to repeated questions only

A 3 Age group (4-5, 6-7 and 8-9-year-olds) x 4 Repetition style (*verbatim*, *gist*, *closed to open*, *open to closed*) x 2 Repetition order (1st repetition, 2nd repetition) X 2 Repetition type (answerable, unanswerable) analysis of variance (ANOVA) with repeated measures was applied to the children's responses.

There were effects of age: $F(2, 157) = 18.74, p < .001$), style: $F(3, 155) = 4.69, p < .01$) and type: $F(1, 157) = 322.86, p < .001$) but no effect of order. There were two-way interactions between style and type: $F(3, 155) = 47.13, p < .001$), style and order: $F(3, 155) = 4.17, p < .01$), and type and order: $F(1, 157) = 25.63, p < .001$). There was a three-way interaction between style, type and order: $F(3, 155) = 6.11, p < .001$). There were no interactions with age.

The effect of age showed an increase in accuracy of responses with age: 4-5-year-olds $M = 0.78$ (39.1%), $SD = 2.92$, 6-7-year-olds $M = 0.95$ (47.6%), $SD = 3.83$, and 8-9-year-olds $M = 1.06$ (52.1%), $SD = 4.45$. A Tukey HSD comparison on response accuracy showed a significant difference ($p < .001$), between the 4-5-year-olds and both 6-7 and 8-9-year-olds (the difference between the 6-7 and the 8-9-year-olds was close to significance $p = .053$).

The effect of style showed that responses to *open to closed* repetitions were least accurate $M = 0.87$ (43.7%), $SD = 1.21$, accuracy for responses to

verbatim, closed to open and *gist* repetitions were $M = 0.95$ (47.9%), $SD = 1.62$, $M = 0.95$ (47.6%), $SD = 1.62$, and $M = 0.96$ (48.0%), $SD = 1.46$ respectively.

Pairwise comparisons showed a difference between the *open to closed* repetition style and all other styles only, $p < .01$.

The effect for question type showed that responses to answerable questions were $M = 1.32$ (66.4%), $SD = 2.65$, and for responses to unanswerable questions $M = 0.54$ (27.2%), $SD = 3.31$. Pairwise comparisons showed a difference between these question types, $p < .001$.

The three-way interaction showed that accuracy increased through repetitions for answerable questions when the question format changed (from *open to closed* or *closed to open*). There was an increase in accuracy from 52.7% to 67.7% for responses to *closed to open* questions, and from 70.5% to 76.5% for responses to *open to closed* questions (from 64.7% to 64.4%, 66.3% to 65.7% for responses to *verbatim* and *gist* responses respectively). The accuracy of responses for unanswerable questions declined between the first and second repetitions in all question styles although the decline was greatest in responses where the question structure changed. The decline in accuracy of responses was from 31.7% to 29.4% for *verbatim* questions, from 36.2% to 32.7% for *closed to open* questions, from 15.6% to 11.1% for *open to closed* questions and from 30.6% to 28.7% for *gist* questions.

Tests of simple effects were conducted for the accuracy of responses to answerable and unanswerable questions separately. These showed that for responses to answerable questions there was a difference in accuracy between repetition order ($p < .05$). There was also a difference between the *open to closed* repetition style and the other repetition styles ($p < .01$ with *verbatim* and *gist*, $p <$

.001 with *closed to open*). For responses to unanswerable questions the tests of simple effects again showed a difference between *open to closed* repetition style with all other repetition styles ($p < .001$) but no difference between repetition orders.

Changes in responses

We found that there were no changes at all (i.e. responses remained completely consistent across all repetitions) in responses to 56.3% of questions posed. Of these consistent responses 71.6% were accurate, and 28.4% were inaccurate.

Repeated questions resulted in changes (in comparison to the previous response given by the child) in 29.6% of all responses to repetitions, 24.4% of all responses to answerable questions and 34.7% of all responses to unanswerable questions.

Changes in responses

A 3 Age group (4-5, 6-7 and 8-9-year-olds) x 4 Repetition style (*verbatim, gist, open to closed, closed to open*) x 2 Repetition type (answerable, unanswerable) x 2 Repetition order (1st repetition, 2nd repetition) analysis of variance (ANOVA) with repeated measures was applied to the children's responses.

There were main effects of age group: $F(2, 157) = 6.58, p < .01$, of repetition style: $F(3, 155) = 66.80, p < .001$, repetition type: $F(1, 157) = 57.87, p < .001$, and of repetition order: $F(1, 157) = 15.8, p < .001$.

There were two-way interactions between age and type $F(2, 157) = 4.41$, $p < .05$, style and type $F(3, 155) = 13.37, p < .001$, style and order $F(3, 155) = 27.84, p < .001$, and type and order $F(1, 157) = 5.26, p < .05$.

There were three-way interactions between age group, style and order $F(6, 310) = 2.85, p < .05$, and between style, type and order $F(3, 155) = 2.90, p < .05$.

The effect of age showed that changes in responses decreased with age: $M = 0.68$ (34.0%), $SD = 4.36$ for 4-5-year-olds, $M = 0.58$ (29.2%), $SD = 3.35$ for 6-7-year-olds, and $M = 0.52$ (25.8%), $SD = 3.46$ for 8-9-year-olds. A Tukey HSD comparison showed a difference between the 4-5-year-olds and the other age groups, $p < .001$, but not between the 6-7-year-olds and the 8-9-year-olds.

The effect of style showed that changes in structure (*open to closed* $M = 0.77$ (38.4%), $SD = 1.61$, *closed to open* $M = 0.80$ (39.8%), $SD = 1.47$) resulted in more changes in responses than when there was a change in words only (*gist* $M = 0.48$ (24.0%), $SD = 1.48$) or no change at all (*verbatim* $M = 0.32$ (15.9%), $SD = 1.41$). Pairwise comparisons showed a difference between all repetition forms ($p < .001$) with the exception of between *open to closed* and *closed to open* repetition forms.

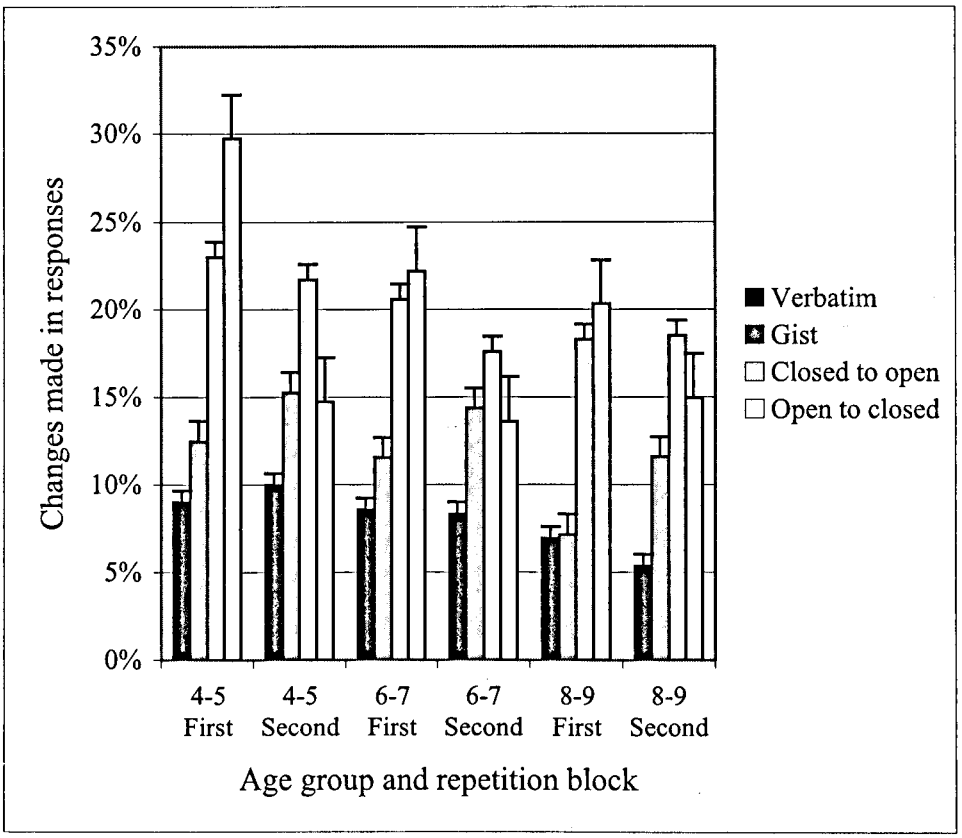
The effect of type showed that there were fewer changes in responses to answerable questions ($M = 0.49$ (24.3%), $SD = 2.36$) than changes in responses to unanswerable questions ($M = 0.70$ (37.7%), $SD = 2.38$). Pairwise comparisons showed a difference in the number of changes in responses between answerable and unanswerable questions ($p < .001$).

The effect of order showed that changes reduced with repetitions from $M = 0.63$ (31.4%), $SD = 2.25$ for the first repetition to $M = 0.55$ (27.6%), $SD = 2.11$ for the second repetition. Pairwise comparisons showed a difference between the first and second repetitions $p < .001$.

The three-way interaction between age group, style and order (see Figure 3:1) showed that changes decreased with age and with second repetitions particularly with the *open to closed* repetition style. Tests of simple effects confirmed differences in change between all age groups in responses to the first repetition (between 4-5-year-olds and 8-9-year-olds $p < .001$, between other age groups $p < .05$) but only between the 4-5-year-olds and the 8-9-year-olds in responses to the second repetition $p < .05$.

Changes were most prevalent in question forms where there was a change in structure; from *open to closed.* Tests of simple effects showed differences in the number of changes made between 4-5-year-olds and 8-9-year-olds in all repetition styles ($p < .05$), between 4-5-year-olds and 6-7-year-olds in *open to closed* repetition style only ($p < .05$), and between the 6-7-year-olds and the 8-9-year-olds in the *gist* repetition style only ($p < .05$).

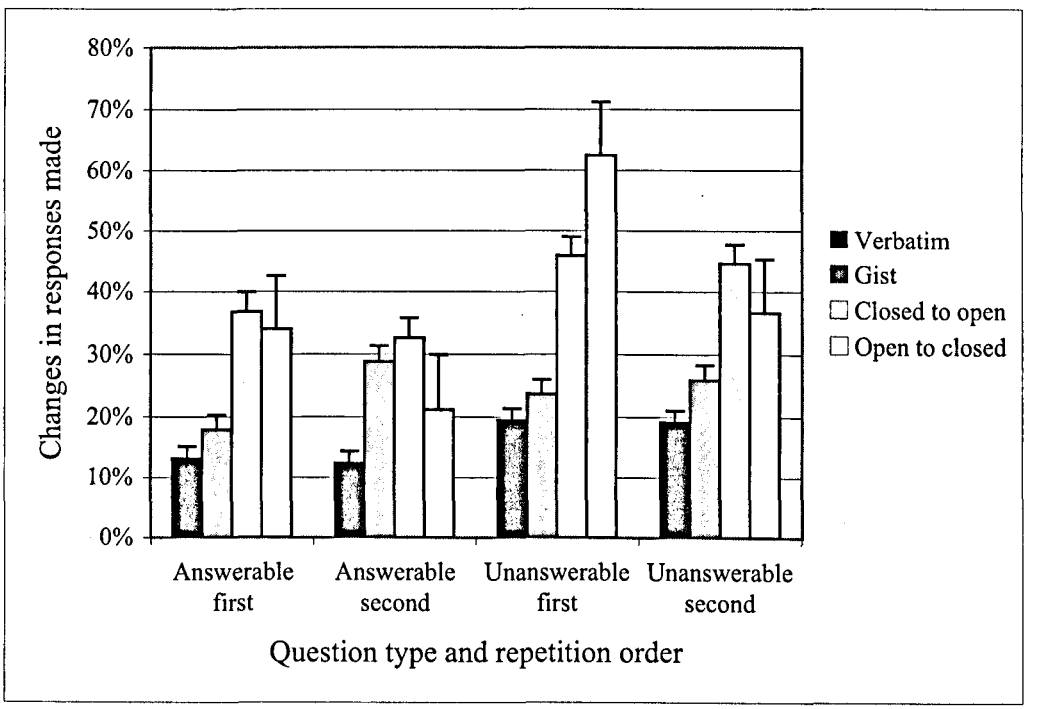
Figure 3:1
Changes in responses according to repetition style, age group and repetition order



The three-way interaction between repetition style, order and question type (see Figure 3:2) showed a difference in number of changes between responses to answerable and unanswerable questions according to repetition style and order. Changes in responses to a first repetition of unanswerable questions were highest particularly amongst question styles that provided a change in structural format. Changes in responses to *verbatim* repetitions also increased with unanswerable repetition types but remained least prone to change regardless of order or question type. Tests of simple effects showed that, with the exception of *gist* repetitions, there was a difference in the number of changes according to repetition style between answerable and unanswerable question types ($p < .05$). Further tests of simple effects were conducted on answerable and unanswerable responses separately. These showed that there was a difference in the number of

changes according to repetition order for unanswerable questions only ($p < .01$). For changes in responses to answerable repetitions tests of simple effects showed that there were differences between all repetition styles ($p < .001$, between *open to closed* and *closed to open* ($p < .01$) with the exception of between *open to closed* and *gist* repetitions. With responses to unanswerable repetitions there were differences between all repetition styles ($p < .001$) with the exceptions of between *verbatim* and *gist*, and between *open to closed* and *closed to open* repetitions.

Figure 3:2
Changes in responses according to repetition style, type of question and repetition order



Desirable and undesirable shifts in responses

The accuracy totals for each repetition order might have masked the effects of changes in responses with an improvement in one response counteracting a decline in another.

An initial inaccurate response that resulted in a novel but also inaccurate response after repetition represented a shift that was, in effect, neutral in terms of

accuracy. This form of shift predominated with $M = 3.91$ (38.8%), $SD = 3.00$ of inaccurate responses changing to a further novel inaccurate response. Undesirable shifts (from accuracy to inaccuracy) $M = 3.41$ (33.9%), $SD = 1.78$ were more frequent than desirable shifts (from inaccuracy to accuracy) $M = 2.74$ (27.3%), $SD = 1.67$.

Patterns of changes

56.3% of responses remained the same throughout the experiment (repetition pattern XXX). The most common pattern of change was to make a change and then sustain this new response for the next repetition (XYY), this form of repetition pattern accounted for 16.7% of all response patterns, 9.4% returned to the original (XYX), 8.9% sustained the original response and then changed to a novel response (XXY), and 8.7% changed with each repetition (XYZ).

Children's comments

Three children (less that 2% of the participants), all aged 4-5, commented on the use of repetition: "That's like the other one", "I heard that one before" and "It's again isn't it, like the other time". Their comments related to recognition of repetition rather than questioning its usage. The interviewer did not respond to these comments and the children continued to answer the remaining questions without changing their responses to the questions about which they commented.

Discussion

This study examined whether accuracy and consistency of children's responses were affected by the question type, repetition style and repetition order of repeated questions. As predicted, the younger children were less accurate and less consistent in their responses to repetition than older children although the

difference in consistency was less marked than in previous experimental research (Howie et al., 2004; Krähenbühl & Blades, 2006; Poole & White, 1991). Consistent with previous research (Memon & Vartoukian, 1996; Moston, 1987), accuracy declined after the initial response.

The effect of repetition on accuracy of responses could be generalised, as there were no interactions with age. Awareness of these question types by the children is suggested by the different patterns of changes made to responses depending on the question type.

Accuracy improved slightly in responses to answerable questions if the questions were asked in a structurally different way, rather than merely with the same or different wording. This improvement, albeit slight, was found mostly in responses to *closed to open* repetitions, which was consistent with previous research which showed the benefits of the use of open-ended questions (Dent & Stephenson, 1979; Hershkowitz, 2001; Lamb et al., 1996; Orbach & Lamb, 2000; Peterson et al., 1999).

The lack of general improvement in accuracy as a result of repetition suggested that a gist memory trace (Reyna & Kiernan, 1994), activated by the initial posing of a question, was unsuccessful in enabling access to the correct response or that other factors inhibited the verbalisation of that correct response. Repetition of unanswerable questions in particular did result in improvements in accuracy as the reluctance to state "I don't know" predominated (Hughes & Grieve, 1980; Waterman et al., 2004), with further incorrect responses as a result.

The limited options of responses to *open to closed* unanswerable questions ("yes", "no" and "I don't know") did not encourage children to

acknowledge that they did not know the answer, which would have been accurate. This reluctance to admit ignorance, combined with a reduction in accuracy after the initial response was given, has implications for interviewing protocols which encourage a move from open to closed questioning styles (Home Office, 2001). Our results indicated that if an interviewer was unaware of whether a question was answerable or unanswerable there was little chance of improvement in accuracy through the use of repetition.

As expected in Study 2, over a quarter of responses to repetitions changed (Krähenbühl & Blades, 2006; Moston, 1987; Poole & White, 1991). Frequency of change in Study 2, an experimental study, was lower than in Study 1 in which nearly half of the responses to repetitions changed. This suggested that consistency related results from experimental studies may have underestimated the number of changes in responses that children made in actual interviews.

Contrary to our expectations responses to *verbatim* repetitions (see Figures 3:1 and 3:2) were most consistent across repetitions. In other words, when the children encountered this style of repetition they were more likely to maintain their original response (irrespective of accuracy) than in responses to other styles of repetition. Overall, once children had made a change to their original response they maintained that change rather than to return to their original response, even though their original response was their most accurate.

The results of Study 2 showed that consistency in children's responses to repetition may have been overestimated in the majority of interview repetition experiments in which only, or mainly, *verbatim* repetition was used (for example Howie et al., 2004; Memon & Vartoukian, 1996; Poole & White, 1991; Powell

& Thomson, 1996). Our results also called into question the caution concerning the use of *verbatim* repetitions in the ABE interviewing protocol (Home Office, 2001) because we found that changes were more frequent in responses to *gist, open to closed* and *closed to open* repetition styles. Changes in responses could be advantageous if the accuracy of those responses increased, but we found that overall accuracy declined after the initial question.

Our accuracy results were consistent with those of Howie et al's (2004) 'no rationale' condition (where children were not informed as to the reason for the repetition of the questions in the interview). However, the number of desirable shifts (towards accuracy) according to age in Howie et al's experiment was not replicated in our results because there was no difference in the number of desirable shifts according to age. This discrepancy between our results and Howie et al's could be accounted for by the implementation of a different questioning regime. Howie et al. repeated their questions once only and did not include 'unanswerable' questions, which would have required a "don't know" response for accuracy. Nevertheless, the quantity of novel inaccurate responses showed a willingness on the part of children to try a new response if a question was repeated, even if they had no idea as to the accurate response.

In summary, the original response was most accurate, and *gist* or *verbatim* repetitions were less likely to affect this original response although when change did occur it was more common in responses to repeated unanswerable questions.

In Study 2 we used 2 repetitions of each question, however, in Study 1 we found that the mean number of repetitions in police interviews was between 3 and 4 repetitions of an initial question. Previous researchers have not considered

the effect of so many multiple repetitions within an interview. So in Study 3 we introduced multiple question repetitions to examine the effect of this factor on the accuracy and consistency of children's responses.

Chapter 4

Study 3

The results from Study 2 showed that responses to *gist* repetition were more accurate and more consistent (with the exception of verbatim repetitions) than responses to other repetition forms. In Study 2, questions were repeated twice, but in Study 1 we found that questions were repeated more frequently. Therefore in Study 3 we included multiple gist repetitions of questions to examine the effect of an increased number of repetitions on children's responses.

Introduction

As discussed in Chapter 1 interviewing techniques such as question repetition, comments, or paraphrasing used in police forensic interviews may distort, influence or change the testimony elicited (Cederborg et al., 2000; Davies et al., 2000; Hershkowitz, 2001; Zajac, Gross, & Hayne, 2003). Such changes in responses may adversely reflect on the credibility of the witness (Davis et al., 1999; Regan & Baker, 1998) and affect the likelihood of a case reaching court (Gallagher & Pease, 2000).

Open-ended questions have been shown to elicit the longest, most detailed and accurate responses (Cronch et al., 2006; Hershkowitz, 2001, 2002; Sternberg et al., 1996). This question form was also found to be least likely to elicit changes in testimony (Lamb & Fauchier, 2001) although as a study of actual forensic transcripts, the researchers were unable to manipulate the balance of the question formats included.

The majority of existing experimental studies have been conducted using a limited number (one or two) of verbatim repetitions of questions (for example

Howie et al., 2004; Memon & Vartoukian, 1996; Moston, 1987, 1990; Poole & White, 1991; Powell & Thomson, 1996; Warren et al., 1991). These studies have not included different intervals between repetitions as a variable but have either repeated the questions in the same order (Howie et al., 2004; Poole & White, 1991; Powell & Thomson, 1996; Warren et al., 1991), or have interspersed the repetitions between other questions without examination of the effect of changes in that dispersal pattern (Memon & Vartoukian, 1996; Moston, 1987). In forensic interviews it is unlikely that all questions can be answered correctly, however, in research cited above only answerable questions were used. Interviewing protocols (Home Office, 1992, 2001) state that interviewees should be told that it is appropriate and acceptable to state "I don't know" in response to a question (if this is the case). Nevertheless, the results from Chapter 3 and from previous researchers have found that children would rather generate an inaccurate response to unanswerable questions, even if the question was bizarre or nonsensical, rather than say that they do not know the correct response (Beuscher & Roebers, 2005; Hughes & Grieve, 1980; Krähenbühl & Blades, 2006; Waterman et al., 2000, 2001, 2004).

