

A Study Space Analysis for Multiple Interviewing of Child Witnesses

Genevieve F. Waterhouse¹

Anne M. Ridley²

Ray Bull³

Rachel Wilcock¹

Author Note

¹Department of Psychology, University of Winchester, Winchester, UK; ²Department of Psychology, London South Bank University, London, UK; ³Department of Criminology and Law, University of Derby, Derby, UK.

Correspondence should be addressed to Genevieve F. Waterhouse, Department of Psychology, Faculty of Humanities and Social Sciences, University of Winchester, Sparkford Road, Winchester, UK. SO22 4NR. Email: genevieve.waterhouse@winchester.ac.uk
Telephone Number: 01962 675147

Suggested Running Head: A SSA for Multiple Interviewing of Child Witnesses

Manuscript Word Length: 4,517

A Study Space Analysis for Multiple Interviewing of Child Witnesses

Abstract

This article presents a Study Space Analysis of 44 published research studies examining the use of multiple interviews with child victims/witnesses. Study Space Analysis is a method of detecting gaps in the existing literature and thus determining whether ecologically valid situations that arise in practice have actually been addressed and studied. The use of this methodology is particularly useful for techniques which are being considered for changes in policy or practice, ensuring that the literature is sufficient to warrant change. Multiple or repeated interviewing has been argued by some authors to be ready for change. However, in the present Study Space Analysis, it is concluded that despite a growing literature, there are still some key variables which require research examination prior to policy change. In particular, research is needed involving samples of children of between 11 and 18 years old, participants with multiple needs, and interviews regarding repeated events. Findings from the supplementary online material also identify the need for studies with longer delays between the event and the initial interview, more than two interviews of a child, phased multiple interviews and interviews conducted by professional interviewers.

Practitioner Messages:

- Some researchers suggest interviewing a child more than once about an event can be beneficial for police investigations as it leads to additional accurate details.
- This study concludes there are key variables yet to be studied under controlled conditions and thus more research is required before comprehensive policy change is recommended.
- Future research is required with adolescent participants, children with multiple needs, and interviews about repeated events.

Keywords: Study Space Analysis; child victims; investigative interviewing; multiple interviewing; police.

A Study Space Analysis for Multiple Interviewing of Child Witnesses

Introduction

Multiple interviewing consists of when a witness/victim is interviewed more than once about the same event(s). Guidelines worldwide discourage using multiple interviews (e.g., England and Wales' 'Achieving Best Evidence in Criminal Proceedings: Guidance on interviewing victims and witnesses, and guidance on using special measures', henceforth ABE, Ministry of Justice, 2011; New Zealand Law Commission, 1996; Scottish Executive, 2011). In some countries, guidelines suggest multiple interviews may only be considered in certain circumstances. For example, ABE guidelines stipulate multiple interviews to be considered only when there is not enough time to fully discuss the event(s) in one interview; when other sources (including the perpetrator) reveal new information that needs discussion with the child; or when the child tells someone they have further information to give the police (Ministry of Justice, 2011).

Multiple interviewing of child witnesses is an area ripe for policy and practice change, or at least expansion. La Rooy *et al.* (2010) argue that the robust literature on reminiscence (when a person recalls new information during a second recall attempt) warrants a change in guidelines encouraging the use of multiple interviews in a broader range of circumstances. Furthermore, Block *et al.* (2013) conducted a cost-effectiveness analysis. They concluded that the economic benefits of additional convictions and reduced victimisation related to increased disclosure from routine second interviews with suspected child victims of sexual abuse outweighed the additional resources required.

La Rooy *et al.* (2010) state that multiple interviews should be used to obtain further information with the limitation that these interviews should be conducted according to best practice (as should all investigative interviews). Although this statement was supported by a brief review of the literature and four case examples, it follows on from La Rooy, *et al.*'s

(2009) more extensive qualitative review. This narrative review of the literature on multiple interviews of children aged 2 to 13 identified that children generally gave less accurate information in second and subsequent interviews than in the first, but that this additional information can still be relevant, accurate, and vital for prosecutions (such as in the case examples in La Rooy *et al.*, 2010). However, although this review concluded further research is required for a full understanding of the benefits and risks of multiple interviewing, it did not systematically examine the research to determine where significant gaps in knowledge remain. The present analysis aims to determine whether there are sufficient studies using ecologically-valid methods to support La Rooy *et al.*'s (2010) recommendations and whether the research is sufficient to extend them to all children, including 13 to 18 year olds.