Previous researchers have found that repeated questions affect children's testimony. However, the number of repetitions and the intervals between repetitions used by previous researchers has not reflected either the form of questions or the frequency of repetitions in police interviews as shown in Study 1. In Study 3 we therefore increased the number of instances that a question was repeated during an interview and controlled the intervals between those repetitions. Due to the number of levels in the independent variables (number of repetitions, intervals between repetitions) the experimental design of Study 3 did

not allow the possibility of an interaction between these two variables, although this was introduced in Studies 4 and 5 by means of a reduction in levels.

Study 3 included gist repetition of answerable and unanswerable questions in interviews with young children between four and nine years of age who had witnessed a staged event. We incorporated the number of repetitions and length of delays between repetitions used in police interviews in Study 1. The repeated questions were varied systematically to find out whether frequency of repetitions or length of delay between repetitions had an effect on the accuracy and consistency of responses.

We expected that accuracy and consistency in responses to repeated questions would increase with age and that accuracy in all age groups would diminish after the initial response (Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Moston, 1987; Poole & White, 1991). Following Howie et al. (2004) we predicted that all children would provide more undesirable shifts than desirable shifts and that younger children would make more undesirable shifts than older children. In the absence of previous research into multiple repetitions on this scale and the effect of intervals between repetitions, we did not make any specific predictions about the effects of these variables.

Method

Participants

Of the 156 children who completed the study there were fifty-seven 8-9-year-olds (32 boys and 25 girls, M = 8 years 5 months, SD = 3.53 months), fifty-one 6-7-year-olds (27 boys and 24 girls, M = 6 years 7 months, SD = 4.35 months) and forty-eight 4-5-year-olds (27 boys and 21 girls, M = 4 years 8 months, SD = 2.92 months). A further five children took part but their data were

not included because four of these withdrew (the children asked to return to their classes) and one child's interview was interrupted.

Procedure

The event, in which a female confederate called Chris talked to a class of children about dog care and her dog called Susie, was the same as in Study 2 (pp. 81-84).

The children were interviewed individually by an unfamiliar adult approximately one week (7, 8 or 9 days later). The teachers informed the children that the interviewer wanted to talk to them about Chris's visit.

The interview commenced with instructions and explanations for the child as described in Study 2 (p. 84).

The children were asked 60 questions in total. There were eight open-ended base questions, half of which were answerable and half unanswerable (from the context of the presentation). Within each type (answerable or unanswerable), half of the questions were about the event and half were about the story. Each of these eight base questions was repeated in gist form a further four times (based on the number of times a single question was repeated in the police interviews in Study 1). Each repetition was separated by the following intervals: zero – no other questions in between, three questions in between, six questions in between, or ten or more questions in between to represent the immediate, short, medium and long term intervals that were possible within the 60 questions posed. For example the question "Where does Chris take Susie for her walks?" was repeated as "Chris takes Susie out for walks, where does she take her?", "When Chris takes Susie out for walks, where do they go?", "Susie is

taken for walks by Chris, where is she taken?" and "Where is Susie taken for her walks with Chris?"

There were 20 filler questions each of which was in a closed question format requiring a "yes", "no" or "don't know" response and which were asked only once. The filler questions were used to create the required intervals between repetitions and were therefore not included in the analyses of repeated questions. To maintain consistency with the repeated questions half of the filler questions were answerable and half were unanswerable, half of the filler questions related to the event and half to the story.

The questions and the question intervals were all counterbalanced. There were 24 different patterns of repetition intervals (base, zero, three, six, ten plus then base, zero, three, ten plus, six etc). Each pattern was used for each answerable and unanswerable question giving 48 different question orders. At least one child in each age group received each of these question orders.

Example of a set of questions used in the interview

'A' refers to answerable questions, 'U' to unanswerable questions. 'F' refers to filler questions. 'R' followed by a number refers to a series of repeated questions: 0 is the original question, 1, 2, 3 and 4 are the repeats of that question.

1. Was there a main road next to the woods? [UF1]
2. Where did Chris keep her tissues? [AR1, 0]
3. Where were Chris's tissues kept? [AR1, 1]
4. Did the chickens in the story get eaten? [AF2]
5. In the story how many chickens were there? [UR2, 0]
6. If you counted the chickens how many would you get? [UR2, 1]
7. In the story did the fox have a sore patch on its leg? [UF3]

8. In the story which animal wanted to eat the ducks? [AR3, 0]

9. What creature wanted to eat the ducks? [AR3, 1]

10. Chris's tissues, where were they kept? [AR1, 2]

11. Were the fleas as big as the dog? [AF4]

12. Did the ducks go swimming in the pond? [AF5]

13. There were chickens but how many were there? [UR2, 2]

14. Where does Chris take Susie for her walks? [UR4, 0]

15. Chris takes Susie out for walks, where does she take her? [UR4, 1]

16. Who wanted to eat those ducks? [AR3, 2]

17. Did Chris wear a hat? [AF6]

18. What colour was Chris's scarf? [AR5, 0]

19. How would you describe Chris's scarf, its colour for example? [AR5, 1]

20. Were there lambs on the farm in the story? [UF7]

21. Chris had a packet of tissues, where did she keep them? [AR1, 3]

22. When Chris takes Susie out for walks, where do they go? [UR4, 2]

23. Was Chris wearing stripy socks under her trousers? [UF8]

24. Tell me the number of chickens there were, what was it? [UR2, 3]

25. Tell me about Chris's tissues, where were they kept? [AR1, 4]

26. The colour of Chris's scarf, what was it? [AR5, 2]

27. The ducks were nearly eaten, who wanted to do that? [AR3, 3]

28. How many chickens did you see? [UR2, 4]

29. Was there a tin of dog food? [AF9]

30. In the story how old was the fox? [UR6, 0]

31. What was the fox's age? [UR6, 1]

32. In the story why did the dog keep on scratching? [AR7, 0]

33. Something was making the dog itchy so that it kept on scratching, what was that? [AR7, 1]

34. An animal wanted to eat the ducks, which animal was that? [AR3, 4]

35. Susie is taken for walks by Chris, where is she taken? [UR4, 3]

36. What was the name of the man who came into the classroom? [UR8, 0]

37. The man's name, what was it? [UR8, 1]

38. Tell me about the fox, how old was he? [UR6, 2]

39. Where is Susie taken for her walks with Chris? [UR4, 4]

40. What was making the dog scratch all the time? [AR7, 2]

41. Did the dog carry the chickens back in a sack? [AF10]

42. Tell me about Chris's scarf, what was it like? [AR5, 3]

43. Had the dog biscuits in the bag gone soft? [UF11]

44. What was the man called? [UR8, 2]

45. Did Chris bring Susie's basket into school to show you? [AF12]

46. The scarf that Chris was wearing, what colour was it? [AR5, 4]

47. Did Susie have only three legs? [AF13]

48. Was it the dog's birthday? [UF14]

49. Did Chris have any other dogs at home? [UF15]

50. Was the fox's tail all tangled? [UF16]

51. Tell me about the poor dog, why did it keep on scratching? [AR7, 3]

52. Was the dog lead purple? [AF17]

53. How old was the fox in the story? [UR6, 3]

54. Did Chris have a pen in her handbag? [UF18]

55. The dog kept on scratching, why was that? [AR7, 4]

56. The man who came in, what was his name? [UR8, 3]

57. Did Susie have clean teeth? [UF19]

58. The fox was how old? [UR6, 4]

59. Did the fox live in the woods? [AF20]

60. What was the man's name? [UR8, 4].

The children's responses were coded and assessed for accuracy and for consistency of responses to question types (as described in Study 2, pp. 88-89), the effect of the interval length between repetitions, and the effect of the number of repetitions. The effect of a change in response on accuracy was also examined. Any comments made by the children relating to repetition during the interview were recorded.

An independent rater coded 10% of the results; these codings were then subjected to a kappa test. The results for accuracy and change were $\kappa = 0.99$ and $\kappa = 0.95$ respectively (both significant $p < .001$).

Results

Preliminary analyses of all responses

Analyses between female and male participants and between different lengths of delay (7, 8 or 9 days) in each age group for all responses revealed no differences. Therefore these factors were not considered further.

Accuracy of all responses

A one-way between-subjects ANOVA was performed on the accuracy of responses to all questions. There was an effect of age group: $F(2, 153) = 28.93$, $p < .001$. Accuracy increased with age: for 4-5-year-olds $M = 22.25$ (37.8%), $SD = 7.86$, 6-7-year-olds $M = 28.92$ (48.2%), $SD = 7.18$ and 8-9-year-olds $M = 32.67$ (54.5%), $SD = 6.11$. A Tukey HSD test showed a difference between the 4-5-

year-olds and the 6-7 and 8-9-year-olds, $p < .001$, and between the 6-7-year-olds and the 8-9-year-olds, $p < .05$.

Effect of repetition frequency on accuracy of responses to repeated questions

We considered whether the number of times that a question was asked influenced the accuracy of a child's responses. A 3 Age group (4-5, 6-7, 8-9-year-olds) x 2 Question type (answerable, unanswerable) x 5 Repetition order (base, first, second, third and fourth repetitions) analysis of variance (ANOVA) with repeated measures was conducted on the children's responses to repeated questions.

There were main effects for age group: $F(2, 153) = 18.01$, $p < .001$, for question type: $F(1, 153) = 20.09$, $p < .001$, and for repetition order, $F(4, 150) = 9.82$, $p < .001$. There were interactions between age and question type: $F(2, 153) = 4.04$, $p < .05$, and between question type and repetition order, $F(4, 150) = 23.44$, $p < .001$.

The effect of age showed that accuracy increased with age: 4-5-year-olds $M = 1.60$ (40.1%), $SD = 6.53$, 6-7-year-olds $M = 2.03$ (50.8%), $SD = 6.25$ and 8-9-year-olds $M = 2.31$ (57.8%), $SD = 5.33$. A Tukey HSD comparison showed a difference between the 4-5-year-olds and the 6-7-year-olds and the 8-9-year-olds, $p < .01$ and $p < .001$ respectively, and between the 6-7-year-olds and the 8-9-year-olds, $p < .05$.

The effect of order showed that there was a decline in accuracy of responses until the fourth repetition: $M = 2.15$ (54.2%), $SD = 1.37$, $M = 2.03$ (51.1%), $SD = 1.47$, $M = 1.94$ (49.0%), $SD = 1.54$, $M = 1.90$ (47.9%), $SD = 1.44$, $M = 1.90$ (47.9%), $SD = 1.55$ for base, first, second, third and fourth repetitions respectively. Pairwise comparisons showed a difference between the base

question and all other repetitions ($p < .001$), between the first repetition and second repetition ($p < .05$) and between the first repetition and all other repetitions ($p < .001$) but no other differences between the later repetitions.

The effect of question type showed the accuracy of responses to answerable questions $M = 2.21$ (55.9%), $SD = 4.15$ and responses to unanswerable questions $M = 1.76$ (44.2%), $SD = 5.00$. These results serve to demonstrate a level of accuracy for each question type group. It was not possible to counterbalance the answerable and unanswerable questions for difficulty. Therefore, although Pairwise comparisons showed a difference between these question types ($p < .001$) this was irrelevant.

The interaction between age group and question type showed that although accuracy for both answerable and unanswerable questions increased with age, the difference in accuracy between question types also increased with age: responses to unanswerable questions were lower than responses to answerable questions by 4.4%, 8.4% and 20.8% for 4-5, 6-7, and 8-9-year-olds respectively. Tests of simple effects were conducted on responses to answerable and unanswerable questions separately. These showed that the age differences were a consequence of responses to answerable questions only. For such questions the differences between all age groups were: 8-9-year olds and the other age groups $p < .001$, and between 4-5-year-olds and 6-7-year-olds $p < .01$.

Figure 4.1
Accuracy of responses according to question type and repetition order

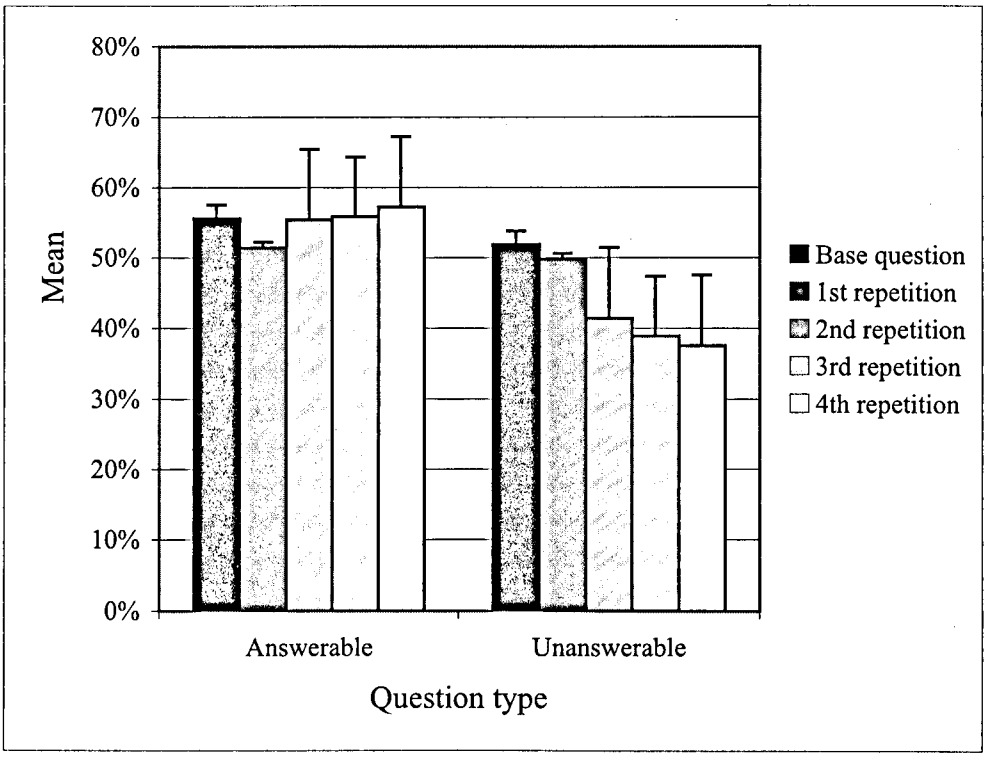


Figure 4.1 shows the interaction between question type and repetition order. For answerable questions, accurate responses declined with the first repetition but returned to and subsequently exceeded the original level with further repetitions. For responses to unanswerable questions accuracy declined with each repetition. Tests of simple effects were conducted on responses to answerable and unanswerable questions separately. The results showed a difference depending on question type. There was a difference in accuracy of responses to answerable questions between the first and fourth repetitions ($p <$ .05). For responses to unanswerable questions differences were found between the accuracy of the initial response (base) and responses to all repetitions ($p <$ .001) with the exception of the first repetition. The second, third and fourth repetitions also showed a difference to the first repetition ($p < .05$ with the second repetition, $p < .001$ with the third and fourth repetitions).

Effect of delay on accuracy of responses to repeated questions

We also considered the effect of delay between repetitions on children's responses through a comparison of the accuracy of responses for all the repeated questions that were asked after an interval of zero questions, after three intervening questions, six intervening questions or after an interval with ten or more filler questions. We used a 3 Age group (4-5, 6-7, 8-9-year-olds) x 2 Question type (answerable, unanswerable) x 5 Interval (base, zero, three, six, ten plus) analysis of variance (ANOVA) with repeated measures to examine the children's responses to repeated questions.

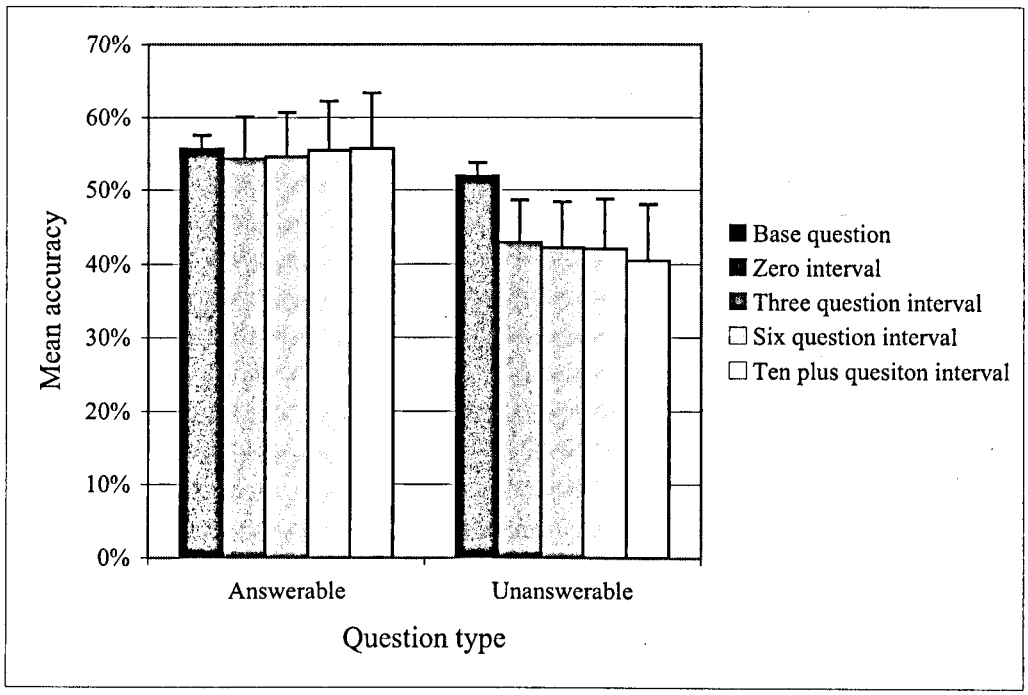
The results for effects of age group and question type were as for the previous analysis on the effect of repetition frequency. There was also an effect for interval: F (4, 150) = 8.5, $p < .001$, and an interaction between question type and interval: $F(4, 150) = 7.13$, $p < .001$.

Overall repetition interval showed little effect on accuracy once repetition had taken place (base $M = 2.15$ (54.2%), $SD = 1.37$, zero $M = 1.95$ (49.2%), $SD = 1.51$, three question interval $M = 1.94$ (48.9%), $SD = 1.55$, six question interval $M = 1.95$ (49.3%), $SD = 1.49$, and ten plus question interval $M = 1.92$ (48.6%), $SD = 1.47$. Pairwise comparisons showed differences between the accuracy of the responses to the base question and all other repetitions $p < .001$ only, but no differences between the repetition intervals.

The interaction between question type and repetition interval showed that although there were no differences in accuracy with different intervals between repetitions for answerable questions, there was a decline in accuracy for responses to unanswerable questions after the base question (see Figure 4.2).

Tests of simple effects showed a difference between the accuracy of responses to unanswerable base questions with other response intervals $p < .01$ only.

Figure 4.2
Accuracy of responses according to question type and repetition interval



Change to previous response

We found that there were no changes at all (i.e. responses remained completely consistent across all repetitions) in responses to 50.8% of questions. Of these consistent responses 74.6% were accurate, and 25.4% were inaccurate.

Responses to repeated questions resulted in changes (in comparison to the previous response given by the child) in 28.4% of all responses to repetitions, 28.3% of responses to answerable questions and 28.5% of all responses to unanswerable questions.

To find out if the frequency of repeated questions affected children's changed responses, a 3 Age group (4-5, 6-7, 8-9-year-olds) x 2 Question type (answerable, unanswerable) x 4 Repetition order (first, second, third and fourth repetitions) analysis of variance (ANOVA) with repeated measures was applied to the changes in children's responses.

There were main effects for age group: $F(2, 153) = 24.66$, $p < .001$, and for repetition order: $F(3, 151) = 3.75$, $p < .05$. There was no effect of question type and no interactions.

The effect of age showed that the number of changes decreased with increasing age: 4-5-year-olds $M = 1.62$, (40.6%), $SD = 0.83$, 6-7-year-olds $M = 1.08$, (26.8%), $SD = 0.68$, and 8-9-year-olds $M = 0.71$, (17.70%), $SD = 0.56$. A Tukey HSD comparison showed a difference between the 4-5-year-olds and the 6-7 and 8-9-year-olds, $p < .001$, and between the 6-7-year-olds and the 8-9-year-olds, $p < .05$.

The effect of repetition order showed that changes decreased as the number of repetitions increased: changes between base to first repetition $M = 2.47$, (30.8%), $SD = 1.75$, first to second repetition $M = 2.25$, (28.1%), $SD = 1.75$, second to third repetition $M = 2.10$, (26.3%), $SD = 1.77$, and third to fourth repetition $M = 2.06$, (25.7%), $SD = 1.74$. Tests of simple effects showed differences in changes in responses between base to first repetitions and both second to third and third to fourth repetition orders only, ($p < .01$).

To find out if delay had an affect on children's likelihood of changing responses, a 3 Age group (4-5, 6-7, 8-9-year-olds) x 2 Question type (answerable, unanswerable) x 4 Delay (intervals of zero, three, six, ten plus) analysis of variance (ANOVA) with repeated measures was applied to the changes in children's responses.

The main effect for age group was as reported above. There was also an effect for delay, $F(3, 151) = 5.83$, $p < .001$. There was no effect of question type and no interactions.

The effect of delay showed that change decreased as the interval between repetitions increased: base to zero interval $M = 2.47$, (30.9%), $SD = 1.72$, zero to three interval $M = 2.11$, (26.4%), $SD = 1.80$ and three to six interval $M = 2.01$, (25.2%), $SD = 1.63$, with the exception of the longest interval of six to ten plus $M = 2.28$, (28.5%), $SD = 1.82$, tests of simple effects showed a difference in changes between the base to zero and zero to three intervals ($p < .05$), and between the base to zero and the three to six intervals ($p < .001$) only.

Desirable and Undesirable shifts

The above results do not demonstrate movements towards (desirable shifts) or away (undesirable shifts) from accuracy in relation to the previous response given by the child. An inaccurate response may change each time the question is posed but remain inaccurate (novel inaccurate). Repetition may have a detrimental effect on the accurate responses or indeed a positive effect on the inaccurate. Across age groups 1.4% of changed responses were desirable shifts, 16.2% were undesirable shifts, and 82.4% were novel inaccurate shifts.

Children's comments

Twenty-one children (13.5%) made a total of 27 comments on the repetition of questions (4-5-year-olds = 4, 6-7-year-olds = 7, 8-9-year-olds = 10). Twenty-five of these comments recognised that repetition had taken place, by saying for example, "It's the same question this is"; all of these children continued to answer the questions. Two children (1.2% of the children interviewed) questioned the repetition procedure and asked why the questions were being repeated: "Why are you talking to me about the same stuff again and again and again?" (5-year-old), "Why are they where tissues, where tissues, where tissues?" (7-year-old). Both children were told, "Please answer the

question", both continued without changing their responses, and completed the interview without further comment.

Discussion

Study 3 examined the accuracy and consistency of children's responses to questions that were repeated on multiple occasions with different intervals between repetitions. As predicted our results showed that accuracy and consistency performance increased with age (Howie et al., 2004; Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Moston, 1987) and that consistency was greater than in responses to repetitions in police interviews discussed in Study 1. The first response given by all children was the most accurate (Memon & Vartoukian, 1996; Moston, 1997) and accuracy declined with the first and second repetitions (as in Study 2), but then remained stable.

The reluctance to answer unanswerable questions accurately was found in all age groups, with differences in accuracy scores being attributed to the increasing ability with age to answer the answerable questions correctly. Figure 4.1 showed how the accuracy of responses to unanswerable questions declined with the first and second repetitions but then remained stable. Similarly, the decline in accuracy for responses to unanswerable questions also took place with the initial repetition; the interval between repetitions did not have an effect on accuracy (see Figure 4.2). These results suggested that children were not willing to say that they "don't know" the answer more than once or twice; to say "I don't know" once may be difficult enough but it may be even harder for children to maintain such a response when asked multiple repeated questions.

As in Study 2 we found that a quarter of responses to repetitions were inconsistent with the previous response given. The decrease in changes with an

increase in age was in accordance with previous experimental research (Ghetti et al., 2002). The frequency of these changes exceeded those found in responses to open-ended questions in forensic transcripts (Lamb & Fauchier, 2001; Orbach & Lamb, 2001) but were lower than changes in responses to all forms of repetition found in Study 1. In accordance with results from Study 2 when changes in responses were made these rarely made an improvement in accuracy but almost always resulted in a decline in accuracy (Howie et al., 2004) or else simply provided a further inaccurate response.

In most interviews this level of inconsistency might be perceived by legal representatives and possibly jury members to have a negative impact on the perception of a child's credibility (Davis et al., 1999; Regan & Baker, 1998) and may present a reason for not continuing an investigation (Gallagher & Pease, 2000).

Our results for consistency showed no effects of question type (answerable or unanswerable), nor any interaction with age group, repetition interval, or repetition delay on changes to responses through repetition. This suggested that it was the act of repetition itself rather than the repeated question type, the way in which the repetition was conducted, or the age of the recipient that affected the consistency of responses to repetitions.

Although some of the children articulated their awareness of the question repetition, they rarely challenged or queried its use. Children have been shown to be cautious in 'correcting' adult distortions of their responses when responding to perceived interviewer expectations (Moston, 1990; Roberts & Lamb, 1999; Sternberg et al., 1996; Westcott & Kynan, 2006). However, we found that the children who did ask why repetition was being used did not subsequently change

their responses. In the rapport stage of police interviews (Home Office, 2001) instructions and ground rules are discussed with the child. Children are encouraged to say if they do not understand a question or if they believe the interviewer has misunderstood their response. During this stage of the interview, it may also be advisable for the child to be told or given practice at challenging interviewer behaviour (both generally and specifically in respect to the use of question repetition) as the responses to the repetitions in the present study appeared to be in relation to the act of repetition rather than to the content of that repeated question.

A change in response may not necessarily affect the accuracy of that response (changes may represent numerous inaccurate responses). Nevertheless it is important to establish conditions where changes in response accuracy take place in order to develop appropriate interviewing techniques. Accuracy in responses to unanswerable questions declined with repetition (see Figures 4.1 and 4.2), giving rise to undesirable changes. This pattern was linked to the children's reluctance to say that they "don't know" as a response to a question. To say "I don't know" once may be difficult enough, but it may be even harder for children to maintain such a response when asked multiple repeated questions.

When repetition is deemed necessary, *verbatim* form is cautioned against in interviewing protocols and it is suggested that repetition should take a different form (Home Office, 2001). Our results confirmed that *gist* repetition did not have an adverse effect on responses to answerable questions but there remains a problem with unanswerable questions with the main differences in both accuracy and change being found in the first two repetitions. Therefore Study 4 used two *gist* repetitions of answerable and unanswerable questions at

either a short or long interval to establish whether a there was an interaction between repetition delay and repetition order on the accuracy and consistency of children's responses to questions.

Chapter 5

Study 4

In Studies 2 and 3 we examined the effects of repetition patterns and quantity on children's accuracy and consistency. Those studies included, in replication of Poole and White (1991), a 1-week delay between event and interview. In Study 4 we incorporated two interview conditions: interview within 48-hours, and interview after 1-week to establish whether accuracy or consistency was affected by the length of delay before an interview. We also reduced the range of levels of repetitions and intervals between repetitions in order to examine whether these factors interacted.

Introduction

The prevalence of abuse of children is unknown. Children who have been abused may not disclose their experiences for diverse reasons (Goodman-Brown et al., 2003), may only disclose weeks, months or years after the abuse has happened (Sjöberg & Lindblad, 2002; Smith et al., 2000), or may wait until adulthood (Berliner & Conte, 1995; London et al., 2005; Smith et al., 2000). Regardless of the delay between the event taking place and disclosure the ensuing judicial procedure is similar.

The ABE (Home Office, 2001) interviewing protocol outlined in Study 1 (pp. 52-54) advocated interviewing as soon as possible after an allegation was made. In practice in England and Wales the time between an event and a court case may take months despite recommendations such as the Pigot Report, which stated that proceedings involving child witnesses should advance "as rapidly as is consonant with the interests of justice" (Home Office, 1989, Para. 2.14). Whilst minimising delay is encouraged, trials involving child witnesses may take

place several months or more after the alleged incidents (Flin et al., 1992b) and there is an average delay of 11.6 months for cases to come to the Crown Court (Plotnikoff & Woolfson, 2004).

Delay between an event and the disclosure or responses to questions in a forensic interview can have an impact on the recall of the memory for that event. A deterioration in accuracy and consistency in recall of an event occurs over time between repeated interviews after both laboratory based events (Gobbo et al., 2002; Poole & White, 1991, 1993; Roebers & Schneider, 2000), and after real-life experienced medical events (Salmon & Pipe, 2000; Tizzard-Drover & Peterson, 2004).

Changes in children's event report over time may reflect changes either in memory storage, information retrieval or both of these processes. In fuzzy-trace theory (described in Chapter 1 pp. 37-39), Reyna and Brainerd (1995) proposed that when event memories were stored they were represented as verbatim (the surface content) and gist (interpretative content) traces. Initially verbatim representations would be cued although, with time, the basis of recall would shift to gist representations.

Researchers whose experiments included different delay intervals before an initial interview found a reduction in accuracy with an increase in time (Jones & Pipe, 2002; Poole & White, 1991; Powell & Thomson, 1996; Salmon & Pipe, 2000). However, in respect to consistency Poole and White (1991) found that consistency was greater in a delayed initial interview (a delay of 1-week) than in an immediate initial interview (in effect no delay). An immediate interview after an event is a rare occurrence in interviews after incidences of abuse (Goodman-Brown et al., 2003). Goodman-Brown et al. (2003) found four tenths of the

children who alleged abuse disclosed within 48-hours following the last assault, and nearly one fifth of children disclosed between 48-hours and two-weeks. Therefore this study examined accuracy and consistency of responses to repeated questions using these timescales.

In Study 1 we found that repetition may legitimately be used to gather more detailed information or to clarify a point. The results from Studies 2 and 3 indicated that repetition interval, repetition order, and question type affected the accuracy and consistency of responses after repeated questions. The results also showed that the changes in responses made by children were generally detrimental to accuracy, and that children were reluctant to acknowledge when they did not know the answer to an unanswerable question (Beuscher & Roebers, 2005; Krähenbühl & Blades, 2006; Price & Connolly, 2004; Waterman et al., 2000, 2001, 2004). In Studies 2 and 3 we did not manipulate the timing of the interview to find out if the delay before the interview had any effect on children's responses to repeated questions.

In Study 4 we included two interview delay conditions: either after a period of less than 48-hrs, the most common length of delay before disclosure (Goodman-Brown et al., 2003), or after a period of 1-week (Krähenbühl & Blades, 2006; Poole & White, 1991) as used in Studies 2 and 3.

In Study 3 we found that the main effect of repetition order took place within the first two repetitions. In Study 1 we referred to interviewing protocols (Home Office, 1992, 2001) that suggested a delay in the repetition of questions was advantageous in preventing changes in responses as a result of compliance to perceived interviewer demands. In Study 4 therefore, we included two repetitions and used a short and long delay between repetitions to establish

whether the variables of repetition order and repetition delay would interact in their effect on children's responses.

We predicted that accuracy of responses would deteriorate with a longer delay compared to a shorter delay before the initial interview (Poole & White, 1991; Powell & Thomson, 1996; Salmon & Pipe, 2000). In accordance with Poole and White we expected consistency to be greater in the delayed interview than in the immediate interview. Regardless of delay, as found in Studies 2 and 3, and in accordance with previous research we expected accuracy and consistency in responses to questions would increase with age. In accordance with the results of Studies 2 and 3 it was expected that the majority of shifts in accuracy (Howie et al., 2004) would be undesirable shifts but that the majority of changes would be novel inaccurate changes. We expected, in accordance with Studies 2 and 3 that the most common pattern of shifts would be XYY when a child's response changed after the initial question and then remained with that changed response through subsequent repetitions.

Method

Participants

In this study there were 305 participants: one hundred and two 5-year-olds (48 boys and 54 girls, mean age = 5 years 4 months, *SD* = 3.8 months), ninety-six 7-year-olds (42 boys and 54 girls, mean age = 7 years 4 months, *SD* = 3.7 months), and one hundred and seven 9-year-olds (51 boys and 56 girls, mean age = 9 years 2 months, *SD* = 3.4 months). No children withdrew from the experiment.

Procedure

Children observed the staged scripted event on dog care described in Study 2 (see pp. 81-84). The presentation was audio taped to detect any differences that may affect the children's responses to the subsequent interview questions.

An unfamiliar adult interviewed the children individually either within 48-hours (condition 1) or 1-week (condition 2) later.

Materials and coding

The children were interviewed under two conditions: either within 48-hours (145 children) or after 1-week (160 children). The interviews followed the same structure regardless of the condition.

In the interview a total of 40 questions were asked: eight open-ended questions (four answerable and four unanswerable) were repeated a further two times (24 questions). An example of an answerable repeated question was "Where did Chris keep her tissues?", which was repeated as "Where were Chris's tissues kept?", and "Chris had a packet of tissues, where did she keep them?" An example of an unanswerable repeated question was "Chris takes Susie [her dog] out for walks, where does she take her?", which was repeated as "Susie is taken for walks by Chris, where is she taken?", and "Where does Chris take Susie for her walks?" Four of the repeated questions were repeated with an initial "long" interval of 17 intervening questions and then a second "short" immediate repetition, and four in the opposite pattern.

A further 16 filler non-repeated closed questions (of which half were answerable and half unanswerable) were used both to promote a less artificial range of questions than an interview consisting of entirely repeated questions and

to enable the required intervals between the repeated questions. An example of an answerable question was "Was there a tin of dog food?", and an unanswerable example was "Did Chris have any other dogs at home?"

The order of repetition delays was counterbalanced to provide eight different questioning patterns to which the eight question repetitions were applied. This gave rise to 64 different interview orders; one set of questions is given below.

The labelling of the questions is as follows: F = filler question, R = repeated question (numbered 1-8), B = original question, I or II = first or second repetition order, S or L = short or long repetition interval, A or U = answerable or unanswerable.

1. FU Were there lambs on the farm in the story?
2. FA Was there a tin of dog food?
3. R1BA What colour was Chris's scarf?
4. R2BU In the story how many chickens were there?
5. R2ISU Tell me the number of chickens there were, what was it?
6. FA Did Chris wear a hat?
7. R3BA Where did Chris keep her tissues?
8. FA Did the dog carry the chickens back in a sack?
9. R4BA In the story why did the dog keep on scratching?
10. R4ISA What was making the dog scratch all the time?
11. FU Did Susie have clean teeth?
12. R5BU Chris takes Susie out for walks, where does she take her?
13. FA Was the dog lead purple?
14. R6BU In the story how old was the fox?

15. R6ISU Tell me about the fox, how old was he?

16. R7BU What was the name of the man who came into the classroom?

17. FU Had the dog biscuits in the bag gone soft?

18. R8BA In the story which animal wanted to eat the ducks?

19. R8ISA An animal wanted to eat the ducks, which animal was that?

20. R1ILA The scarf that Chris was wearing, what colour was it?

21. R1IISA What colour would you say Chris's scarf was?

22. R2ILU How many chickens did you see?

23. FU In the story did the fox have a sore patch on its leg?

24. R3ILA Chris had a packet of tissues, where did she keep them?

25. R3IISA Where were Chris's tissues kept?

26. FU Did Chris have any other dogs at home?

27. R4IILA The dog kept on scratching, why was that?

28. FA Did Chris bring Susie's basket into school to show you?

29. R5ILU Susie is taken for walks by Chris, where is she taken?

30. R5IISU Where does Chris take Susie for her walks?

31. FA Did the fox live in the woods?

32. R6IILU What was the fox's age?

33. R7ILU What was the man called?

34. R7IISU The man's name, what was it?

35. FU Was there a main road next to the woods?

36. R8IILA Who wanted to eat those ducks?

37. FU Was it the dog's birthday?

38. FA Did the chickens in the story get eaten?

39. FA Were the fleas as big as the dog?

40. FU Did Chris have a pen in her handbag?

All responses were coded for accuracy, and the responses to repeated questions were coded for consistency (see Study 2, pp. 88-89). Patterns of changes were also recorded to establish whether, after experiencing repetitions, children maintained their original or novel responses, returned to a previous response after a change, or gave a subsequent response to each repetition. The possible response patterns were as follows: XXX, XXY, XYX, XYY, and XYZ where X was the initial response, Y the first novel response, and Z the second novel response.

10% of the transcribed results were selected at random and re-coded to establish inter-rater reliability, this was achieved at $\kappa = 0.99$ and $\kappa = 0.97$ respectively (both $p < .001$).

Results

Preliminary analyses of all responses

Analyses between male and female participants within the two interview conditions revealed no differences so these factors were not considered further.

Accuracy of all responses

A 3 Age group (4-5, 6-7 or 8-9 years) x 2 Interview condition (48-hours, 1-week) analysis of variance (ANOVA) was carried out on the accuracy of responses to both filler and repeated questions.

There was an effect of age group: $F(2,299) = 37.29$, $p < .001$, and of interview condition: $F(1,299) = 8.62$, $p < .01$.

The effect of age group showed that accuracy increased with age: $M =$ 18.23 (45.6%), $SD = 5.32$ (4-5-year-olds), $M = 21.20$ (53.0%), $SD = 4.49$ (6-7-

year-olds) and $M = 23.63$ (59.1%), $SD = 3.85$ (8-9-year-olds). A Tukey HSD comparison showed differences between all three age groups ($p < .001$).

There was an effect of interview condition showing that accuracy was greater in the 48-hours condition $M = 21.83$ (54.6%), $SD = 5.29$, than in the 1-week interview $M = 20.35$ (50.9%), $SD = 4.81$. Pairwise comparisons showed a difference between the interview conditions $p < .01$.

Accuracy of base and repeated questions only (fillers excluded)

A 3 Age group (4-5, 6-7 or 8-9 years) x 2 Interview condition (48-hours, 1-week) x 2 Question type (answerable or unanswerable) x 3 Repetition order (base question, first or second repetition) analysis of variance (ANOVA) with repeated measures was carried out on the responses to repeated questions.

There was an effect of age group: $F(2,299) = 36.01$, $p < .001$, an effect of interview condition: $F(1,299) = 8.58$, $p < .01$, and an effect of repetition order: $F(2,298) = 6.57$, $p < .01$. There was an interaction between question type and repetition order: $F(2,298) = 19.38$, $p < .001$, but no other interactions.

The effect of age group showed that accuracy increased with age: 4-5-year-olds were less accurate $M = 1.94$ (48.3%), $SD = 4.20$ than 6-7-year-olds $M = 2.28$ (56.9%), $SD = 3.56$, who were less accurate than 8-9-year-olds $M = 2.65$ (66.1%), $SD = 3.24$. A Tukey HSD comparison showed differences between all age groups $p < .001$.

There was a mean decline in accuracy from the 48-hours condition $M = 2.39$ (59.9%), $SD = 4.23$ to the 1-week condition $M = 2.19$ (54.9%), $SD = 3.86$. Pairwise comparisons showed a difference between interview conditions $p < .01$.

The effect of repetition order showed a decline in accuracy particularly after the responses to the base question from $M = 2.36$ (59.0%), $SD = 1.52$ to M

= 2.26 (56.5%), *SD* = 1.48 for responses to the first repetition, and *M* = 2.25 (56.3%), *SD* = 1.53 for responses to the second repetition. Pairwise comparisons showed a difference between the base and both the first and second repetitions of $p < .001$ and $p < .01$ respectively.

The interaction between question type and repetition order showed that the accuracy of responses to answerable questions increased with the second repetition. Accuracy of responses in the base questions was *M* = 2.24 (55.9%), *SD* = 0.90, for the first repetition *M* = 2.23 (55.7%), *SD* = 0.96, and for the second repetition *M* = 2.36 (59.1%), *SD* = 1.00. For responses to unanswerable questions accuracy declined with repetition from *M* = 2.48 (62.1%), *SD* = 1.10 for responses to the base questions to *M* = 2.29 (57.2%), *SD* = 1.13 for responses to the first repetitions and *M* = 2.14 (53.5%), *SD* = 1.14 for responses to the second repetitions. Tests of simple effects for responses to answerable questions however, showed no differences between question orders. There were differences for responses to unanswerable questions between the base and first repetitions ($p < .05$) and between base and second repetitions ($p < .001$).

Effect of repetition interval and order on responses to repeated questions only

A 3 Age (4-5, 6-7 or 8-9 years) x 2 Interview condition (48-hours, 1-week) x 2 Question type (answerable or unanswerable) x 2 Repetition order (first or second repetition) x 2 Repetition interval (short or long interval between repetitions) analysis of variance (ANOVA) with repeated measures was carried out on the responses to repeated questions.

There was an effect of age group: $F(2,299) = 36.52$, $p < .001$, and an effect of interview condition, $F(1,299) = 6.99$, $p < .01$. There were no other effects of single factors.

There were two-way interactions between repetition interval and interview condition, $F(1,299) = 5.13, p < .05$, between question type and repetition order $F(1,299) = 19.08, p < .001$, and between question type and repetition interval $F(1,299) = 20.48, p < .001$.

There were three-way interactions between repetition interval, age group and interview condition: $F(2,299) = 3.18, p < .05$, between repetition interval, question type and interview condition: $F(1,299) = 6.02, p < .05$, and between repetition interval, question type and repetition order: $F(1,299) = 20.66, p < .001$.

The effect of age group showed that accuracy increased with age: 4-5-year-olds were less accurate ($M = 0.95$ (47.3%), $SD = 2.90$) than 6-7-year-olds ($M = 1.11$ (55.5%), $SD = 2.47$) who were less accurate than 8-9-year-olds ($M = 1.32$ (65.8%), $SD = 2.20$). A Tukey HSD comparison showed differences between all age groups ($p < .001$).

The effect of interview condition showed a mean decline in accuracy from the 48-hours condition $M = 1.17$ (58.8%), $SD = 2.90$ to the 1-week condition $M = 1.08$ (54.2%), $SD = 2.69$. Pairwise comparisons showed a difference between interview conditions ($p < .01$).

The three-way interaction between age group, interview condition and repetition interval is represented in Table 5:1. The results showed that accuracy increased with age, decreased with interview delay, increased with a long interval between repetitions in the 48-hour condition, but decreased (with the exception of the 4-5-year-olds) with a long interval in the 1-week condition. Tests of simple effects were conducted on short and long intervals separately between interview conditions. There was a difference between interview conditions for the long repetition interval only ($p < .01$).

Table 5:1
The accuracy of responses in each interview condition according to age group and repetition interval

Interview condition	48-hours		1-week	
Repetition interval	short	long	short	long
4-5-year-olds	50.0%	50.5%	44.0%	45.0%
6-7-year-olds	55.1%	59.9%	56.0%	51.7%
8-9-year-olds	67.9%	68.9%	64.1%	62.9%

The three-way interaction between question type, repetition interval and interview condition indicated that accuracy was greater in the 48-hour-condition and was greater with the long repetition interval for all questions except for responses to unanswerable repetitions in the 1-week condition. Tests of simple effects were conducted on answerable and unanswerable questions separately between interview conditions and repetition intervals. These showed that there was a difference in accuracy between interview conditions for the unanswerable long delay responses only (p <.05).