Study Space Analysis (henceforth SSA) is a way of amalgamating and evaluating published research on a subject and thus determining whether the research is sufficiently applicable to practice to warrant changes. Unlike meta-analysis, SSA does not look at whether a technique has a statistically robust effect on outcomes, or indeed the results of the studies at all, but looks at the topics the current research has covered, the breadth of these topics, and their relation to the associated field of practice (Malpass *et al.*, 2008). The benefits of this procedure include detecting whether key variables or conditions have been explored and thus determining whether a topic has sufficiently progressed to warrant evidence-based policy changes.

Policy changes should be based on high quality, methodologically rigorous studies that address diverse variables related to both theory and changing ecological conditions (Malpass *et al.*, 2008). For example, when examining the effectiveness of a new investigative interviewing technique, it is important to compare the new technique to current practice and other similar evidence-based improvements. Furthermore, it is important to determine whether the technique compares favourably for an array of participants; varying in

age, participation in the to-be-remembered event (e.g., participant or observer, victim or witness), and gender, among other factors. If an interviewing technique were brought into practice, it is likely to be used with a variety of witnesses and victims in different situations (for example, different lengths of delay between crime and interview). Therefore, it should have been empirically tested with a similarly broad group of mock-interviewees under varying conditions. Policy decisions made on a literature which omits some of these participants or conditions will be based on incomplete understanding and may limit the beneficial effects the technique could have, or even result in less well-conducted interviews for those unstudied participants/conditions.

SSA reveals gaps in the literature by creating a merged visual representation of all of the relevant studies, their independent, dependent, and cross-study variables (methodological factors which vary between studies but not within the study, such as whether the event was live or video), and their relationships (Malpass *et al.*, 2008). Matrices are created with the frequencies of each individual variable plotted against the other variables. Areas of the matrices with low or null frequency counts demonstrate a lack of research.

This methodology has been used for a number of topics within investigative psychology. Malpass *et al.* (2008) included exemplary SSAs for eyewitness identification line-ups and for alcohol and eyewitness memory. Memon *et al.* (2010) conducted a SSA on the Cognitive Interview and more recently, Brubacher *et al.* (2015) completed another for ground rules in child interviews. These SSAs revealed strengths and weaknesses within the research. For example, Memon *et al.*'s (2010) SSA highlighted the under-representation of non-student populations in Cognitive Interviewing research, as well as the reliance on filmed to-be-remembered events (henceforth TBR). Thus, the SSA methodology has previously been used to good effect in this field.

The present analysis will apply this methodology to multiple interviewing of child victims/witnesses. The SSA will include studies that have examined multiple interviewing in the laboratory setting within broad parameters (see the methods section for inclusion criteria). However, it will then evaluate these studies on their application to child abuse investigations specifically. In particular, this SSA will determine if the literature sufficiently examines motivational reasons for multiple interviewing, namely *why* multiple interviews should be conducted in child abuse investigations. It will also examine the representativeness of the samples in the existing studies (i.e., age, developmental differences) and the ecological validity of the methodologies used (e.g., event variables such as length, repetition, participation). Thus, this SSA will aid policy makers to determine whether the expansion of the use of multiple interviews to circumstances not currently described in interviewing guidelines is warranted.

Method

Studies

The studies included were obtained via online searches of the PsycINFO and PsycARTICLES databases. Searches using the terms ‘child’, ‘interview’, ‘memory’, ‘multiple’, ‘repeat’, ‘twice’, ‘three’, or ‘four’ were conducted, along with searches in which ‘child’ was replaced with ‘adolescent’ and ‘teenage’. Additionally, relevant publications from the reference lists of key multiple interviewing publications (including La Rooy *et al.*, 2009) were obtained. Published studies from peer-reviewed journals were included if they:

1. Involved more than one interview of a participant recalling information about the same event.
2. Included a child sample (i.e., some participants were under 18 years of age).