There was also a three-way interaction between question type, repetition order, and repetition interval. This interaction (as represented in Table 5:2) showed that accuracy for a first repetition after a longer interval declined for answerable questions but improved if the question was unanswerable. The effect on accuracy of second repetitions was in the opposite direction; with an increase after a longer interval for answerable questions and a decrease of nearly 20% with a longer interval for unanswerable questions.

Table 5:2
The accuracy of responses to answerable and unanswerable question types according to repetition interval and order

Question order	Base	First repetition		Second repetition	
Question interval		short	long	short	long
Answerable	56.0%	59.6%	51.7%	51.4%	66.8%
Unanswerable	62.1%	50.4%	63.8%	63.2%	43.6%

Tests of simple effects were conducted on responses to answerable and unanswerable questions separately. These showed that there was a difference between the repetition interval for first and second repetitions for all question types, $p < .001$, with the exception of answerable first repetitions where $p < .05$.

Effect on consistency

Consistency examined whether there had been a semantic change to the previous response given. When there was a grammatical change only this response was not coded as being inconsistent with the previous response. In summary, responses remained completely consistent across all repetitions in responses to 61.3% of questions. Of these consistent responses 73.7% were accurate, and 26.3% were inaccurate.

Repeated questions resulted in changes (in comparison to the previous response given by the child) in 26.8% of all responses, 23.4% of all responses to answerable questions and 30.2% of all responses to unanswerable questions.

A 3 Age (4-5, 6-7 or 8-9 years) x 2 Interview condition (48-hours, 1-week) x 2 Question type (answerable or unanswerable) x 2 Repetition order (first or second) x 2 Repetition interval (short or long) analysis of variance (ANOVA) with repeated measures was carried out on the changes in responses to repeated questions.

There were effects of age group: $F(2,299) = 39.85$, $p < .001$, interview condition: $F(1,299) = 12.74$, $p < .001$, of question type: $F(1,299) = 20.28$, $p < .001$, and an effect of repetition order: $F(1,299) = 6.24$, $p < .05$. There was a two-way interaction between repetition order and repetition interval: $F(1,299) = 25.09$, $p < .001$. There were no interactions with age group or interview condition.

The effect of age showed that the number of changes decreased with age: 4-5-year-olds changed more responses $M = 0.72$ (36.3%), $SD = 3.17$ than 6-7-year-olds $M = 0.56$ (28.4%), $SD = 2.63$, who changed more responses than 8-9-year-olds $M = 0.33$ (16.4%), $SD = 2.05$. A Tukey HSD comparison showed differences between 8-9-year-olds and other age groups ($p < .001$), and between 4-5 and 6-7-year-olds ($p < .01$) .

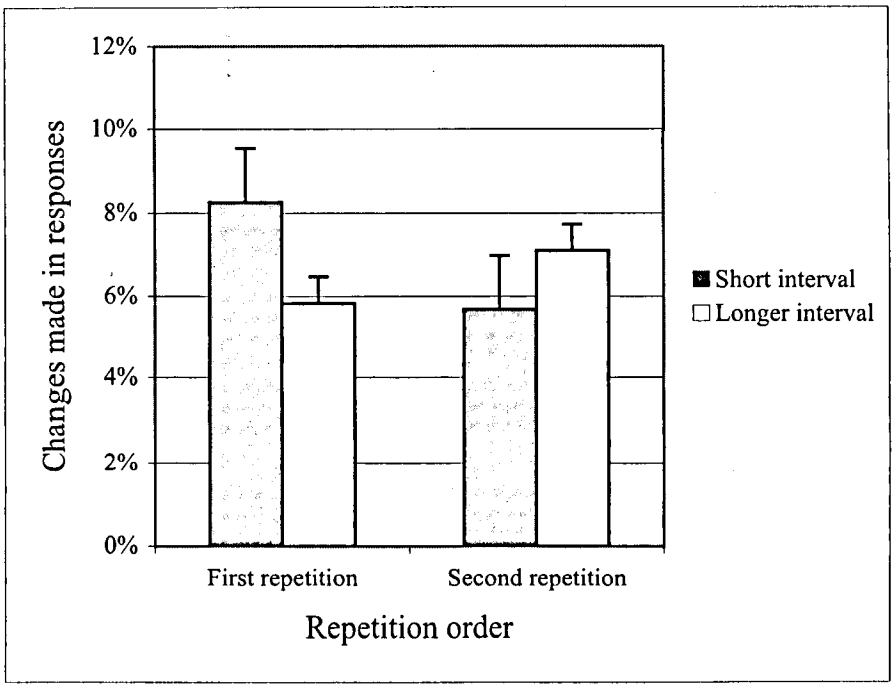
The effect of interview condition showed an increase in the number of changes with an increase in delay before the interview: changes in responses in the 48-hours condition $M = 0.47$ (23.4%), $SD = 2.70$, changes in the 1-week condition $M = 0.60$ (29.9%), $SD = 3.09$. Pairwise comparisons showed a difference in changes made according to interview condition ($p < .001$).

The effect of question type showed that fewer changes were made in responses to answerable questions $M = 0.47$ (23.4%), $SD = 1.58$ than in responses to unanswerable questions $M = 0.61$ (30.2%), $SD = 2.01$. Pairwise comparisons showed a difference between changes in responses to answerable and unanswerable questions ($p < .001$).

The effect of repetition order showed a decrease in consistency with repetition with more changes in responses to first repetitions $M = 0.56$ (28.2%), $SD = 1.70$ than in responses to second repetitions $M = 0.51$ (25.5%), $SD = 1.58$. Pairwise comparisons showed a difference between repetition orders ($p < .05$).

The two-way interaction between repetition order and repetition interval is shown in Figure 5:1. Consistency increased with a longer interval between repetitions if the question was a first repetition but decreased with a longer interval if the question was a second repetition.

Figure 5:1
The effect of repetition order and interval on consistency in responses



Tests of simple effects showed a difference between the first and second repetitions for short interval changes (p <.001), and for longer interval changes (p <.05).

Desirable and Undesirable shifts

The majority of shifts were novel inaccurate shifts M = 1.78 (42.3% of all shifts), SD = 1.76 where an inaccurate response was changed to a different, but still inaccurate response. There were more undesirable shifts M = 1.32 (31.5% of all shifts), SD = 1.18 where accurate responses became inaccurate, than desirable shifts M = 1.10 (26.2% of all shifts), SD = 1.10 where inaccurate responses became accurate.

Table 5:3 shows that the direction of shifts differs according to age. The number of shifts made (in relation to all shifts for each individual age group) to an initially accurate response to an inaccurate response decreased with age. The youngest children were more likely than older children to attempt to find further responses regardless of the inaccuracy of their initial responses.

Table 5:3
Shift direction within each age group as a percentage of shifts made

Age group	Shift direction		
	Novel inaccurate	Becomes accurate	Becomes inaccurate
4-5-year-old	47.8%	22.8%	29.4%
6-7-year-old	35.3%	28.2%	36.4%
8-9-year-old	24.1%	33.2%	42.8%

Patterns of changes

Shifts towards and away from accuracy may mask whether a child returned to a previous response. 61.3% of responses followed the XXX pattern with no change throughout. The most common pattern of changes (13.3%) was where a novel response was subsequently maintained (XYY). A change made in the final response only (XXY) accounted for 10.8% of response patterns, a return to the original response after a change (XYX) 7.7% and a novel response given to each repetition (XYZ) 6.9%.

Children's comments

Twelve children (4%) commented on the question repetition (five 4-5-year-olds, two 6-7-year-olds, and five 8-9-year-olds). Nine of the comments directly mentioned repetition e.g. "We did that question already", "It's again the same question", "It keeps repeating". Two 4-5-year-olds referred to their responses stating, "I just told you", "I already said that one, I said I'd forgotten". Only one child (in the 8-9-year-old age group) attempted to stop the repetition "You already asked me that, now would you stop it [laughs]..." Regardless of the comments made all of the children continued without any intervention on the interviewer's behalf. Of these 12 children 10 did not change their response after making their comments.

Discussion

In Study 4 we examined whether the effect of delay before interviewing would affect children's accuracy and consistency in responses to repeated answerable and unanswerable questions. The results indicated that delay and age of child were significant factors in accuracy, and in consistency of responses to repeated questions.

As predicted, and consistent with results in Studies 2 and 3, accuracy increased with age (Krähenbühl & Blades, 2006; Memon & Vartoukian, 1996; Poole & White, 1991) but declined with a delayed interview (Poole & White, 1991; Powell & Thomson, 1996; Salmon & Pipe, 2000). Contrary to expectations, a delayed interview resulted in a decrease in consistency. There were no interactions between age group and interview condition in consistency, which suggested that the effect of a delayed interview was similar for all age groups.

As with previous research (Howie et al., 2004) and in accordance with results of Studies 2 and 3, undesirable shifts outnumbered desirable shifts, although most common were shifts to a novel inaccurate response. Our results (shown in Table 5:3) suggested that higher numbers of movements towards inaccuracy and the willingness to provide alternative inaccurate responses negated the benefits of limited increases in accuracy through repetition.

An explanation for the decrease in accuracy and consistency with time is provided by fuzzy-trace theory (Brainerd & Reyna, 1995) as the memory recall after a delay became increasingly reliant on gist memory traces. These memory traces would provide a representation of the event, which became more generalised. As such the memory of the event may inadvertently give rise to

related or approximate details but not necessarily accurate responses to specific questions.

Consistency declined in the delayed interview, which given the number of novel inaccurate responses would not necessarily be recognised by examination of the accuracy scores alone. This has been an issue largely ignored by previous research despite the association between consistency and credibility (Davis et al., 1999). Unlike our results, Poole and White (1991) found an increase in consistency between the immediate and 1-week single delayed interview, which is not explained in their discussion. The discrepancy between our results and those of Poole and White could be attributed to our use of unanswerable questions, responses to which are more likely to be inconsistent than to answerable questions. Our results showed that accuracy of unanswerable questions declined with a second repetition in the 1-week condition. This suggested that children's reluctance to say "I don't know" and willingness to formulate novel (and therefore inconsistent) responses was exacerbated with time.

The interaction between question type and repetition order reflected the difficulties children have in stating when they do not know the answer to a question. These results suggested that other factors such as social awareness of conversation expectations or compliance affected the children in different ways according to age. It has been recognised that children respond to perceived interviewer demands (Schwarz & Roebers, 2006; Zajac & Hayne, 2006). In our experiment younger children were unwilling to say that they "don't know" to a greater extent than the older children who may have been more confident about stating this response.

The impact of question type on accuracy of responses was affected by repetition order and repetition intervals. In Studies 2 and 3 children have been shown to be reluctant to give a "don't know" response to unanswerable questions despite explicit encouragement to do so (Beuscher & Roebers, 2005; Krähenbühl & Blades, 2006; Waterman et al., 2004). Our results (as shown in Tables 5:1 and 5:2) indicated that a child was more likely provide an incorrect response to an unanswerable question when the delay before the first question repetition was short. When delay before the second repetition was short it is as if the child realised that they were manufacturing responses and "gave up", finally giving the accurate "don't know" response. With answerable questions the first short repetitions and second long interval repetitions led to an increase in accuracy. The increase in accuracy for responses to repeated answerable questions might serve to remind children of the event and provide cues that enhance their recall for details.

Figure 5:2 showed that the first short repetition interval was most prone to change. When the first repetition occurred immediately after the original question it is likely that the child could remember their original response but understood the repetition to suggest that their original response was not appropriate. A change in response made at this stage was almost twice as likely to be sustained (resulting in a XYY repetition pattern) than return to the original response (XYX), which was generally more accurate than subsequent responses. With a second repetition it was the longer repetition interval that resulted in more changes in responses, which suggested that the child did not recall their previous response, or again believed that the interviewer was implicitly requesting a novel response.

In Study 4 the children were not put under direct pressure through suggestive questioning or by comments made by the interviewer about their responses (Finnilä et al., 2003; Warren et al., 1991) and yet their responses sometimes changed when the questions were repeated. As with Beuscher and Roebers (2005) it appeared that direct intervention in the form of a warning about saying "I don't know" to unanswerable questions (Home Office, 2001) did not facilitate accuracy in responses to unanswerable questions. This suggested that implicit in repetition was an element of inadequacy of the first response to which children acquiesced by changing their responses.

Our results indicated that to interview as early as possible, as suggested by the Pigot Report (Home Office, 1989), is advisable in terms of the accuracy and for the consistency of responses, but also indicated that even an early interview is not immune to problems of inaccuracy and inconsistency. In respect to consistency we found that the type of the questions and the way in which those questions were asked affected responses regardless of the time delay between event and interview or the age of the child.

In Studies 2, 3 and 4 considered the effects of question repetition after children have witnessed an event. In Study 5 we used a participatory event to establish whether results we had found when children had witnessed an event would also apply when children were more directly involved in the event.

Chapter 6

Study 5

The child eye-witness interviewing protocols such as ABE currently used in England and Wales (Home Office, 2001) do not distinguish between the child interviewee as witness and the child as participant or victim. In Studies 2, 3 and 4 the participants observed a staged event; they were, in effect, participating as witnesses only. In Study 5 the event required the children's active involvement, as participants. The interviewing regime to establish accuracy and consistency of recall in this study was the same as that used in the 1-week condition of Study 4 to enable comparison with the results of that study and ascertain the effect of participation rather than observation on recall elicited through questions.

Introduction

Accuracy and consistency of recall increased with age in interviews related to actual experienced events (Ghetti et al., 2002; Peterson et al., 2001) and in laboratory experiments (Gobbo, 2000; Goodman, Bottoms, Rudy, Davis, & Schwartz-Kenney, 2001; Hardy & Van Leeuwen, 2004; Krähenbühl & Blades, 2006; Powell, Thomson, & Ceci, 2003). In recall of the same information provided in different event modalities (see Chapter 1 pp. 33-34), participation, in comparison to observation, resulted in greater accuracy of responses to questions concerning that event (Roebers et al., 2004; Shrimpton et al., 1998; Thierry & Spence, 2004; Tobey & Goodman, 1992). However, these studies did not include unanswerable questions to which the accurate response would be "I don't know" (or similar). Previous research has shown children's reluctance to state that they do not know an answer (Peterson & Grant, 2001; Waterman et al., 2004), this reluctance was demonstrated by the poor accuracy of responses to unanswerable

questions in Studies 2, 3 and 4. However, given that accuracy of recall generally improved with participation it was possible that children would be more willing to answer unanswerable questions correctly when they had been directly involved in the event.

To date, the only study that incorporated repeated questions within a single interview after a participatory event was that of Powell and Thomson (1996). Powell and Thomson focused on the effects of repeated events on subsequent recall; however, in one of their conditions they included a single event with an interview just under one week later that contained question repetition. The repetition of the questions (all of which were answerable) was verbatim, occurred once only, and provided forced choice responses, that were coded for accuracy and for consistency with previous responses. Powell and Thomson found that accuracy and consistency increased with age (participants were aged 4-5 or 6-8-years-old) and that approximately one tenth of responses to repetitions changed.

In Study 5 we expected that the increase in accuracy and consistency of responses with age, as found in Studies 2, 3 and 4, would also be found for a participatory event (Peterson et al., 2001; Quas & Schaaf, 2002; Sussman, 2001). We also expected that this age related accuracy would be found in responses to repeated questions within an interview following a participatory event (Powell & Thomson, 1996). Accuracy was expected to decline after the initial repetition (Krähenbühl & Blades, 2006; Moston, 1987) and as shown in Studies 2, 3 and 4 the decline in accuracy would be found primarily in responses to unanswerable questions, which were also more likely to change (Beuscher & Roebers, 2005; Waterman et al., 2001). In accordance with the results of Studies 2 and 4 we

expected the most common pattern of responses to repetitions would be XYY where the initial repetition resulted in a novel response, which was then maintained through response to a further repetition. Undesirable shifts were expected to exceed desirable shifts (Howie et al., 2004) but, as found in Studies 2, 3 and 4, the most common change was expected to be a novel inaccurate change.

Method

Participants

This study was conducted with 163 participants. There were fifty-one 4-5-year-olds (19 boys and 32 girls, mean age = 4 years 5 months, *SD* = 3.3 months), fifty-three 6-7-year-olds (20 boys and 33 girls, mean age = 6 years 6 months, *SD* = 3.7 months), and fifty-nine 8-9-year-olds (34 boys and 25 girls, mean age = 8 years 6 months, *SD* = 3.3 months). Although no children withdrew, data from a further two 8-9-year-olds were omitted because one interview was interrupted and the other interview was terminated before completion because the child repeatedly replied, "don't know" before the interviewer had finished asking each of the questions.

Procedure

Groups of six children participated in a craft activity using a decoupage technique during morning sessions at their schools. The activity took approximately 30 minutes to complete. The nature of the activity was discussed with class teachers to establish whether it was a feasible activity with the age groups involved. A pilot study of the activity had been conducted with three children in each age group to ensure that the equipment and activity were appropriate for the age groups. The activity was not scripted but the procedure of

the activity, the clothing and equipment provided by the confederate and specific points and actions incorporated were the same in each session. The confederate had pre- and post-event checklists to ensure that all relevant aspects of the event had been included and in the same order,

Pre-event list

Clothing:

1. Wear trousers with a pocket
2. Wear your watch so that it is visible but the watch face is obscured
3. Wear a long sleeved sweatshirt – any colour except red
4. Whatever you wear under your sweatshirt make sure it is not visible

Equipment:

5. Put the garden string in your trouser pocket
6. Have your handbag with a packet of tissues in it
7. Make sure you have got your own scissors (these were distinctive being jewelled and in the shape of a bird)
8. Check decoupage box for:

    Demonstration card

    Blank cards

    Silicone glue

    Glue stick

    Cocktail stick

    Wrapping paper with repeating pictures

    Stickers

    Hologram paper

    Shape templates

Pen / pencil

Children's scissors

During the activity

State the following:

1. "My own finished cards are taken to a shop to be sold."
2. "I made cards with my husband Peter last week."
3. "You need to use special glue to do decoupage so that the pictures are raised up."
4. "You can buy special sheets to use for decoupage pictures and cards from the stall at the market in Leek."
5. "I have not brought the envelopes, they are at home."

Do the following:

6. Use the cocktail stick to place the silicone glue on the cards to create the decoupage effect
7. Take the string out of your pocket, state "That wasn't what I wanted" replace the string and continue with the activity
8. Take a packet of paper tissues out of your handbag, use a tissue, replace the packet and close the bag

An unfamiliar adult interviewed the children individually one week later. The teachers told the children that the interviewer wanted to talk to them about the decoupage activity.

The children in the participatory event took their craftwork home after all interviewing was completed. Receipt of the craftwork was unrelated to participation in the interview.

Materials and coding

The decoupage card making activity followed the same procedure for all age groups although the 8-9-year-olds were required to cut out their pictures where as the younger children used pre-cut pictures for their cards. The confederate guided and assisted the children as appropriate and incorporated activities, equipment and information in accordance with the pre- and post- event checklists. At the end of the activity the children did not receive their cards but were told that the cards would be returned to school the following week when the envelopes would be available. This was to reduce the possibility of discussion of the activity before the interview had taken place.

In the individual interviews the children were asked 40 questions, half of which related to information provided in the activity and half to observation of the confederate and her equipment. Of these, half of the questions were answerable, and half unanswerable. Eight open-ended questions (four answerable and four unanswerable) were repeated a further two times (24 questions). An example of an answerable repeatable question was "What does Chris do with her own finished cards?", which was repeated as "When her own cards are finished what does Chris do with them?", and "Tell me, what does Chris do with her own cards when they are finished?" An example of an unanswerable repeatable question was "What colour shirt did Chris wear under her sweatshirt?", which was repeated as "Chris was wearing a shirt under her sweatshirt – what colour was it?", and "Tell me, what was the colour of the shirt that Chris wore under her sweatshirt?" Four of the repeated questions were repeated with an initial interval of 17 ("long") intervening questions and then an immediate ("short") repetition and four in the opposite pattern. This pattern also

enabled half of the repetition delay intervals to form the first repetitions and half the second repetitions.

Sixteen non-repeated closed filler questions were used to provide the required intervals between the repeated questions and to promote a less artificial range of questions than an interview consisting of entirely repeated questions. Half of these related to information provided in the activity, half to observation, half were answerable and half unanswerable. An example of an answerable filler question was "Can you buy sheets for decoupage cards at the stall at the market in Leek?", and an unanswerable filler question was "Did Chris have a bruise on her elbow?" The order of repetition delays was counterbalanced to provide eight different questioning patterns to which the eight question repetitions were applied giving rise to 64 different interview orders. The filler questions were also randomised and used in between the repetitions to establish the required intervals. An example of a full set of questions is provided showing repetition order and intervals used:

F = filler question, R = repeated question, B = original question, I or II = first or second repetition, S or L = short or long repetition interval, A or U = answerable or unanswerable, numbers represent each of the 8 repeated questions. Chris was the name of the confederate.

FU	Did Chris have a plaster on her toe?
RBA1	Who did Chris make cards with at home last week?
FU	Was Chris's favourite card one with a mouse on it?
RBA2	What does Chris do with her own finished cards?
RISA2	Tell me, what does Chris do with her own cards when they are

	finished?
FA	Can you buy sheets for decoupage pictures to make cards from the stall at the market in Leek?
RBU1	When Chris makes cards at home which room does she make them in?
RISU1	Which room in her house does Chris make her cards in?
FU	Did Chris have a bruise on her elbow?
RBA3	Chris used something to put on the glue, what did she use?
FU	Was Chris hoping to find some fish paper to make into pictures for her cards?
RBA4	What did Chris take out of her pocket?
RISA4	Tell me what Chris took out of her pocket?
FA	Was Chris's card blue?
RBU2	What fruit did Chris have in her bag?
RBU3	What colour shirt did Chris wear under her sweatshirt?
RISU3	Tell me, what was the colour of the shirt that Chris wore under her sweatshirt?
FU	Was Chris's watch telling the right time?
RILA1	Who made cards at home with Chris last week?
RIISA1	Last week Chris made cards with someone at home - who was that?
RBU4	How old is Chris?
RIILA2	When her own cards are finished what does Chris do with them?
FA	Did Chris bring the envelopes with her?
FA	Did Chris show you her knitting?

RIILU1	Where does Chris make her cards when she is at home?
FA	Did you need to use a special glue for the cards?
RILA3	The glue was put on, what was this done with?
RIISA3	What did Chris use to put on the glue?
FU	Was there a packet of sweets in Chris's handbag?
RIILA4	Chris took something out of her pocket, what was it?
FU	Did Chris find birds difficult to make into decoupage pictures and cards?
RILU2	The fruit in Chris's bag, what sort was it?
RIISU2	Chris had fruit in her bag what fruit was it?
FA	Did Chris have a red sweatshirt on?
RIILU3	Chris was wearing a shirt under her sweatshirt - what colour was it?
FA	Were the cards kept in a plastic bag?
FU	Did Chris buy the cards in Hanley?
RILU4	Chris is older than you, how old is she?
RIISU4	What is Chris's age?
FA	Were the scissors that Chris used her own scissors?

The children's responses to all questions were coded for accuracy. Responses to repeated questions were also coded for consistency, direction and patterns of change as described in Study 2 (pp. 88-89). Any comments that the children made relating to the questions and repetition in general were recorded.

10% of transcripts were selected at random from each age group and re-coded to establish inter-rater reliability. Re-coding was by a trained rater who

was familiar with the activity used in the event but naïve to the hypotheses. The results for accuracy and change were both $\kappa = 0.98$ ($p < .001$).

Results

Preliminary analysis of all responses

Analyses for gender or interview delay (7 or 8 days) differences revealed no differences so these factors were not considered further.

Accuracy of all responses

This analysis was performed to check that the set of questions although different in wording resulted in similar differences in levels of accuracy in comparison to previous experiments.

A one-way between-subjects ANOVA was carried out on the accuracy of both the filler and repeated questions responses. There was an effect of age group: $F(2,160) = 36.01$, $p < .001$. Accuracy increased with age: 4-5-year-olds $M = 11.41$, (28.5%), $SD = 4.80$, 6-7-year-olds $M = 17.43$, (43.6%), $SD = 5.59$, and $M = 20.07$, (50.2%), $SD = 5.77$ for the 8-9-year-olds. A Tukey HSD comparison showed a difference between the 4-5-year-olds and other age groups ($p < .001$), and a difference between the 6-7-year-olds and 8-9-year-olds ($p < .05$).

Accuracy of base and repeated questions only

A 3 Age group (4-5, 6-7 or 8-9 years) x 2 Question type (answerable or unanswerable) x 3 Repetition order (base question, first or second repetition) analysis of variance (ANOVA) with repeated measures was carried out on the responses to repeated questions.

There was an effect of age group: $F(2,160) = 19.20$, $p < .001$, an effect of question type: $F(1,160) = 85.11$, $p < .001$, and an effect of repetition order: $F(2,159) = 3.24$, $p < .05$. There were two-way interactions between question type

and age: $F(2,160) = 12.13, p < .001$, and between question type and repetition order: $F(2,159) = 8.03, p < .001$.

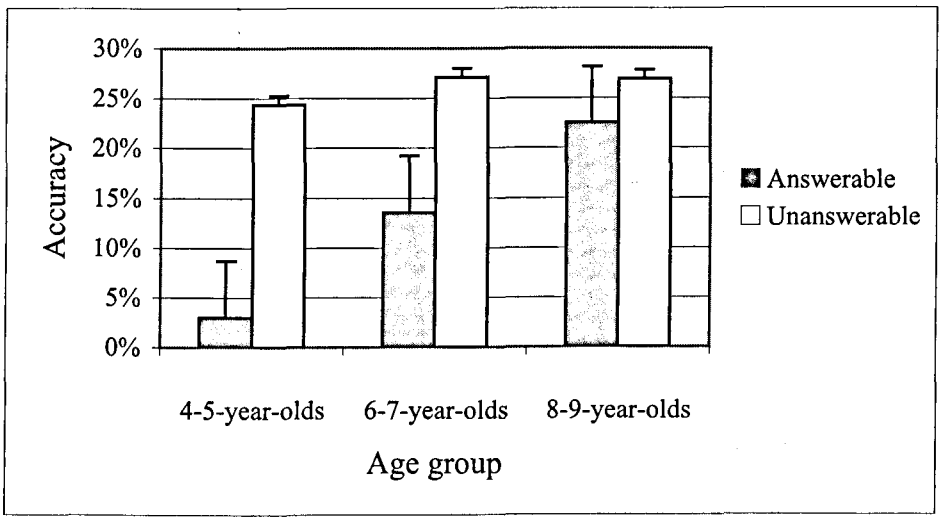
The effect of age group showed that accuracy increased with age: 4-5-year-olds were less accurate $M = 1.10$ (27.4%), $SD = 1.01$, than 6-7-year-olds $M = 1.63$ (40.7%), $SD = 1.16$, who were less accurate than 8-9-year-olds $M = 1.98$ (49.4%), $SD = 1.18$. A Tukey HSD comparison showed differences between the 4-5-year-olds and other age groups ($p < .001$) and between 6-7 and 8-9-year-olds ($p < .05$).

The effect of question type showed the accuracy of responses to answerable questions $M = 1.04$ (27.0%), $SD = 0.77$, and for responses to unanswerable questions $M = 2.09$ (52.4%), $SD = 0.92$. Pairwise comparisons showed a difference between the question types ($p < .001$).

The effect of repetition order showed a decline in accuracy after the responses to the base question from $M = 1.63$ (41.3%), $SD = 1.72$ to $M = 1.53$ (38.7%), $SD = 1.72$ for responses to the first repetition, and $M = 1.54$ (39.1%), $SD = 1.88$ for responses to the second repetition. Pairwise comparisons showed a difference between the base and both the first and second repetitions ($p < .05$) but not between the first and second repetitions.

The interaction between age group and question type showed that the differences in accuracy between responses to answerable and unanswerable questions decreased with age (Figure 6.1). Figure 6.1 showed that accuracy for responses to answerable questions increased with age whereas accuracy for responses to unanswerable questions was similar across age groups.

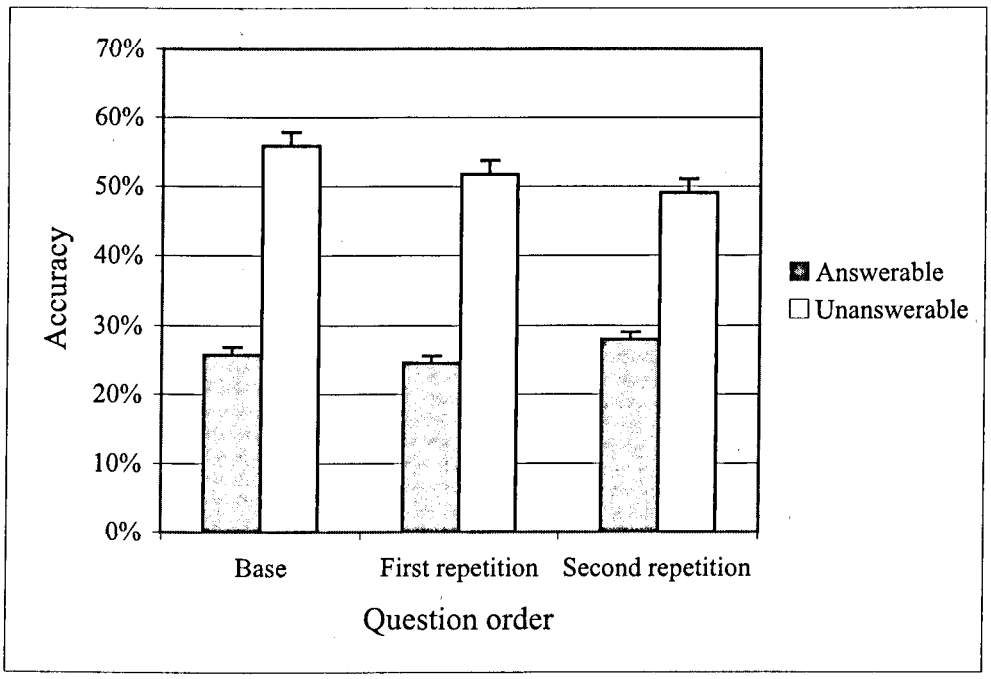
Figure 6:1
The accuracy of children's responses to answerable and unanswerable questions



Tests of simple effects showed a difference between the accuracy of question types (p <.001).

The interaction between question type and repetition order indicated that after an initial decline, responses to answerable questions increased in accuracy with the second repetition (see Figure 6.2). For responses to unanswerable questions there was a decline in accuracy with an increase in repetition.

Figure 6.2
The effect of question order on responses to answerable and unanswerable questions



Tests of simple effects were showed differences in accuracy between answerable and unanswerable responses for each question order (p <.001).

Effect of repetition delay and order on responses to repeated questions only (base excluded)

A 3 Age (4-5, 6-7 or 8-9 years) x 2 Question type (answerable or unanswerable) x 2 Repetition order (first or second repetition) x 2 Repetition interval (short or long interval between repetitions) analysis of variance (ANOVA) with repeated measures was carried out on the responses to repeated questions. There was an effect of age group: $F(2,160) = 18.46$, p <.001, and an effect of question type: $F(1,160) = 64.63$, p <.001. There were no other effects of single factors.

There were two-way interactions between age group and question type: $F(2,160) = 10.79$, p <.001, between question type and repetition order: $F(1,160)$

= 10.66, p <.001, and between repetition order and repetition interval: $F(1,160)$ = 6.98, p <.01.

There was a three-way interaction between age group, question type and repetition interval: $F(2,160) = 4.03$, p <.05.

The effect of age group showed that accuracy increased with age: 4-5-year-olds were less accurate $M = 0.53$ (26.3%), $SD = 0.70$ than 6-7-year-olds $M = 0.79$ (39.6%), $SD = 0.83$, who were less accurate than 8-9-year-olds $M = 0.98$ (49.0%), $SD = 0.81$. A Tukey HSD comparison showed a difference between the 4-5-year-olds and the other age groups (p <.001), and a difference between the 6-7 and 8-9-year-olds (p <.05).

The effect of question type showed accuracy for responses to answerable questions $M = 0.53$ (27.2%), $SD = 0.52$ and for responses to unanswerable questions $M = 1.01$ (50.5%), $SD = 0.66$. Pairwise comparisons showed a difference between the question types (p <.001).

The two-way interaction between repetition order and repetition interval showed that accuracy of responses improved with first repetitions when the interval between repetitions was long. Conversely, accuracy of responses to second repetitions declined with the longer intervals between repetitions (see Table 6:1). Tests of simple effects were conducted for short and long repetitions separately. These tests showed differences between repetition orders for long repetitions only (p <.05).

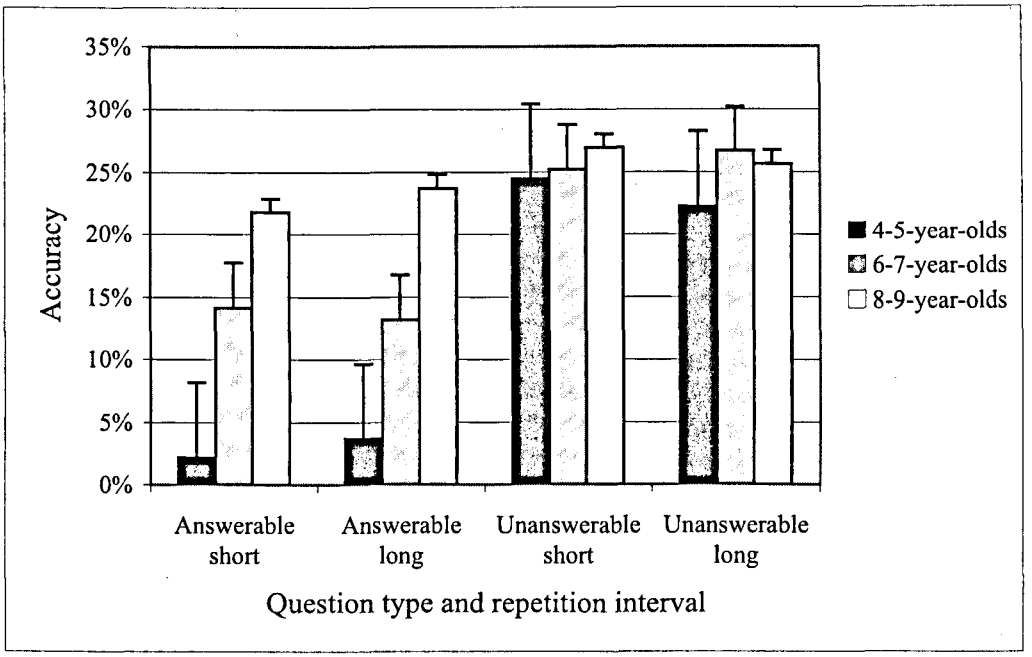
Table 6:1
The interaction between repetition order and repetition interval on the accuracy of responses

Repetition interval	Repetition order	
	First repetition	Second repetition
Short	34.9% (1.09)	41.7% (1.16)
Long	41.4% (1.08)	35.5% (1.13)

Standard deviations in parentheses

The three-way interaction between age group, repetition interval and question type showed the discrepancy between age groups in accuracy of responses to answerable questions which was greatly reduced in the differences between age groups responses to unanswerable questions (see Figure 6.3). For 4-5-year-olds and 8-9-year-olds accuracy improved slightly with a long interval for answerable questions but reduced slightly with the long interval for unanswerable questions. For 6-7-year-olds the converse was found. Tests of simple effects were conducted for each age group separately to determine differences in accuracy between short repetition interval answerable and unanswerable responses, and between long repetition interval answerable and unanswerable responses. Differences were found for both long and short repetition intervals for 4-5-year-olds and 6-7-year-olds (p <.001) but not for 8-9-year-olds.

Figure 6.3
The accuracy of answerable and unanswerable responses according to repetition interval and age of child



Effect on consistency

Consistency examined whether there had been a semantic change to the previous response given, a grammatical change only was not coded as being inconsistent with the previous response. We found that there were no changes at all (i.e. responses remained consistent throughout the interview) in responses to 51.9% of questions. Of these consistent responses 55.5% were accurate, and 44.5% were inaccurate.

Repeated questions resulted in changes (in comparison to the previous response given by the child) in 34.9% of all responses, 39.8% of all responses to answerable questions and 29.9% of all responses to unanswerable questions.

A 3 Age (4-5, 6-7 or 8-9 years) x 2 Question type (answerable or unanswerable) x 2 Repetition order (first or second repetition) x 2 Repetition interval (short or long interval between repetitions) analysis of variance (ANOVA) with repeated measures was carried out on the responses to repeated questions.

There was an effect of age group: $F(2,160) = 26.99$, $p < .001$, and of question type: $F(1,160) = 23.27$, $p < .001$. There were two-way interactions between repetition interval and age group: $F(2,160) = 7.56$, $p < .001$, and between repetition order and repetition interval: $F(1,299) = 8.82$, $p < .01$. There was a three-way interaction between question type, repetition order and repetition interval: $F(1,160) = 5.26$, $p < .05$.

The effect of age group showed that the number of changes in responses decreased with age. 4-5-year-olds made more changes in responses $M = 0.93$ (46.4%), $SD = 0.94$ than 6-7-year-olds $M = 0.76$ (37.9%), $SD = 0.87$, who made more changes in responses than 8-9-year-olds $M = 0.44$ (22.1%), $SD = 0.58$. A Tukey HSD comparison showed a difference between 8-9-year-olds and both 6-7 and 4-5-year-olds ($p < .001$) but not between the 4-5 and 6-7-year-olds.

There was an effect of question type, with more changes in responses to answerable questions $M = 0.81$ (39.8%), $SD = 0.50$ than in changes in responses to unanswerable questions $M = 0.61$ (29.9%), $SD = 0.54$. Pairwise comparisons showed a difference in changes of responses between answerable and unanswerable questions ($p < .001$).

The two-way interaction between repetition interval and age group showed a decrease in the number of changes in both short and long intervals between repetitions with an increase in age. The difference in the number of changes according to age was greater among short interval repetitions than among long interval repetitions. Tests of simple effects showed a difference between the 8-9-year-olds and both the 4-5 and 6-7-year-olds ($p < .001$) for short interval repetitions. For long interval repetitions there were differences between

all age groups: between 4-5 and 6-7-year-olds (p <.01), between 4-5 and 8-9-year-olds (p <.001), and between 6-7 and 8-9-year-olds (p <.05).

The three-way interaction between question type, repetition order and repetition interval is shown in Table 6:2. This interaction showed that responses to answerable questions were changed more often particularly in response to first short or second long repetitions. For unanswerable repetitions the long intervals resulted in slightly more changes in responses than the short intervals but this difference was not as pronounced as the difference in changes in response to answerable questions. Tests of simple effects were conducted for changes between answerable and unanswerable responses; there were differences between changes in responses to first repetitions p <.01, second repetitions p <.001, short repetition intervals p <.001, and long repetition intervals p <.05.

Table 6:2
The changes in responses to answerable and unanswerable questions according to repetition order and repetition interval

	Question type and repetition order			
Repetition interval	Answerable first	Unanswerable first	Answerable second	Unanswerable second
Short	48.2% (0.82)	30.1% (0.80)	33.7% (0.76)	27.9% (0.88)
Long	32.5% (0.73)	32.5% (0.76)	44.8% (0.84)	29.1% (0.74)

Standard deviations in parentheses

Shifts in responses as a result of repetition

Directions of shifts

The most common shift in all age groups was a novel inaccurate shift M = 3.39 (61.1%), SD = 1.06. Undesirable shifts where an accurate response became inaccurate were more common M = 1.17 (21.1%), SD = 1.00 than desirable shifts where an inaccurate response became accurate M = 0.99 (17.8%),

SD = 1.07. Shifts according to changes made by each age group as a percentage of the shifts made by that age group are shown in Table 6:3.

Table 6:3
The direction of shifts as a percentage of the shifts made by each age group of child

	Shift direction		
Age group	Novel inaccurate	Becomes inaccurate	Becomes accurate
4-5-year-olds	72.3% (3.63)	16.4% (1.01)	11.3% (0.92)
6-7-year-olds	57.0% (2.72)	23.4% (1.10)	19.9% (1.17)
8-9-year-olds	45.9% (1.74)	27.3% (1.08)	26.8% (1.07)

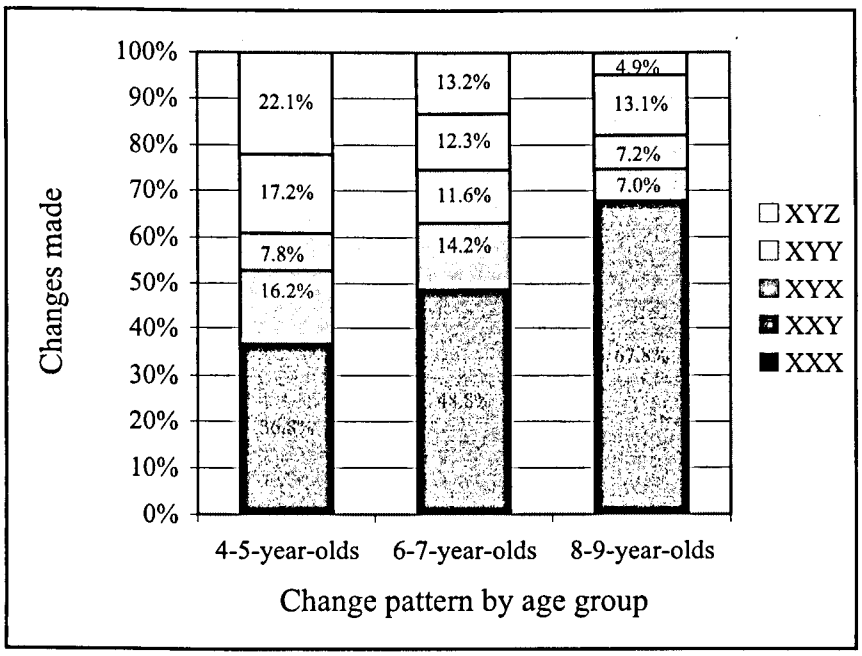
Standard deviations in parentheses

Patterns of changes

The children remained completely consistent in 51.1% of their responses to repetitions. The most common pattern of changes in responses to repetitions was where a novel response was subsequently maintained (XYY 14.2%). A novel response to each repetition (XYZ) was found in 13.4% of responses and a novel response to the second repetition only was found in 12.4% of responses to repetitions. The least common response shift pattern was where an original response was returned to after an intervening novel response (XYX 8.9%).

The 5-year-olds gave novel responses to all repetitions (XYZ 22%) compared to (XYZ 13%) with 7-year-olds, and (XYZ 5%) with 9-year-olds. The distribution of response patterns is shown in Figure 6:4.