3. Compared children's responses in multiple interviews; either through direct statistical comparisons, or by comparing the information provided in different interviews (e.g., coding children's recall as repeated or novel).
4. Had a minimum sample size of 40 participants.
5. Included ecologically-valid interviews in terms of:
 - a. Including some free recall of the TBR event.
 - b. Involving face-to-face recall (e.g., not completed via telephone or written).
 - c. Attempting to replicate multiple investigative interviews rather than cross-examinations.
 - d. Not aiming to create false memories. Studies that included some misleading questions without this specific aim were, however, included.
6. Did not analyse field interviews. Studies that examined real forensic interviews of children were excluded because children's accuracy often cannot be determined. Additionally, the interviews are not standardised. Thus, various confounding interview variables may have affected the results, including differing interview quality.
7. Were published in English.

These criteria were chosen in order to include only studies of relatively high ecological validity, with a reasonable sample size, and high levels of control over confounding variables.

No time frame was specified for the search and so all relevant papers published before August 2018 were included.

The initial literature search revealed 149 articles. After screening for relevance via abstracts available electronically (including examination of the full text where necessary), 105 were removed from the analysis for not meeting the criteria. This left 44 published research articles. All articles were available electronically. One article included two

appropriate experiments, another included three, and four experiments were extensions of other experiments included in the analysis (i.e., studies that re-interviewed the same sample, or re-coded and analysed the data from another experiment). Thus, 43 independent samples were included in the analysis.

Procedure

For the studies, their independent (IV), dependent (DV) and cross-study variables (CV) were identified. Separate matrices were created for each of these types of variable and all the appropriate variables for the relevant studies were listed in each matrix (e.g., all 43 samples' independent variables in one, their dependent in the next, and their cross-study variables in the third). The independent variables were listed along the top of each matrix, and frequency counts were entered for each independent variable against its corresponding independent, dependent and cross-study variables (i.e., IV x IV, DV x IV, CV x IV).

For studies which were extensions of previous experiments, only the new aspects of the experiment were included. For example, Pipe *et al.* (1999; Experiment 1) was an extension of Pipe and Wilson's (1994) work. Pipe and Wilson's (1994) study was entered into the matrices using the standard procedure, with age of child, participant involvement in the event, and type of interview as three independent variables. The dependent variables included the number of correct details and errors provided in free recall for the first and second interviews. Children's responses to specific questions and their overall accuracy during free recall were also measured as dependent variables. The cross-study variables related to the type of TBR event used (e.g., whether it was staged, a life experience or a video), the type of interviews (e.g., question types and the experience of the interviewer), and the timing of the interviews. The new aspects of Pipe *et al.*'s (1999) first experiment were then added. The main extension of the study involved a further interview of the sample.

However, this was only conducted for the younger age group and participation in the event was no longer considered as an independent variable. Thus, the new dependent variables were only added for the 'type of interview' independent variable and not age or event participation. Pipe *et al.* (1999) also re-analysed the prior interviews and so the frequencies for proportion of new accurate and new repeated details for the second and third interviews were increased by one. Therefore, Pipe *et al.*'s (1999) study was not treated as a separate study but as a continuation, and so the factors previously examined were not repeated within the SSA, only new variables added in.

Results

The Independent Variables

The SSA identified a wide array of independent variables. The majority of the 35 categories included independent variables that were only examined in one or two studies (for example, the use of social support in multiple interviews, or whether the TBR event was conducted by an acquaintance or a stranger, Goodman *et al.*, 1991; Lepore & Sescio, 1994 respectively). However, some of the independent variables were included more frequently. In particular, age (e.g., Gobbo *et al.*, 2002; all of Peterson and colleagues' studies), initial retention interval (i.e., time between the TBR event and the first interview; Gross & Hayne, 1999; Pipe *et al.*, 2004; Powell & Thomson, 1997; Salmon & Pipe, 2000), number of interviews the child experienced (particularly whether experiencing an intervening interview between two interviews affected memory in the last interview, e.g., Baker-Ward *et al.*, 1990; Ornstein *et al.*, 2006; Peterson, 1999), delay between the first two interviews (e.g., Baker-Ward *et al.*, 1993; Ornstein *et al.*, 1992; Powell & Thomson, 1997) and whether the child participated directly in the TBR event or merely observed it (Baker-Ward *et al.*, 1990; Gobbo *et al.*, 2002; Pipe & Wilson, 1994). Most studies included more than one independent

variable and often manipulated the types of interviews experienced (such as including human body diagrams, or suggestive questions, e.g., Brown *et al.*, 2012; Bjorklund *et al.*, 2000 respectively). Multiple interviews with child interviewees have, therefore, been studied under a variety of interviewing conditions. However, in the majority, no more than three studies of each condition have been conducted.