Figure 6:4
The distribution of shift in response patterns made according to age group



Children's comments

7 children (4%) commented on the question repetition (one 4-5-year-old, two 6-7-year-olds, and four 8-9-year-olds). All of the comments mentioned repetition e.g. "You've already asked that", "Exactly the same", "You said that just". Regardless of the comments made all of the children continued without any intervention on the interviewer's behalf. None of these children changed their response (in comparison to their previous response) after the question to which they commented.

Discussion

In Study 5 we investigated what effect active involvement in an event might have on children's accuracy and consistency in responses to repeated questions. As expected the accuracy and consistency of children's responses to repetitions increased with age (Ghetti et al., 2002; Krähenbühl & Blades, 2006; Peterson et al., 2001; Powell et al., 2003). There was also a decline in accuracy

after the first repetition which is consistent with previous research (Krähenbühl & Blades, 2006; Moston, 1987) and the results of Studies 2, 3 and 4. However, within these general findings there were notable differences in the patterns (of accuracy and consistency) of responses according to the question type (answerable or unanswerable questions) between the responses given by the children in Study 5 and in Powell and Thomson (1996), and to the responses given to repeated questions by children in previous studies who had observed a stimulus event only.

Approximately one third of all responses to repetitions changed, compared to the previous response given. This amount of change exceeded the one in ten changes in responses in Powell and Thomson's (1996) results for verbatim repetitions after a practical activity event. The difference in these results may be linked to the repetition style; in Study 5 we used gist repetition, whereas Powell and Thomson used verbatim repetition. Study 1 showed that verbatim repetitions resulted in greater consistency in responses to repetitions than in responses to other repetition styles, including gist repetition, which was used in Study 5. Verbatim repetition is cautioned against in the current interviewing protocol used in England and Wales (Home Office, 2001) because such repetition might result in a change in response due to interviewer pressure rather than because the change is an accurate response to repetition. Our results suggested that gist repetition was more likely than verbatim repetition to result in a change in response. The accuracy results (and in accordance with Studies 2, 3 and 4) for gist repetition suggested that a change in response was often detrimental to accuracy.

Our results showed that when, and how often, a question was repeated affected accuracy. Repetitions that were the first repetition after a long delay or a second repetition after a short delay (see Table 6:1) were not only more accurate than other repetitions, but were also less likely to be changed (see Table 6:2).

There were many similarities between the patterns of results of Study 5 and the results from the observational event Studies 2, 3 and 4. The majority of changes were novel inaccurate changes when an incorrect response changed to a further different incorrect response. Similarly, undesirable shifts in responses exceeded desirable shifts (Howie et al., 2004). The similarities in results suggested that the effects of repetition are generalisable and not specific to context although to establish this conclusively would require further experiments where the event modality is counterbalanced.

Unlike previous research incorporating unanswerable questions with an instruction to respond to unanswerable questions appropriately (Beuscher & Roebers, 2005), Study 5 found that accuracy in responses to unanswerable questions was unaffected by age. In interviews following a witnessed event Beuscher and Roebers found that children responded accurately in over a half of responses to unanswerable questions and that accuracy in responses to these questions decreased with age. Beuscher and Roebers suggested that greater accuracy in the younger children showed that their ability to generate alternative answers was limited. However, after a practical activity involving limited participation (in the form of a discussion), Waterman, Blades and Spencer (2004) found that accuracy increased with age and that the accuracy was in approximately four fifths of responses to open-ended questions. Our results suggested that although a practical activity may improve accuracy to

unanswerable questions this positive effect was negated with repetition of the questions, particularly among older children.

Responses to half of the questions asked (through all repetitions) were maintained consistently (resulting in XXX coding), which was higher than in observation event Studies 2, 3 and 4. This suggested that when children actively took part in the event they were not only more willing to say that they did "not know" an answer to a question, but that they were more certain of that response, and would be less willing to change it in response to repetitions than when they had observed an event. However, approximately half of the fully consistent responses were inaccurate. This showed that consistency did not equate with accuracy, children were just as likely to remain consistent with an inaccurate response as with an accurate response.

Our results showed that question repetition was not beneficial to accuracy; whilst responses may in part improve accuracy, overall accuracy declined. Changes in responses either reduced accuracy or merely provided a further inaccurate response. Changes of any kind made children appear inconsistent, which may be perceived to be detrimental to their credibility (Davis et al., 1999). Most salient, however, was that consistency did not equate with accuracy; responses that remained consistent did so regardless of the accuracy of those responses.

Chapter 7

General Discussion

We have shown that question repetition, in diverse forms, is a frequent, necessary and primarily commendable feature of interviewing practice without which an interviewer might be unable to gain further information, clarification, or elicit a verbal response at all. Our studies have revealed that the use of question repetition results in frequent changes in response in relation to the previous response elicited. The effects of these changes are related to the form of repetition used, whether the question is answerable or unanswerable, and the way in which the repetition is conducted.

Repeating questions may have a positive effect in providing the additional or clarified information sought by the interviewer. However, the results of Studies 2, 3, 4 and 5 indicated that changes to responses were more likely to be detrimental to accuracy than beneficial (Howie et al., 2004). The most common change in response to repetition was to provide a further novel inaccurate response which had no effect on accuracy but indicated inconsistency.

Our experimental studies have shown that the detrimental effects of repetition may be greater than was recognised in previous research. The quantity and frequency of repetition found in Study 1 showed that experimental studies using single repetitions (primarily in verbatim format) of a limited number of questions have done little to replicate, even on a reduced scale, actual interviewing practices (Finnilä et al., 2003; Moston, 1987; Poole & White, 1991; Powell & Thomson, 1996; Warren et al., 1991). Similarly, reliance solely on examination of total accuracy scores for repetitions did not determine the number of changes made in responses. This deficiency had been addressed, to a certain

extent by Howie et al. (2004) but Howie et al. did not fully examine the amount of inconsistency found in children's responses but examined changes in accuracy only.

Our results indicate that the perceived benefits of reducing contradictions (Lamb & Fauchier, 2001) and the negative impact inconsistency has on children's credibility as witnesses (Davis et al., 1999; Gallagher & Pease, 2000; Regan & Baker, 1998) may be misplaced. In accordance with Quas et al. (2007) we have demonstrated that consistency is not necessarily an indicator of accuracy, particularly when the child has been actively involved (as in Study 5) in the event under discussion. The issue of whether consistency equates with accuracy has begun to be addressed for adult participants in an experimental context (Pezdek et al., 2007). If the link between consistency and accuracy was found to be unverifiable with children this would necessitate a considerable shift in police and public (especially in the role of jurors) perception of the implications attributed to consistency and inconsistencies in testimony.

The results of Study 1 on original transcripts of police interviews with young children alleging abuse raised many questions concerning the dynamics of the interviewing situation, the roles of the interviewer and the child, and how the interviewing format affected the information elicited. The interviewer influenced the child's responses, both directly and indirectly, by the questioning regime implemented (Hershkowitz, 2001; Orbach & Lamb, 2000). However, the child interviewed was not a passive recipient reacting exclusively to the interviewer. On the contrary, the child's response was influenced by broader social and conversational conventions, an individual perception of what was being asked,

how this was done, why, and what the child believed would be an appropriate or acceptable response (Freeman et al., 1981; Grice, 1989; Siegal, 1991).

The effective communication described by Grice (1989), whereby the utterances of the interviewer are true, informative, relevant and clear, may be assumed by both interviewer and child to be in place, especially after the inclusion of instructions or ground rules. Although the interviewer may be clear about his or her intention, this may not be as clear to a child who has limited experience of situations where a relatively unfamiliar adult in a position of authority asks for responses to questions the child may not wish to discuss, does not know, or which have already been answered in earlier interviews or questions. However, possibly for the reasons outlined by Siegal (1991), the child seemed to understand or made the assumption that if a question was repeated their original response was inadequate and subsequently provided an alternative. What remains unclear is how much age related accuracy and consistency (Beuscher & Roebers, 2005; Ghetti et al., 2002; Gobbo et al., 2002; Poole & White, 1991, 1993) can be attributed to developmental advances, and how much is the result of a lack of shared understanding of the interviewing process. Interviewer and child may both assume that they have a mutual understanding of the interview process but this has not been established with any certainty.

The aim of the ground rules and instructions in interviews is to create shared understanding, but children did not appear to follow this guidance, failing in particular to state when they do not know an answer to a question (Beuscher & Roebers, 2005; Krähenbühl & Blades, 2006; Peterson et al., 1999; Price & Connolly, 2004), even when the question was open ended (Waterman et al., 2001, 2004). The ABE, MOGP and NICHD interviewing protocols (Home

Office, 1992, 2001; Orbach et al., 2000) state that it is both acceptable and appropriate for a child to say "I don't know" in response to a question. In Study 1 we were unable to verify whether or not the child being interviewed was able to answer a question or not, but in Studies 2, 3, 4 and 5 responses were verifiable. Nevertheless the accuracy of responses to unanswerable questions barely rose above half in any of the studies.

The transcripts used in Study 1 were conducted under the MOGP (Home Office, 1992) interviewing protocol which expressed concern with the use of question repetition. Nevertheless, Study 1 showed that repetition was frequently used, mostly for appropriate reasons, and could not be avoided when the child did not answer the initial question. When a child remained silent, the question was repeated, often until a response was elicited. To continue repeating questions in this way could have changed the implicature of the role of questioning for that portion of conversation (Grice, 1989); the conversation had moved from the basis of responding accurately to the requirement that a response, any response, was required if the interview was to move onwards.

This result was particularly pertinent in understanding the responses of the younger children in Study 1, who often remained silent, providing no verbal response as their initial response. It is possible that this first silent response, as shown by the results of Studies 2, 3, 4 and 5, and by Memon and Vartoukian (1996), was the most accurate response and that either the child did not know the answer or was unable or unwilling to verbalise the response. Repetition of questions following initial silence did generally result in a response being given but as a second (or third, fourth etc.) response these may have been less accurate than the initial silent response.

The younger children in Study 1 had the shortest interviews but experienced most frequent question repetitions. These factors, in conjunction with the infrequent use of the free narrative stage (Davies et al., 2000; Sternberg et al., 2001; Warren et al., 1996) in which the information elicited was expected to be least tainted by other influences (Hershkowitz, 2001; Lamb et al., 2003a; Orbach & Lamb, 2000) and the use of a majority of closed format questions indicated that the interviewers had difficulties conducting interviews in accordance with the protocol guidelines, particularly with younger children (Cederborg et al., 2000; Lamb et al., 1996; Thoresen et al., 2006).

The relatively few comments made by the children in Studies 2, 3, 4 and 5 about the use of repetition cannot be assumed to represent an awareness and/or acceptance on the part of the child of a shared understanding about the purpose of the interview in general and the question repetition in particular. To encourage an understanding of the requirements specific to a police interview, 'ground rules' are established in the rapport stage of the interview (Home Office, 1992, 2001; Orbach & Lamb, 2000). However, although discussion of ground rules is advocated in the MOGP (Home Office, 1992) there is no explicit expectation or guidance on the assessment of children's understanding, acceptance and use of those rules.

Through a small follow-up study we gained understanding of police interviewers perceptions of their own interviewing experiences. Appendix 3 includes a summary of semi-structured interviews with experienced police interviewers trained under the ABE interviewing protocol (Home Office, 2001). We established that the introduction of the ground rules is a procedure that police interviewers find difficult to implement fully. We also found that police

interviewers had a lack of awareness about the extent to which question repetition was used and its effect on children's responses. These officers considerably underestimated the amount of repetition used in forensic interviews and were unaware of the potential problems related to its use regarding accuracy and consistency. Difficulties associated with the use of repetition were described only in terms of reducing interviewee cooperation or concern that the repetitive style of questioning would be considered by others to be inappropriate. The officers were, however, very clear on the importance of using open-ended questions. Future researchers might consider more extensive interviews with police interviewers as a means of providing insights into interviewing practices.

The use of open-ended questions is advocated by interviewing protocols (Home Office, 1992, 2001; Orbach et al., 2000) as advantageous in eliciting accurate, complete and detailed information from children (Dent & Stephenson, 1979; Peterson et al., 1999). However, the results of Study 1 showed that, as is common in other countries (Cederborg et al., 2000; Lamb et al., 1996; Thoresen et al., 2006), closed questions were more frequently used than open-ended questions although the majority of repetitions found were in gist format. Studies 3, 4 and 5 indicated that even if the recommended question and repetition forms were used, half of all responses to unanswerable questions would remain inaccurate.

The accuracy of responses to unanswerable questions was not affected by active participation in the event. Previous researchers have demonstrated that active participation has a positive effect on accuracy of recall (Baker-Ward et al., 1990; Gobbo et al., 2002; Tobey & Goodman, 1992). However, we did not find any difference in accuracy for responses to repeated unanswerable questions

between Study 5, when the children participated in the event, and Studies 2, 3 and 4 where children observed a staged presentation.

Despite occasional improvements in accuracy following question repetition (Goodman et al., 1991), we did not find a specific form of repetition or implementation pattern that consistently improved accuracy. Nevertheless, we showed that for witnessed events, second repetitions, particularly after a long interval between repetitions, were beneficial to accuracy for responses to repeated answerable questions but detrimental for responses to repeated unanswerable questions. Further investigation using a single event in participatory and non-participatory form would be required to confirm this result.

Theories associated with distortions of memory recall provided partial explanation for the changes in responses children made when they encountered repeated questions. The experimental results did not fully support any single theory but this was due, at least in part, to the experimental designs that were not primarily concerned with examining the impact of such theories.

Recognition of children's source monitoring (Roberts, 2000) was only possible with the few children in Studies 2, 3, 4 and 5 who verbalised their awareness that repetition had taken place. In a total of 784 participants only 43 children (5.5%) made a comment regarding the repetition of questions. Such comments generally demonstrated recognition of repetition. Following this recognition it was unclear whether the children's responses to subsequent question repetitions were their 'actual' responses sourced directly from their memories of the event itself, or from the memory of the previous response the children had given to questions earlier in the interview. The three children in Studies 3 and 4 who actually challenged the use of repetition did not change their

responses to those questions that gave rise to their challenges. In general, children changed between a quarter and a third of all of their responses to repeated questions, which suggested that either their source monitoring skills were relatively poor across all age groups (although improved with age) or that changes to responses should be attributed to other influences.

Initially it appeared that our results in Studies 2 and 4 could be explained by fuzzy-trace representations where initial encoding is in verbatim form (Brainerd & Reyna, 1995; Reyna & Brainerd, 1995, 1998). The effect of the encoded representation of information in children's recall could be explained on two levels. First, the representation accessed for recall was formed in relation to the previously experienced event. Second, the representation was formed during the interview and accessed the questions that had been asked and the responses given.

In the short timescale within a single interview, fuzzy-trace theory would emphasise the use of verbatim encoding (Brainerd & Reyna, 1995; Davies, 1995; Reyna & Brainerd, 1995, 1998), which would enable children to recognise that a question was repeated and therefore would link further repetitions to their initial response thus enabling consistency. The level of consistency in responses to repetitions was highest amongst verbatim repetitions in Study 2. Our results showed that responses to verbatim repetitions were changed least often suggesting that the children did indeed access their previous responses. Changes in responses were therefore not made because children did not remember what they had said before as their answer to questions about the event but were made as a reaction to the act of repetition.

In fuzzy-trace theory delays of approximately one week between event and recall would lead to representations being encoded in gist form (Reyna & Kiernan, 1994). This gist representation would enable children to access their memories of the event relatively accurately even if the repeated questions asked were in gist form. Changes in responses to such gist repetitions would not be attributed to the delay between event and interview as the children had access to a stored representation. As such, consistency (although not necessarily accuracy) should not be affected by the time delay would result from the act of repetition itself. However, the results of Study 4 showed that a delayed interview resulted in lower consistency than an earlier interview. This suggested that the representation accessed in order to respond to repeated questions was in relation to the event rather than to the interview, which therefore does not support the inferences drawn from fuzzy-trace theory.

The theory that most appropriately explained our results was that of schema theory (Bartlett, 1932; Roberts, 2002), which assumes that multiple experiences result in the formation of a cohesive representation of the event. According to this theory the formation of a script provides an expectation of what to expect from discussions with adults and what is expected when a question is repeated. The lack of accuracy and consistency by all age groups in responses to repeated questions suggested that children aged 4-9 have developed two related scripts in relation to appropriate responses to an adult's question. First, a script was formed that stated that an answer providing information of some description is required regardless of whether access to that information is possible or not. Second, a further script that stated that if a question is repeated then this is an adult's covert or implicit demand for a response that is different to

the previous response. The effect of these scripts was particularly noticeable in relation to children's reluctance to answer repetitions of unanswerable questions accurately (Beuscher & Roebers, 2005; Waterman et al., 2001, 2004) and children's willingness to change their responses to both answerable and unanswerable repeated questions (Krähenbühl & Blades, 2006).

The results of Study 3 in particular demonstrated the effect of a developmental shift in the acquisition of scripts (Farrar & Goodman, 1990; Fivush, Kuebli, & Clubb, 1992). In Study 3 there was an interaction between question type (answerable and unanswerable) and age in relation to accuracy (Beuscher & Roebers, 2005; Krähenbühl & Blades, 2006; Waterman et al., 2001, 2004). Whilst the older children were more accurate than the younger children in their responses to both answerable and unanswerable questions the discrepancy between the accuracy of answerable and unanswerable questions was far greater amongst the older children. This suggested that the script that a response to unanswerable question repetitions was required regardless of accuracy was more highly established in the older children's reactions to multiple repetitions. The children continued to search for a response they believed to be acceptable giving rise to accuracy that was low relative to their accuracy for answerable questions. This suggests that there is little virtue in terms of accuracy (and also in respect to consistency) in the use of multiple repetitions with children when the question is unanswerable.

Future research is required to establish how repetition can be used to elicit accurate and detailed information, maintain credibility, and to create an interviewing context in which both the interviewer and the child share understanding of the implications of repetition, how the interview will proceed,

and what is actually expected. It will be necessary to develop strategies to use in future interviewing protocols when a child does not respond in a complete or appropriate manner, because, as we have shown, question repetition is not conducive to achieving clear, accurate and consistent information in eyewitness interviews.

In conclusion, the studies in this thesis have shown how difficult it is to question children appropriately and yet elicit the required information. It is the understanding of the use of repetition that needs to be addressed. It is impractical to either avoid repetition altogether or to use repetition only in the limited and complex prescribed circumstances that may be beneficial to children's responses.

References

A.M.Macdonald (Ed.). (1981). *Chambers twentieth century dictionary*. Bath: The Pitman Press.

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology, 60*, 57-83.

Acredolo, C. (1995). Intuition and gist. *Learning and Individual Differences, 7*, 83-86.

Agnew, S. E., Powell, M. B., & Snow, P. C. (2006). An examination of the questioning styles of police officers and caregivers when interviewing children with intellectual disabilities. *Legal and Criminological Psychology, 11*, 35-53.

Ainsworth, P. B. (1998). *Psychology, law & eyewitness testimony*. Sussex: Wiley & Sons Ltd.

Akehurst, L., Köhnken, G., & Hofer, E. (2001). Content credibility of accounts derived from live and video presentations. *Legal and Criminological Psychology, 6*, 65-83.

Akehurst, L., Milne, R., & Köhnken, G. (2003). The effects of children's age and delay on recall in a cognitive or structured interview. *Psychology, Crime & Law, 9*, 97-107.

Aldridge, N. C. (1999). Enhancing children's memory through cognitive interviewing: An assessment technique for social work practice. *Child and Adolescent Social Work Journal, 16*, 101-126.

Alexander, K. W., Goodman, G. S., Schaaf, J. M., Edelstein, R. S., Quas, J. A., & Shaver, P. R. (2002). The role of attachment and cognitive inhibition in children's memory and suggestibility for a stressful event. *Journal of Experimental Child Psychology, 83*, 262-290.

Allwood, C. M., Granhag, P. A., & Jonsson, A.-C. (2006). Child witnesses' metamemory realism. *Scandinavian Journal of Psychology, 47*, 461-470.

American Prosecutors Research Institute. (2003). *Finding words: Half a nation by 2010: Interviewing children and preparing for court*: American Prosecutors Research Institute.

Bahrick, L. E., Parker, J. F., Fivush, R., & Levitt, M. (1998). The effects of stress on young children's memory for a natural disaster. *Journal of Experimental Psychology: Applied, 4*, 308-331.

Baker-Ward, L., Gordon, B. N., Ornstein, P. A., Larus, D. M., & Clubb, P. A. (1993). Young children's long-term retention of a pediatric examination. *Child Development, 64*, 1519-1533.

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. *Cognitive Development, 5*, 55-69.

Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Berliner, L., & Conte, J. R. (1995). The effects of disclosure and intervention on sexually abused children. *Child Abuse and Neglect, 19*, 371-384.

Beuscher, E., & Roebers, C. E. M. (2005). Does a warning help children to more accurately remember an event, to resist misleading questions, and to identify unanswerable questions? *Experimental Psychology, 52*, 1-10.

Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Douglas Brown, R., Park, C. L., Ernst, K., et al. (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.

Bowen, C. J., & Howie, P. M. (2002). Context and cue cards in young children's testimony: A comparison of brief narrative elaboration and context reinstatement. *Journal of Applied Psychology, 87*, 1077-1085.

Brainerd, C. J., & Reyna, V. F. (1995). Fuzzy-Trace theory: Some foundational issues. *Learning and Individual Differences, 7*, 145-162.

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-Trace theory and children's false memories. *Journal of Experimental Child Psychology, 71*, 81-129.

Brown, D., & Pipe, M.-E. (2003a). Individual differences in children's event memory reports and the narrative elaboration technique. *Journal of Applied Psychology, 88*, 195-206.

Brown, D., & Pipe, M.-E. (2003b). Variations on a technique: Enhancing children's recall using narrative elaboration. *Applied Cognitive Psychology, 17*, 377-399.

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*, 520-554.

Bull, R. (1992). Obtaining evidence expertly: The reliability of interviews with child witnesses. *The International Digest of Human Behaviour Science and Law, 1*, 3-36.

Burgwyn-Bailes, E., Baker-Ward, L., Gordon, B., & Ornstein, P. A. (2001). Children's memory for emergency medical treatment after one year: The impact of individual difference variables on recall and suggestibility. *Applied Cognitive Psychology, 15*, 25-48.

Butler-Sloss. (1987). Report of the inquiry into child abuse.

Candel, I., Merckelbach, M., & Muris, P. (2000). Measuring interrogative suggestibility in children: Reliability and validity of the Bonn Test of Statement Suggestibility. *Psychology, Crime & Law, 6*, 61-70.

Cassel, W. S., & Bjorklund, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: An ecologically based short-term longitudinal study. *Law and Human Behavior, 19*, 507-532.

Cassel, W. S., Roebers, C. E. M., & Bjorklund, D. F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology, 61*, 116-133.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.

Ceci, S. J., Huffman, M. L., & Smith, E. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition, 3*, 388-407.

Cederborg, A.-C. (2004). Factors affecting children's disclosure of secrets in an investigatory interview. *Scandinavian Journal of Psychology, 45*, 197-205.

Cederborg, A.-C., Orbach, Y., Sternberg, K. J., & Lamb, M. E. (2000). Investigative interviews of child witnesses in Sweden. *Child Abuse and Neglect, 24*, 1355-1361.

Chae, Y., & Ceci, S. J. (2005). Individual differences in children's recall and suggestibility: The effect of intelligence, temperament, and self-perceptions. *Applied Cognitive Psychology, 19*, 383-407.

Chen, E., Zeltzer, L. K., Craske, M. G., & Katz, E. R. (2000). Children's memories for painful cancer treatment procedures: Implications for distress. *Child Development, 71*, 933-947.

Clarke-Stewart, K. A., Malloy, L. C., & Allhusen, V. D. (2004). Verbal ability, self-control, and close relationships with parents protect children against misleading suggestions. *Applied Cognitive Psychology, 18*, 1037-1058.

Cole, C. B., & Loftus, E. F. (1987). The memory of children. In M.J.Ceci, M.P.Toglia & D.F.Ross (Eds.), *Children's eyewitness memory* (pp. 178-208). New York: Springer Verlag.

Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology, 93*, 207-223.

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.

Cordón, I. M., Saetermoe, C. L., & Goodman, G. S. (2005). Facilitating children's accurate responses: Conversational rules and interview style. *Applied Cognitive Psychology, 19*, 249-266.

Cowan, N. (1998). Children's memories according to Fuzzy-Trace theory: An endorsement of the theory's purpose and some suggestions to improve its application. *Journal of Experimental Child Psychology, 71*, 144-154.

Coxon, P., & Valentine, T. (1997). The effects of the age of eyewitness on the accuracy and suggestibility of their testimony. *Applied Cognitive Psychology, 11*, 415-430.

Cronch, L. E., Viljoen, J. L., & Hansen, D. J. (2006). Forensic interviewing in child sexual abuse cases: Current techniques and future directions. *Aggression and Violent Behavior, 11*, 195-207.

Davies, G. M. (1995). Fuzzy-Trace theory and eyewitness memory. *Learning and Individual Differences, 7*, 111-114.

Davies, G. M., Marshall, E., & Robertson, N. (1998). *Child Abuse: Training investigating officers*. London: Home Office Research, Development and Statistics Directorate.

Davies, G. M., & Westcott, H. L. (1999). *Interviewing child witnesses under the Memorandum of Good Practice: A research review*. London: Home Office.

Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child abuse victims. *Psychology, Crime & Law, 6*, 81-97.

Davies, G. M., Wilson, C., Mitchell, R., & Milsom, J. (1995). *Videotaping children's evidence: An evaluation*. London: Home Office.

Davis, G., Hoyano, L., Keenan, C., Maitland, L., & Morgan, R. (1999). *An assessment of the admissibility and sufficiency of evidence in child abuse prosecutions*. London: Home Office.

Dent, H. (1992). The effects of age and ability on eyewitnessing ability. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 1-13). Chichester: John Wiley and Sons Ltd.

Dent, H. R., & Stephenson, G. M. (1979). An experimental study of the effectiveness of different techniques of questioning witnesses. *British Journal of Social and Clinical Psychology, 18*, 41-51.

Donaldson, M. (1978). *Children's minds*. New York: Norton.
Dunsmore, J. C., Halberstadt, A. G., Eaton, K. L., & Robinson, M. L. (2005). Mothers' typical and event-specific positive expressions influence children's memory for events. *Social Development, 14*, 339-360.
Endres, J., Poggenpohl, C., & Erben, C. (1999). Repetitions, warnings and video: Cognitive and motivational components in preschool children's suggestibility. *Legal and Criminological Psychology, 4*, 129-146.
Erdmann, K., Volbert, R., & Böhm, C. (2004). Children report suggested events even when interviewed in a non-suggestive manner: What are the implications for credibility assessment? *Applied Cognitive Psychology, 18*, 589-611.
Farrar, M. J., & Goodman, G. S. (1990). Developmental differences in the relation between scripts and episodic memory: Do they exist? In R. Fivush & J.A.Hudson (Eds.), *Knowing and remembering in young children* (pp. 30-64). New York: Cambridge University Press.
Finnilä, K., Mahlberg, N., Santilla, P., Sandnabba, K., & Niemi, P. (2003). Validity of a test of children's suggestibility for predicting responses to two interview situations differing in their degree of suggestiveness. *Journal of Experimental Child Psychology, 85*, 32-49.
Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.
Fisher, R. P., Geiselman, R. E., & Amador, M. (1989). Field test of the cognitive interview: Enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology, 74*, 722-727.
Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G.S.Goodman & B.L.Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 1-24). New York: The Guildford Press.
Fivush, R., Haden, C. A., & Salimah, A. (1995). Structure and coherence of preschoolers' personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology, 60*, 32-56.
Fivush, R., Kuebli, J., & Clubb, P. A. (1992). The structure of events and event representations: A developmental analysis. *Child Development, 63*, 188-201.
Fivush, R., Peterson, C., & Schwarzmueller, A. (2002). Questions and answers: The credibility of child witnesses in the context of specific questioning techniques. In M.L.Eisen, J.A.Quas & G.S.Goodman (Eds.), *Memory and suggestibility in the forensic interview*. Hillsdale, N.J.: Erlbaum.
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*. California: Sage publications.
Flin, R., Boon, J., Knox, A., & Bull, R. (1992a). The effect of a five-month delay on children's and adults' eyewitness memory. *British Journal of Psychology, 83*, 323-336.
Flin, R., Bull, R., Boon, J., & Knox, A. (1992b). Children in the witness-box. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 167-189). Chichester: John Wiley & Sons, Ltd.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined events: A developmental comparison. *Child Development, 54*, 51-60.

Freeman, N. H., Sinha, C. G., & Condliffe, S. G. (1981). Collaboration and confrontation with young children in language comprehension testing. In W.P.Robinson (Ed.), *Communication in child development* (pp. 63-88). London: Academic Press.

Fritzley, V. H., & Lee, K. (2001). *Young children's response biases to yes/no questions.* Paper presented at the Society for Research in Child Development, Minneapolis, Minnesota.

Fritzley, V. H., & Lee, K. (2003). Do young children always say "Yes" to Yes/No questions? A metadevelopment study of the affirmation bias. *Child Development, 74*, 1297-1313.

Gallagher, B., & Pease, K. (2000). *Understanding the attrition of child abuse and neglect cases in the criminal justice system*: ESRC End-of-Award Report, Grant number R 000 236 891.

Garven, S., Wood, J. M., Malpass, R. S., & Shaw, J. S. (1998). More than suggestion: The effect of interviewing techniques from the McMartin preschool case. *Journal of Applied Psychology, 83*, 347-359.

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, 39*, 3-27.

Geddie, L., Fradin, S., & Beer, J. (2000). Child characteristics which impact accuracy of recall and suggestibility in preschoolers: Is age the best predictor? *Child Abuse and Neglect, 24*, 223-235.

Gee, S., Gregory, M., & Pipe, M.-E. (1999). What color is your pet dinosaur?' The impact of pre-interview training and question type on children's answers. *Legal and Criminological Psychology, 4*, 111-128.

Geiselman, R. E., Fisher, R. P., Mackinnon, D. P., & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology, 70*, 401-412.

Geiselman, R. E., Saywitz, K. J., & Bornstein, G. K. (1990). *Cognitive interviewing techniques for child witnesses and witnesses of crime* (No. SJI-BB-11J-D-016): State Justice Institute.

Geiselman, R. E., Saywitz, K. J., & Bornstein, G. K. (1993). Effects of cognitive questioning techniques on children's recall performance. In G.S.Goodman & B.L.Bottoms (Eds.), *Child victims, child witnesses: understanding and improving testimony* (pp. 71-93). New York: The Guilford Press.

Ghetti, S., Goodman, G. S., Eisen, M. L., Qin, J., & Davis, S. L. (2002). Consistency in children's reports of sexual and physical abuse. *Child Abuse and Neglect, 26*, 977-995.

Gilstrap, L. L. (2004). A missing link in suggestibility research: What is known about the behavior of field interviewers in unstructured interviews with young children. *Journal of Experimental Psychology: Applied, 10*, 13-24.

Gilstrap, L. L., & Ceci, S. J. (2005). Reconceptualizing children's suggestibility: Bidirectional and temporal properties. *Child Development, 76*, 40-50.

Gilstrap, L. L., Fritz, K., Torres, A., & Melinder, A. (2005). Child witnesses: Common ground and controversies in the scientific community. *William Mitchell Law Review, 32*, 59-79.

Gobbo, C. (2000). Assessing the effects of misinformation on children's recall: How and when makes a difference. *Applied Cognitive Psychology, 14*, 163-182.

Gobbo, C., Mega, C., & Pipe, M.-E. (2002). Does the nature of the experience influence suggestibility? A study of children's event memory. *Journal of Experimental Child Psychology, 81*, 502-530.

Goodman, G. S., Batterman-Faunce, J. M., Schaff, J. M., & Kenney, R. (2002). Nearly 4 years after the event: children's eyewitness memory and adults' perceptions of children's accuracy. *Child Abuse and Neglect, 26*, 849-884.

Goodman, G. S., Bottoms, B. L., Rudy, L., Davis, S. L., & Schwartz-Kenney, B. M. (2001). Effect 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., & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology, 12*, 1-19.

Goodman, G. S., Quas, J. A., Batterman-Faunce, J. M., Riddlesberger, M. M., & Kuhn, J. (1994). Predictors of accurate and inaccurate memories of traumatic events experienced in childhood. *Consciousness and Cognition, 3*, 269-294.

Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses. *Psychology, Public Policy and Law, 5*, 255-281.

Goodman, G. S., & Reed, R. S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior, 10*, 317-332.

Goodman, G. S., Sharma, A., Thomas, S. F., & Considine, M. G. (1995). Mother knows best: Effects of relationship status and interviewer bias on children's memory. *Journal of Experimental Child Psychology, 60*, 195-228.

Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., et al. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development, 57*, 1-163.

Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P. H., & Gordon, D. S. (2003). Why children tell: a model of children's disclosure of sexual abuse. *Child Abuse and Neglect, 27*, 525-540.

Goodnow, J. J. (1984). On being judged "intelligent". *International Journal of Psychology, 19*, 391-406.

Gordon, B. N., Baker-Ward, L., & Ornstein, P. A. (2001). Children's Testimony: a review of research on memory for past experiences. *Clinical Child and Family Psychology Review, 4*, 157-181.

Greenhoot, A. F., Ornstein, P. A., Gordon, B. N., & Baker-Ward, L. (1999). Acting out the details of a pediatric check-up: The impact of interview condition and behavioral style on children's memory reports. *Child Development, 70*, 363-380.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, Mass.: Harvard University Press.

Hardy, C. L., & Van Leeuwen, S. A. (2004). Interviewing young children: effects of probe structures and focus of rapport-building talk on the qualities of young children's eyewitness statements. *Canadian Journal of Behavioural Science, 36*, 155-165.

Hershkowitz, I. (2001). Children's responses to open-ended utterances in investigative interviews. *Legal and Criminological Psychology, 6*, 49-63.

Hershkowitz, I. (2002). The role of facilitative prompts in interviews of alleged sex abuse victims. *Legal and Criminological Psychology, 2002*, 63-71.

Hershkowitz, I., Horowitz, D., & Lamb, M. E. (2005). Trends in children's disclosure of abuse in Israel: A national study. *Child Abuse and Neglect, 29*, 1203-1214.

Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1997). The relationships among interviewer utterance, CBCA scores and the richness of children's responses. *Legal and Criminological Psychology, 2*, 169-176.

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 and Neglect, 30*, 753-769.

Holliday, R. E. (2003a). The effect of a prior cognitive interview on children's acceptance of misinformation. *Applied Cognitive Psychology, 17*, 443-457.

Holliday, R. E. (2003b). Reducing misinformation effects in children with cognitive interviews: Dissociating recollection and familiarity. *Child Development, 74*, 728-751.

Holliday, R. E., & Albon, A. J. (2004). Minimising misinformation effects in young children with cognitive interview mnemonics. *Applied Cognitive Psychology, 18*, 263-281.

Home Office. (1989). *Report of the Advisory Group on Video-Recorded Evidence (The Pigot Report)*. London: Home Office.

Home Office. (1992). *Memorandum of good practice on video recorded interviews with child witnesses for criminal proceedings*. London: Her Majesty's Stationary Office.

Home Office. (1998). *Speaking up for justice: Report of the Interdepartmental Working group on the treatment of vulnerable or intimidated witnesses in the Criminal Justice System.*

Home Office. (2001). *Achieving best evidence in criminal proceedings*. London: Communication Directorate.

Howie, P. M., & Dowd, H. J. (1996). Self-esteem and the perceived obligation to respond: Effects on children's testimony. *Legal and Criminological Psychology, 1*, 197-209.

Howie, P. M., Sheehan, M., Mojarrad, T., & Wrzesinska, M. (2004). "Undesirable" and "desirable" shifts in children's responses to repeated questions: Age differences in the effect of providing a rationale for repetition. *Applied Cognitive Psychology, 18*, 1161-1180.

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.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Jones, C., & Pipe, M.-E. (2002). How quickly do children forget events? A systematic study of children's event reports as a function of delay. *Applied Cognitive Psychology, 16*, 755-768.

Jones, C. H., & Powell, M. B. (2005). The effect of event context on children's recall of non-experienced events across multiple interviews. *Legal and Criminological Psychology, 10*, 83-101.

Kebbell, M. R., Milne, R., & Wagstaff, G. F. (1999). The cognitive interview: A survey of its forensic effectiveness. *Psychology, Crime & Law, 5*, 101-115.

Korkman, J., Santtila, P., & Sandnabba, N. K. (2006). Dynamics of verbal interaction between interviewer and child interviews with alleged victims of child sexual abuse. *Scandinavian Journal of Psychology, 47*, 109-119.

Krähenbühl, S. J., & Blades, M. (2006). The effect of question repetition within interviews on young children's eyewitness recall. *Journal of Experimental Child Psychology, 94*, 57-67.

La Rooy, D., Pipe, M.-E., & Murray, J. E. (2005). Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology, 90*, 235-254.

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., & Garretson, M. E. (2003). The effects of interviewer gender and child gender on the informativeness of alleged child sexual abuse victims in forensic interviews. *Law and Human Behavior, 27*, 157-171.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., et al. (1996). Effects of investigative utterance types of Israeli children's responses. *International Journal of Behavioral Development, 19*, 627-637.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1998). Conducting investigative interviews of alleged sexual abuse victims. *Child Abuse and Neglect, 22*, 813-823.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology, 86*, 997-1005.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., Stewart, H., & Mitchell, S. (2003a). 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. (2003b). Differences between accounts provided by witnesses and alleged victims of child sexual abuse. *Child Abuse and Neglect, 27*, 1019-1031.

Larsson, A., Granhag, P., & Spjut, E. (2003). Children's recall and the cognitive interview: Do the positive effects hold over time? *Applied Cognitive Psychology, 17*, 203-214.

Larsson, A. S., & Granhag, P. A. (2005). Interviewing children with the cognitive interview: Assessing the reliability of statements based on

observed or imagined events. *Scandinavian Journal of Psychology, 45*, 49-57.

Leichtman, M. D., & Ceci, S. J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*, 568-578.

Lindsay, D. S. (2002). Children's source monitoring. In H.L.Westcott, G.M.Davies & R.H.C.Bull (Eds.), *Children's testimony* (pp. 83-98). Sussex: John Wiley & Sons Ltd.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy and Law, 11*, 194-226.

Markham, R., Howie, P., & Hlavacek, S. (1999). Reality monitoring in auditory and visual modalities: Developmental trends and effects of cross-modal imagery. *Journal of Experimental Child Psychology, 72*, 51-70.

Marsh, E. J. (2007). Retelling is not the same as recalling. *Current Directions in Psychological Science, 16*, 16-20.

Marx, M. H., & Henderson, B. B. (1996). A Fuzzy Trace analysis of categorical inferences and instantial associations as a function of retention interval. *Cognitive Development, 11*, 551-569.

McIndoe, A., & Walsh, P. (2000). Crisis in the air. *British Medical Journal, 321*, 509-514.

Melnyk, L., & Bruck, M. (2004). Timing moderates the effects of repeated suggestive interviewing on children's eyewitness memory. *Applied Cognitive Psychology, 18*, 613-631.

Memon, A. (1999). Interviewing witnesses: The cognitive interview. In A. Memon & R.Bull (Eds.), *Handbook of the psychology of interviewing* (pp. 343-355). Chichester: John Wiley & Sons Ltd.

Memon, A., Holliday, R., & Hill, C. (2006). Pre-event stereotypes and misinformation effects in young children. *Memory, 14*, 104-114.

Memon, A., & Vartoukian, R. (1996). The effects of repeated questioning on young children's eyewitness testimony. *British Journal of Developmental Psychology, 87*, 403-415.

Merritt, K. A., Ornstein, P. A., & Spicker, B. (1994). Children's memory for a salient medical procedure: implications for testimony. *Pediatrics, 94*, 17-23.

Miller, P. H., & Bjorklund, D. F. (1998). Contemplating Fuzzy-Trace theory: The gist of it. *Journal of Experimental Child Psychology, 71*, 184-193.

Milne, R., & Bull, R. (1999). The cognitive interview. In R.Milne & R.Bull (Eds.), *Investigative interviewing: Psychology and practice* (pp. 33-53). Chichester: John Wiley & Sons Ltd.

Milne, R., & Bull, R. (2003). Does the cognitive interview help children to resist the effects of suggestive questioning? *Legal and Criminological Psychology, 8*, 21-38.

Moston, S. (1987). The suggestibility of children in interview studies. *First Language, 7*, 67-78.

Moston, S. (1990). How children interpret and respond to questions: Situational sources of suggestibility in eyewitness interviews. *Social Behaviour, 5*, 155-167.

Mulder, M. R., & Vrij, A. (1996). Explaining conversation rules to children: an intervention study to facilitate children's accurate responses. *Child Abuse and Neglect, 20*, 623-631.

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 & Law, 6*, 331-351.

Nadel, L. (1994). Multiple memory systems: What and why, an update. In D.L.Schacter & E.Tulving (Eds.), *Memory systems* (pp. 39-63). Cambridge: The MIT Press.

Ochsner, J. E., Zaragoza, M. S., & Mitchell, K. J. (1999). The accuracy and suggestibility of children's memory for neutral and criminal eyewitness events. *Legal and Criminological Psychology, 4*, 79-92.

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 and 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 and Neglect, 23*, 91-98.

Orbach, Y., & Lamb, M. E. (2000). Enhancing children's narratives in investigative interviews. *Child Abuse and Neglect, 24*, 1631-1648.

Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse and Neglect, 25*, 323-333.

Ornstein, P. A., Baker-Ward, L., Gordon, B., & Merritt, K. (1997). Children's memory for medical experiences: Implications for testimony. *Applied Cognitive Psychology, 11*, 87-104.

Ornstein, P. A., Baker-Ward, L., Gordon, B. N., Pelphrey, K. A., Staneck Tyler, C., & 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.

Perlman, N. B., Ericson, K. I., Esses, V. M., & Isaacs, B. J. (1994). The developmentally handicapped witness. Competency as a function of question format. *Law and Human Behavior, 18*, 171-187.

Perry, N. W., McAuliff, B. D., Tam, P., Claycomb, L., Dostol, C., & Flanagan, C. (1995). When lawyers question children: Is justice served? *Law and Human Behavior, 19*, 609-629.

Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review, 22*, 370-402.

Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development, 67*, 3045-3070.

Peterson, C., & Biggs, M. (1997). Interviewing children about trauma: Problems with "specific" questions. *Journal of Traumatic Stress, 10*, 279-290.

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, 33*, 118-127.

Peterson, C., Moores, L., & White, G. (2001). Recounting the same events again and again: consistency across multiple interviews. *Applied Cognitive Psychology, 15*, 353-371.

Peterson, C., Pardy, L., Tizzard-Drover, T., & Warren, K. L. (2005). When initial interviews are delayed a year: Effect on children's 2-Year recall. *Law and Human Behavior, 29*, 527-541.

Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology, 15*, 7-24.