Sample Representativeness

The following tables show the number of studies that have included certain design features and the age group of the participants involved in that study. Some studies included an age group that spanned more than one age range, and in some cases different ages constituted the independent variable. Thus, the age groups include as many of the contrasting age groups as possible so the following tables present every age group in all of the experiments and every variable they experienced.

No studies of the effects of multiple interviews included a sample of children aged 14 to 18 years (see Table 1). Only five independent samples included participants aged 11 to 13 years. The most studied age group was children between five and eight-years-old, closely followed by children aged three to five. Furthermore, the vast majority of children studied in the research included here were typically-developing children (just two studies included children with intellectual disabilities; Brown *et al.*, 2015; Henry & Gudjonsson, 2003).

Ecological Validity

It could be argued that multiple interviewing is of greatest value to child abuse investigations. These cases often do not progress to court (National Society for the Prevention of Cruelty to Children, 2014), have serious psychological and social consequences for children (Norman *et al.*, 2012; Tyler, 2002), and rely heavily on children's testimony as

the only source of information (other than the perpetrator; Malloy *et al.*, 2011). Thus, there are strong arguments for the allocation of monetary and time resources for such cases, especially for techniques that elicit further evidential leads. Hence, the most appropriate TBR events for experimental studies would be those replicating aspects of child abuse cases, within ethical boundaries. When making this comparison, this SSA reveals areas where the ecological validity of studies could be improved.

As can be seen from Table 1, the majority of studies used a live interaction as their TBR event. These were staged for the experiment itself (e.g., visits from pirates or magic shows; Jack *et al.*, 2012, La Rooy *et al.*, 2005), thus replicating a situation where the child is the sole source of information about a private child-adult event. Alternatively, they were naturally occurring events (e.g., Peterson and colleagues' studies, 1996, 2005, 2010). Both these types of event had the benefit of the child actually participating, rather than watching or hearing about the event, however, the latter studies were of higher ecological validity. For example, a significant sub-sample used children's real visits to the doctor or medical emergencies as their TBR event. These incidents could be argued to have numerous similarities with child abuse; they may involve negative emotions (including pain and/or stress), the child is often touched by an adult, sometimes in intimate areas, the child is directly involved in the event, and it can last a significant amount of time. The disadvantage of using this type of TBR event is the lack of control over it; children's experiences differ which may cause differences in their recall due to the event itself and not the interviewing techniques. Also, assessing the accuracy of recall is problematic. Furthermore, these real-life events were not standardised in length of time, and so some may have been long (such as experiences of a hurricane) and some quite short (such as one-off dentist visits), which may in turn affect memory strength. Experiments with younger age group samples (2 to 5 years), in particular, very rarely involved interviews about events of a known, standardised length.

Further, when studies did use standardised length events, these were mainly very short, which may be unlikely to replicate abuse (e.g., 18.2% of the known standardised length TBR events lasted less than two minutes). On the other hand, 33.3% of known-length events were 30 to 60 minutes.

Only five experiments involved repeated events (Jack *et al.*, 2012; Powell & Thomson, 1997; Price *et al.*, 2016, Experiments 1 to 3). Repeated abuse occurs frequently, in around a quarter to a third of cases in some studies (Bottoms *et al.*, 2007; Goodman *et al.*, 1992). Thus, understanding the impact of multiple interviewing in these situations could be especially beneficial.

Very few of the studies included in this SSA used a crime as the TBR event. When crimes were used, they were presented in video format and thus children did not participate in the event. However, this is likely to be due to the ethical issues related to a child witnessing a live crime event.

Examining the last section of Table 1, many authors who used a more standardised event also made attempts to include forensically-appropriate aspects in their events. This included unaccompanied interaction with an adult, touch (including, in some doctor visits, touch of the genital area) or a photograph being taken of the child, and some involved removing clothes from a toy. However, a large number of studies did not include any abuse-related aspects and only one study with one age group involved a TBR event that the child was asked to keep secret.

[Table 1 placed here]

Why re-interview?

An excerpt of the DV x CV matrix is presented in Table 2. Most studies included more than one dependent variable in their design. For Table 2, the dependent variables were

defined broadly; the variables include general measures of each variable, as well as measures that were specified by the type of question asked or by the topic of the detail. For example, the dependent variable ‘Number of Details’ includes measures of the total number of details provided, but also measures of the number of details provided in response to open questions only. Each study, however, was only coded once for each cell even if it measured more than one form of this variable (e.g., total number of details and number of details about people).