Pezdek, K., Sperry, K., & Owens, S. M. (2007). Interviewing witnesses: The effect of forced confabulation on event memory. *Law and Human Behavior, 31*, 463-478.

Piaget, J. (1929). *The child's conception of the world*. London: Routledge & Kegan Paul.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge & Kegan Paul.

Piaget, J., & Szeminska, A. (1952). *The child's conception of number*. London: Routledge & Kegan Paul.

Pipe, M.-E., Sutherland, R., Webster, N., Jones, C., & La Rooy, D. (2004). Do early interviews affect children's long-term event recall? *Applied Cognitive Psychology, 18*, 823-839.

Plotnikoff, J., & Woolfson, R. (2004). *In their own words: The experiences of 50 young witnesses in criminal proceedings*. London: NSPCC.

Poole, D. A., & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology, 60*, 129-154.

Poole, D. A., & White, L. T. (1991). Effects of question repetition on the eyewitness testimony of children and adults. *Developmental Psychology, 27*, 975-986.

Poole, D. A., & White, L. T. (1993). Two years later: Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology, 29*, 844-853.

Powell, M. B., Fisher, R. P., & Wright, D. B. (2005). Investigative interviewing. In N.Brewer & K.Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 11-42). New York: Guildford Press.

Powell, M. B., & Roberts, K. P. (2002). The effect of repeated experience on children's suggestibility across two question types. *Applied Cognitive Psychology, 16*, 367-386.

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., & Thomson, D. M. (1996). Children's memory of an occurrence of a repeated event: Effects of age, repetition, and retention interval across three question types. *Child Development, 67*, 1988-2004.

Powell, M. B., Thomson, D. M., & Ceci, S. J. (2003). Children's memory of recurring events: is the first event always the best remembered? *Applied Cognitive Psychology, 17*, 127-146.

Pratt, C. (1990). On asking children - and adults - bizarre questions. *First Language, 10*, 167-175.

Price, H. L., & Connolly, D. A. (2004). Event frequency and children's suggestibility: A study of cued recall responses. *Applied Cognitive Psychology, 18*, 809-821.

Quas, J. A., Davis, E. L., Goodman, G. S., & Myers, J. E. B. (2007). Repeated questions, deception, and children's true and false reports of body touch. *Child Maltreatment, 12*, 60-67.

Quas, J. A., Goodman, G. S., Bidrose, S., Pipe, M.-E., & Craw, S. (1999). Emotion and memory: Children's long-term remembering, forgetting and suggestibility. *Journal of Experimental Child Psychology, 72*, 235-270.

Quas, J. A., Goodman, G. S., Ghetti, S., & Redlich, A. D. (2000). Questioning the child witness: what can we conclude from the research thus far? *Trauma, Violence, and Abuse, 1*, 223-249.

Quas, J. A., & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology, 83*, 304-338.

Regan, P. C., & Baker, S. J. (1998). The impact of child witness demeanor on perceived credibility and trial outcome in sexual abuse cases. *Journal of Family Violence, 13*, 187-195.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-Trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1-75.

Reyna, V. F., & Brainerd, C. J. (1998). Fuzzy-Trace theory and false memory: New frontiers. *Journal of Experimental Child Psychology, 71*.

Reyna, V. F., & Kiernan, B. (1994). Development of gist versus verbatim memory in sentence recognition: Effects of lexical familiarity, semantic content, encoding instructions, and retention interval. *Developmental Psychology, 30*, 178-191.

Ricci, C. M., Beal, C. R., & Dekle, D. J. (1996). The effect of parent versus unfamiliar interviewers on children's eyewitness memory and identification accuracy. *Law and Human Behavior, 20*, 483-500.

Roberts, K. P. (2000). An overview of theory and research on children's source monitoring. In K.P.Roberts & M.Blades (Eds.), *Children's source monitoring* (pp. 11-58). Mahwah: Lawrence Erlbaum Associates, Inc.

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. (1995). Children's discrimination of memories for actual and pretend actions in a hiding task. *British Journal of Developmental Psychology, 13*, 321-333.

Roberts, K. P., & Blades, M. (1998). The effects of interacting in repeated 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., & Lamb, M. E. (1999). Children's responses when interviewers distort details during investigative interviews. *Legal and Criminological Psychology, 4*, 23-31.

Roberts, K. P., & Powell, M. B. (2006). The consistency of false suggestions moderates children's reports of a single instance of a repeated event:

Predicting increases and decreases in suggestibility. *Journal of Experimental Child Psychology, 94*, 68-89.

Roebers, C. E. M., Gelhaar, T., & Schneider, W. (2004). "It's magic!" The effects of presentation modality on children's event memory, suggestibility, and confidence judgments. *Journal of Experimental Child Psychology, 87*, 320-335.

Roebers, C. E. M., & Schneider, W. (2000). The impact of misleading questions on eyewitness memory in children and adults. *Applied Cognitive Psychology, 14*, 509-526.

Roebers, C. E. M., & Schneider, W. (2002). Stability and consistency of children's event recall. *Cognitive Development, 17*, 1085-1103.

Roebers, C. E. M., & Schneider, W. (2005). The strategic regulation of children's performance and suggestibility. *Journal of Experimental Child Psychology, 91*, 21-44.

Salmon, K., & Pipe, M.-E. (2000). Recalling an event one year later: the impact of props, drawing and a prior interview. *Applied Cognitive Psychology, 14*, 99-120.

Saywitz, K. J. (1995). Improving children's testimony: The question, the answer and the environment. 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. 113-140). Thousand Oaks: Sage Publications.

Saywitz, K. J., Jaenicke, C., & Camparo, L. (1990). Children's knowledge of legal terminology. *Law and Human Behavior, 14*, 523-535.

Saywitz, K. J., & Snyder, L. (1993). Improving children's testimony with preparation. In G.S.Goodman & B.L.Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 117-146). New York: The Guildford Press.

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.

Schwarz, S., & Roebers, C. E. M. (2006). Age differences in the effects of social influence on children's eyewitness performance and their metacognitive monitoring. *Journal of Experimental Child Psychology, 94*, 229-248.

Seidler, K. M., & Howie, P. M. (1999). Motivational factors in children's reporting of events: The influence of age and expected reinforcement contingency. *Journal of Applied Developmental Psychology, 20*, 101-118.

Shapiro, L. R., Blackford, C., & Chen, C.-F. (2005). Eyewitness memory for a simulated misdemeanor crime: The rôle of age and temparament in suggestibility. *Applied Cognitive Psychology, 19*, 267-289.

Shapiro, L. R., & Purdy, T. L. (2005). Suggestibility and source monitoring errors: Blame the interview style, interviewer consistency, and the child's personality. *Applied Cognitive Psychology, 19*, 489-506.

Shrimpton, S., Oates, K., & Hayes, S. (1998). Children's memory of events: Effects of stress, age, time delay and location of interview. *Applied Cognitive Psychology, 12*, 133-143.

Siegal, M. (1991). *Knowing children: Experiments in conversation and cognition*. Hove: Lawrence Erlbaum Associates Ltd.

Siegal, M. (1999). Language and thought: the fundamental significance of conversational awareness for cognitive development. *Developmental Science, 2*, 1-14.

Siegal, M., Waters, L. J., & Dinwiddy, L. S. (1988). Misleading children: Causal attributions for inconsistency under repeated questioning. *Journal of Experimental Child Psychology, 45*, 438-456.

Sjöberg, R. L., & Lindblad, F. (2002). Delayed disclosure and disrupted communication during forensic investigation of child sexual abuse: a study of 47 corroborated cases. *Acta Paediatrica, 91*, 1391-1396.

Smith, D. W., Letourneau, E. J., Saunders, B. E., Kilpatrick, D. G., Resnick, H. S., & Best, C. L. (2000). Delay in disclosure of childhood rape: Results from a national survey. *Child Abuse and Neglect, 24*, 273-287.

State of Michigan Governor's Task Force on Children's Justice, & Department of Human Services. (1998). *Forensic interviewing protocol.* Michigan: FIA-779.

Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135-154). Dordrecht, Netherlands: Kluwer Academic Publications.

Steller, M., & Boychuk, T. (1992). Children as witnesses in sexual abuse cases: Investigative interview and assessment techniques. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 47-72). Chichester: John Wiley and Sons.

Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). The Memorandum of Good Practice: Theory versus application. *Child Abuse and Neglect, 25*, 669-681.

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., & Sunshine, N. (1996). The relation between investigative utterance type 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., et al. (1997). Effects of introductory style on children's abilities to describe experiences of sexual abuse. *Child Abuse and Neglect, 21*, 1133-1146.

Steward, M. S., Bussey, K., Goodman, G. S., & Saywitz, K. J. (1993). Implications of developmental research for interviewing children. *Child Abuse and Neglect, 17*, 25-37.

Steward, M. S., Steward, D. S., Farquhar, L., Myers, J. E. B., Reinhart, M., Welker, J. J. N., et al. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development, 57*, 1-232.

Sussman, A. L. (2001). Reality monitoring of performed and imagined interactive events: developmental and contextual effects. *Journal of Experimental Child Psychology, 79*, 115-138.

Sutherland, R., & Hayne, H. (2001). Age-related changes in the misinformation effect. *Journal of Experimental Child Psychology, 79*, 388-404.

Thierry, K. L., & Spence, M. J. (2004). A real-life event enhances the accuracy of preschoolers' recall. *Applied Cognitive Psychology, 18*, 297-309.

Thomsen, Y., & Berntsen, D. (2005). Knowing that I didn't know: Preschoolers' understanding of their own false belief is a predictor of assents to fictitious events. *Applied Cognitive Psychology, 19*, 507-527.

Thoresen, C., Lønnum, K., Melinder, A., Stridbeck, U., & Magnussen, S. (2006). Theory and practice in interviewing young children: A study of Norwegian police interviews 1985-2002. *Psychology, Crime & Law, 12*, 629-640.

Tizzard-Drover, T., & Peterson, C. (2004). The influence of an early interview on long-term recall: A comparative analysis. *Applied Cognitive Psychology, 18*, 727-743.

Tobey, A. E., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse and Neglect, 16*, 779-796.

Toth, S. L., & Cicchetti, D. (1998). Remembering, forgetting, and the effects of trauma on memory: A developmental psychopathology perspective. *Development and Psychopathology, 10*, 589-605.

United Nations. (1989). Convention on the rights of the child: Office of the High Commissioner for Human Rights.

Vrij, A. (2002). Deception in children: A literature review and implications for children's testimony. In H.L.Westcott, G.M.Davies & R.H.C.Bull (Eds.), *Children's testimony* (pp. 175-194). Chichester: John Wiley & Sons Ltd.

Warren, A., Hulse-Trotter, K., & Tubbs, E. C. (1991). Inducing resistance to suggestibility in children. *Law and Human Behavior, 15*, 273-285.

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.

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. (2001). Interviewing children and adults: The effect of question format on the tendency to speculate. *Applied Cognitive Psychology, 15*, 521-531.

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.

Welch-Ross, M. K. (1995). Developmental changes in preschoolers' ability to distinguish between memories of performed, pretended, and imagined actions. *Cognitive Development, 10*, 421-441.

Welch-Ross, M. K. (1999). Interviewer knowledge and preschoolers' reasoning about knowledge states moderate suggestibility. *Cognitive Development, 14*, 423-442.

Westcott, H., & Kynan, S. (2006). Interviewer practice in investigative interviews for suspected child sexual abuse. *Psychology, Crime & Law, 12*, 367-382.

Westcott, H., Kynan, S., & Few, C. (2006). Improving the quality of investigative interviews for suspected child abuse: A case study. *Psychology, Crime & Law, 12*, 77-96.

Westcott, H. L., & Jones, J. (Eds.). (1997). *Perspectives on the Memorandum.* Hants, England: Arena.

Working together. (1988; revised edition 1991). HMSO.

Wyllie, S., & Brown, K. (2003). *A flea in the ear.* London: Anderson Press Ltd.

Youth Justice and Criminal Evidence Act, (1999).

Zajac, R., Gross, J., & Hayne, H. (2003). Asked and answered: Questioning children in the courtroom. *Psychiatry, Psychology and Law, 10*, 199-209.
Zajac, R., & Hayne, H. (2006). The negative effect of cross-examination style questioning on children's accuracy: Older children are not immune. *Applied Cognitive Psychology, 20*, 3-16.

Appendix 1

The script

Hello children, my name is Chris, I am very pleased to be able to be here today and hope that you will learn a little more about looking after dogs. How many of you have got a dog at home? Well, I am going to tell you about my dog Susie. She is (Chris sniffs). Here is (Chris sniffs again) - oh excuse me, I have got a bit of a cold, I'll just get a tissue (Chris reaches into her handbag, takes out a tissue, uses it and replaces it in the bag). Here is a photograph of Susie (shows photograph). Susie's a smooth haired lurcher and she is a very fast runner. A lurcher is one particular breed of dog. As you can see from this poster (holds up the poster) there are many different breeds of dog. They come in all shapes and sizes, some are really hairy like this one and some have got really grumpy faces like the boxer.

But it doesn't matter whether your dog is small, big, hairy or grumpy; all dogs need the same sort of looking after. What I am going to tell you now is important for all dogs.

All dogs need food and fresh water everyday. I go to the supermarket to buy food for Susie. It comes in tins like this (shows can of food), this is chicken in jelly, mmm sounds yummy. Or else you can get dried food like this (shows bag of food), this is moist meal with beef, carrots and cereals. You get to know how much food your dog will need. You must be careful not too give your dog too much food, or too little. Too much and it might look more like an elephant and too little and it might disappear completely!

Now, dogs also need to have exercise, this is very important especially for a dog like Susie – just like you people dogs need to have playtime or break, well dogs need that too. If you are lucky enough to have a big garden then your dog can run around there, but if not, then your dog has to go for a walk.

(At this point Peter enters in a slightly agitated manner and goes straight over to Chris, says "you left the car in the wrong place, where are the keys" and, without waiting for an answer, starts rummaging in Chris's handbag to get the keys. Chris needs to respond as though she is a bit peeved but not that much –

this is the ambiguous part. Peter finds the keys then goes out as if in rather a hurry but not saying anything else.)

Right, now, where was I, oh yes...I take Susie out for a walk 3 times a day which keeps us both fit! When you take your dog for a walk you need to take a lead – this purple lead is Susie's, look how it works (here Chris selects a child to help demonstrate pulling out a portion of the lead). You hold this and pretend to be Susie, you haven't got a collar on have you. If we go for a walk Susie can have a long lead like this (demonstrates with a child). Thank you. Or a short lead (demonstrates again), when we have not got much room. Susie has to go on her lead because she cannot cross the road safely, she might otherwise get lost and although she is a very friendly dog not all people and other dogs like her so I sometimes have to be able to keep her near to me.

The other reason that dogs need to go out is so that they can go to the toilet. I expect that you all know how horrid it is to tread in dog mess but this wouldn't happen if people looked after their dogs properly. When your dog has been to the toilet you need to clear up the mess. You can use one of these (show the "pooper scooper") to pick up the mess and then pull over the bag, tie a knot in it and put that in the rubbish bin. Of course, when you get home you should wash your hands well. Actually, you should always wash your hands after touching your pets, dogs included.

Now, if a dog is not well then it needs to go to see the vet, which is a doctor for animals. Susie got poorly once, do you know what she did, she swallowed a fishhook once. She had to have an x-ray and the vet said to give her cotton wool sandwiches to get rid of the sharp hook, now she's fine.

And lastly, the best thing to do to look after your dog is to give it lots of love and attention; be its best friend.

I have brought a story to read to you, which I think you will enjoy; it's quite funny and shows how you can use your brains to sort out problems.

(Read the book "A Flea in the Ear" showing the picture after each page of text is read.)

Thank you for listening so well. If you have time you could draw a picture of your dog for me, maybe your teacher could send it to me.

Teacher – (whilst Chris starts to pack up) Thank you for coming to talk to us Chris.

The Teacher then speaks to the children so as to reduce their opportunity to ask Chris any questions.

Appendix 2

Police training

"Incompetency comes down to the person asking the questions." Lin Cross – course leader, National Consultant Trainer in Child and Vulnerable Adult abuse, course leader of a MAIVIC (Multi-Agency Investigation and Video Interview Course) at Staffordshire Police Headquarters.

The training course consisted of 3 modules: 1. Generic issues (4 days), 2. Investigations with children (4 days), and 3. Investigations with vulnerable adults (4 days). The following lists of objectives are included through kind permission of Jenny Blewitt and Lin Cross (training course leaders).

The objectives for module 1 are to: identify vulnerable persons, outline the principles of multi-agency investigations, explain communication issues, recognise categories of abuse, plan and prepare an investigative interview, identify relevant documentation, outline the Four Phase Approach to interviewing, recognise issues surrounding medical examinations, describe issues relating to Pre-trial therapy for witnesses and outline responsibilities of those involved in conducting a video interview.

The objectives for module 2 are to: recognise issues surrounding language development in children, describe appropriate circumstances in which to conduct a video interview, outline specific areas for consideration when communicating with disabled children, recognise the issues concerning Sex Offenders in relation to the interviewing process, conduct an Investigative Interview to a recognised standard, summarise the areas of concern for a non-abusing carer, outline the powers to protect children, describe the support

mechanism for witnesses attending court and explain the principles of a Child Protection conference.

Appendix 3

Semi-structured interviews with practicing police officers

Three police officers were asked a series of questions (in 2006) to examine their perceptions of the interviewing process in general and their awareness of the use of question repetition. All 3 officers had been trained in either 2004 or 2005, had conducted between 3 and 5 interviews with children aged between 4 and 16 years per month. None of the officers had received any further training after the initial training course nor had any been previously asked to discuss their interviewing practices.

The officers volunteered to be interviewed. The responses to questions were not audio taped but were recorded in note form. The summary below is based on those notes, copies of which were sent to the officers concerned whose consent was obtained for the reproduction of the summary in this thesis.

Interview questions and results

1) Could you outline the easier, and more difficult factors involved in interviewing?

An interview is easier to conduct when the interviewing officer has personal knowledge of the child and the alleged incident. Older and more mature children are generally easier to talk with as they understand what is required of them and what is meant by the questions asked. With younger children there is a problem with communicating in a manner that keeps their attention and yet covers the information required. The ground rules are too long and awkward (they sound false) although they are easier to establish with older children. A balance is needed between eliciting the level of detail necessary for clarity without prompting or leading the child in any way.

One of the most difficult and frustrating aspects of interviewing is when the rapport between the interviewer and the child does not develop despite every effort made. It was emphasised that rapport is necessary with all interviewees but how this is done, and how long it takes, depends on the child's age and level of confidence.

A positive aspect is the immediate job satisfaction experienced when the interview goes well. Interviewing was described as a situation where one can learn from others in order to become more skilled.

2) What would you consider a highly successful interview, and vice versa, a really disappointing interview?

A highly successful interview is one when the relationship is such that the child discloses relevant information, which will lead to a successful prosecution. This result also implies that the child has understood what is needed and is able to give a lot of detail. For an interview to be successful the rapport will need to be successful.

A disappointing interview is when the interviewers are aware of asking questions which lead the child to "bottle-up" which is then very difficult to redress. There is frustration at being unable to engage some children, particularly younger ones.

3) How aware are you of the type of questions you ask, which do you find more effective in eliciting appropriate information?

The interviewers were "extremely aware" of the questions used, it was described as crucial to "Engage brain before you speak". It is important not to overwhelm the child with questions; it is usually more successful to have as natural a conversation as possible. The interviewers were aware of the balance

between having to "pull out" the information and at the same time being aware that it is not acceptable if this is done too much – this may lead to problems with the CPS or defence.

The 5 'wh-' questions are kept in mind; the questions asked will be framed round those questions. The questions used will be age-related, younger children will need much simpler explanations and language. Closed questions will tend to be used more frequently with younger children who "may not really understand what is expected of them". Closed questions may help these younger children to understand the gist of what information is required. The interviewers were wary of using direct or leading questions.

4) Do you repeat questions in interviews? If so in what stages, how often, what form do they take and why do you use them?

The interviewers were aware of repeating questions but tried to avoid them or keep them to a minimum. Interviewers estimated that repetitions accounted for approximately 5% of the questions asked. Repetition is appropriate in order to check whether the child has understood the question or to clarify a particular point; in these situations the interviewer will explain why the repetition takes place. The most common form of repetition is in gist form but there is no point in trying this technique more than a couple of times, after that a direct yes/no question will be appropriate.

Repetitions can be used to elicit specific or further information if the child does not answer the initial question or does not answer appropriately. However, you "can't push it too far as you can't force them into saying something they don't want to". Care must be taken not to repeat questions to the extent that it can be perceived as "intimidating the child".

5) Have you any concerns over the use of repetition? If so, what are they and how do you address these concerns?

Question repetition is useful in eliciting information but should be used with caution because too much can "numb" the child towards the interviewer, could confuse the child or give the impression that you do not believe them. The interviewers would not want to jeopardise a case by appearing intimidating. Frequent use of repetition might also be perceived as oppressive. If rephrasing or qualifying the question does not succeed in eliciting the information required then the interviewers will have to accept responses such as "I don't know" or "can't remember" from the child.

6) Is question repetition always appropriate?

The reasons for using repetition depend on the age of the child being interviewed. Older children may not answer a question through choice rather than through a lack of understanding. It is therefore possible to stay on the same subject longer with a younger child but necessary to move forward quicker with an older child. It is possible to justify more repetitions with younger children because they might not understand what you expect from them. It is also appropriate to repeat a question when the interviewer has reason to expect (from other information/evidence) that the child does know the answer and could respond if encouraged to do so.