Table 2 shows that many of the studies in the SSA examined accuracy of the child’s recall in some form. This included calculating number of correct and incorrect details as well as percentage accuracy. Accuracy of children’s recall was often assessed for first, second, and third interviews (when conducted) in a study.

Additionally, the accuracy and consistency of children’s recall across interviews was often examined. This involved separate measures of the accuracy of new and repeated details.

Some studies coded children’s recall according to the centrality of the details (e.g., central vs. non-central; Bjorklund *et al.*, 1998). This can be helpful for determining the extent to which new, accurate information is useful for an investigation. Other studies classified the details children gave chronologically or according to topic types. For example, in Fivush *et al.* (2004), children’s recall of Hurricane Andrew was divided into preparation for the storm, the storm itself, and its aftermath. Thus, some studies have examined the type of information being recalled across multiple interviews, which may be relevant for determining how useful this information could be to an investigation.

[Table 2 placed here]

Additional Results

Further results regarding how well the research addresses who should re-interview, what delay there should be between interviews, and how multiple interviews should be conducted are available in the additional online resources. The question of who should conduct multiple interviews was not well answered by the literature and identified some particular gaps. Only one study compared multiple interviews using the same interviewer with those using different interviewers; only three of the 43 independent samples used professional interviewers; and only one of these compared police interviewers with clinicians. The length of time between the incident and the first interview and between the first and subsequent interviews varied widely and only four studies examined the effect of this variable. The majority of studies included only one repeat interview, limiting any interpretation of the optimal number of follow-up interviews. Furthermore, the vast majority of second interviews involved children answering questions on the same events again. This does not replicate some current practice where different phases of the interview process are conducted in separate interviews or where children are interviewed a second time about a new topic within the same event. These findings, full details of which are given in the supplementary material, suggest the need for research with longer delays between the event and the initial interview, more than two interviews of a child, phased multiple interviews and interviews conducted by professional interviewers.

Discussion

This research has used a Study Space Analysis to identify the topics addressed by research on multiple interviews with children and where additional research is required. In particular, it has found that a variety of interviewing conditions have been examined, but very rarely do more than one or two studies examine the same independent variables, meaning

there is little replication in the field. Additionally, there are some key populations missing from the literature; namely, adolescent samples (12 to 17 year olds) and children with learning difficulties or additional needs (such as ASD). This SSA has also found attempts to produce ecologically-valid research methodologies. However, there are benefits and limitations to the options available within ethical parameters and research examining repeated events is lacking. Finally, the SSA determined that the majority of research examining multiple interviews with children included dependent variables which relate to the possible benefits of conducting multiple interviews (i.e., an increased total recall of the event).

Research shows that adolescents and children with multiple needs are over-represented in the Criminal Justice System. For example, there is evidence that most multiple interviews are conducted with adolescents (i.e., 12 to 17 year olds according to Waterhouse's [2016] findings). Although studies have been conducted with adults and young children, we should not assume that if multiple interviewing is effective for obtaining evidence from these age ranges, it will be for adolescents too. Furthermore, research has suggested that abused children often experience cognitive deficits (Mills *et al.*, 2011) and that children with learning difficulties and other conditions may be more likely to be victimised and thus be involved in legal cases (Sullivan & Knutson, 2000). The under-representation of these types of children present a significant flaw in the representativeness of the multiple interviewing research so far.

In terms of the ecological validity of the research covered in the SSA, most studies had incorporated elements into their methodologies that replicate child abuse-related experiences. These include the event being a private interaction with an adult, touching, and negative emotions. However, for ethical reasons, the ecological validity of studies with experimenter-created TBR events is limited. On the other hand, the lack of control over naturally-occurring TBR events introduces the possibility of confounding variables. Thus,

the combination of the two types of study may be necessary for a developed understanding of the likely impact of multiple interviews on child victim/witness recall. There are also some key characteristics of child abuse which are nearly entirely missing from the multiple interviewing literature. In particular, the lack of research on repeated TBR events and events which a child has been asked to keep secret is a concern for the validity and applicability of the research findings so far.

Conversely, the dependent variables that have been employed in the research included in this SSA adequately address the question of why multiple interviews should be conducted. The main potential benefit of multiple interviews is obtaining further accurate information about an alleged crime. Therefore, examining the total number of details provided in subsequent interviews in relation to their accuracy, as many of the studies did, provides a fairly comprehensive overview of the potential benefits associated with multiple interviews. However, the amount and proportion of unique and repeated recall was less frequently explored. Therefore, it is difficult to ascertain the relative amount of additional correct information that is obtained via a second interview; a key factor in determining the costs and benefits of a second interview. Additionally, one downside of multiple interviews that has been discussed in the literature is the risk of providing contradictory memories (Krix *et al.*, 2015). Although providing any new information in a second interview could be perceived by some as contradictory (rather than inconsistent), no studies examined direct contradictions in children's recall.

An important dependent variable that has not been examined, but may be crucial for determining why *not* to conduct multiple interviews, is child well-being. An argument against conducting multiple interviews is that they cause additional distress to the child (Plotnikoff & Woolfson, 2001). Although it would be challenging ethically to interview children more than once in an experiment about events that are as traumatic as those

discussed in forensic interviews, studies using existing traumatic events (such as medical emergencies) could examine whether multiple interviews cause more distress to interviewees than single ones.

Although the present study identifies topics on which there is no literature at all, there are some key limitations of the methodology which mean there may still be further gaps due to missing research, poor quality research or inconsistent findings. The current SSA did not include grey literature and only included two databases in the literature search. This may have resulted in some relevant research being missed from the final sample (McGinn, Taylor, McColgan, & McQuilkan, 2016). Additionally, relying on online searches may mean earlier research that was not available online was overlooked. Furthermore, in general, the SSA methodology does not examine the quality or the findings of the research included and thus it is possible that those studies that have been included do not provide consistent, reliable, or valid results. However, La Rooy *et al.* (2010) conclude from their reviews of the literature that if best practice interview guidelines are followed, repeated interviews often result in additional correct details. Furthermore, the key aim of this study was not to examine results but to determine whether the spread of research was sufficient for the policy change suggested in other studies (e.g., Block *et al.*, 2013; La Rooy *et al.*, 2010), and the current findings suggest this is not the case for all types of child victim/witness.

Conclusions

A comprehensive literature on the multiple interviewing of children is gradually developing. The present Study Space Analysis found a large number of studies focused on five to eight-year-old children, with the majority using somewhat ecologically-valid events. However, crucial gaps in the research have been identified, suggesting researchers should be cautious in advocating policy change at this stage. Research with the following

This is an accepted manuscript of an article published by Wiley in *Child Abuse Review*, available online at <https://onlinelibrary.wiley.com/doi/full/10.1002/car.2604>. It is not the copy of record. Copyright © 2020, Wiley.

characteristics is urgently required to inform policy decisions regarding multiple interviewing:

- 11 to 18-year-old participants,
- participants with multiple needs,
- repeated TBR events.

Further recommendations for research can be found in the supplementary material to this study. It can be concluded that multiple interviews may have great potential to improve children's informativeness. However, further research is required prior to changing policy and practice to ensure the research findings apply to more ecologically-valid conditions, a broader population of interviewees and more possible abuse types.

References

* denotes studies used in the Study Space Analysis

*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. doi: 10.2307/1131550

*Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (Experiment 1; 1990). The effects of involvement on children's memory for events. *Cognitive Development, 5*, 55-69. doi: 10.1016/0885-2014(90)90012-I

*Bjorklund, D. F., Bjorklund, B. R., Douglas Brown, R., & Cassel, W. S. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science, 2*, 99-111. doi: 10.1207/s1532480xads0202_4

*Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Douglas Brown, R., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*, 421-433. doi: 10.1002/1099-0720(200009)14:5%3C421::AID-ACP659%3E3.0.CO;2-4

Block, S. D., Foster, E. M., Pierce, M. W., Berkoff, M. C., & Runyan, D. K. (2013).

Multiple forensic interviews during investigations of child sexual abuse: A cost-

effectiveness analysis. *Applied Developmental Science*, 17, 174-183. doi:
10.1080/10888691.2013.836033

Bottoms, B. L., Rudnicki, A. G., & Epstein, M. A. (2007). A retrospective study of factors affecting the disclosure of childhood sexual and physical abuse. In M.-E. Pipe, M. E. Lamb, Y. Orbach, & A.-C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay, and denial* (pp.175-194). New York: Routledge.

*Brown, D. A., Lewis, C. N., & Lamb, M. E. (2015). Preserving the past: An early interview improves delayed event memory in children with intellectual disabilities. *Child Development*, 86, 1031-1047. doi: 10.1111/cdev.12364

*Brown, D., Pipe, M-E., Lewis, C., Lamb, M. E., & Orbach, Y. (2012). How do body diagrams affect the accuracy and consistency of bodily touch across repeated interviews? *Applied Cognitive Psychology*, 26, 174-181. doi: 10.1002/acp.1828

Brubacher, S. P., Poole, D. A., & Dickinson, J. J. (2015). The use of ground rules in investigative interviews with children: A synthesis and call for research. *Developmental Review*, 36, 15-33. doi: 10.1016/j.dr.2015.01.001

*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. doi: 10.1007/BF01499341

- *Fivush, R., McDermott Sales, J., Goldberg, A., Bahrick, L., & Parker, J. (2004).
Weathering the storm: Children's long-term recall of Hurricane Andrew. *Memory, 12*,
104-118. doi:10.1080/09658210244000397
- *Follmer Greenhoot, A., 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. doi:
10.1111/1467-8624.00027
- *Gobbo, C., Mega, C., & Pipe, M. (Experiment 1; 2002). Does the nature of the experience
influence suggestibility? A study of children's event memory. *Journal of Experimental
Child Psychology, 81*, 502-530. doi: 10.1006/jecp.2002.2662
- *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. doi: 10.1075/jnlh.1.1.05chi
- Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., & Prado, L.
(1992). Testifying in criminal court: Emotional effects on child sexual assault victims.
Monographs of the Society for Research in Child Development, 57 (5, Serial No. 229).
doi: 10.2307/1166127
- *Gordon, B. N., & Follmer, A. (1994). Developmental issues in judging the credibility of
children's testimony. *Journal of Clinical Child Psychology, 23*, 283-294.
doi:10.1207/s15374424jccp2303_6

- *Gross, J., & Hayne, H. (Experiment 2; 1999). Drawing facilitates children's verbal reports after long delays. *Journal of Experimental Psychology: Applied*, 5, 265-283. doi: 10.1037/1076-898X.5.3.265
- *Hedrick, A. M., Haden, C. A., & Ornstein, P. A. (2009). Elaborative talk during and after an event: Conversational style influences children's memory reports. *Journal of Cognition and Development*, 10, 188-209. doi: 10.1080/15248370903155841
- *Henry, L. A., & Gudjonsson, G. H. (2003). Eyewitness memory, suggestibility, and repeated recall sessions in children with mild and moderate intellectual disabilities. *Law and Human Behavior*, 27, 481-505. doi: 10.1023/A:1025434022699
- *Jack, F., Simcock, G., & Hayne, H. (2012). Magic memories: Young children's verbal recall after a 6-year delay. *Child Development*, 83, 159-172. doi: 10.1111/j.1467-8624.2011.01699.x
- Krix, A. C., Sauerland, M., Lorei, C., & Rispens, I. (2015). Consistency across repeated interviews: Contrasting police detectives' beliefs with actual eyewitness performance. *PLoS ONE*, 10 (2), e0118641. doi:10.1371/journal.pone.0118641
- *Kulkofsky, S., & Klemfuss, J. Z. (Experiment 1; 2008). What the stories children tell can tell about their memory: Narrative skill and young children's suggestibility. *Developmental Psychology*, 44, 1442-1456. doi: 10.1037/a0012849

La Rooy, D., Katz, C., Malloy, L. C., & Lamb, M. E. (2010). Do we need to rethink guidance on repeated interviews? *Psychology, Public Policy, and Law*, *16*, 373-392. doi: 10.1037/a0019909

La Rooy, D., Lamb, M. E., & Pipe, M.-E. (2009). Repeated interviewing: A critical evaluation of the risks and potential benefits. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 327-361). New Jersey: Wiley.

*La Rooy, D., Pipe, M.-E., & Murray, J. E. (Experiment 1; 2005) Reminiscence and hypermnesia in children's eyewitness memory. *Journal of Experimental Child Psychology*, *90*, 235-254. doi:10.1016/j.jecp.2004.11.002

*La Rooy, D., Pipe, M.-E., & Murray, J. E. (2007). Enhancing children's event recall after long delays. *Applied Cognitive Psychology*, *21*, 1-17. doi: 10.1002/acp.1272

*Lepore, S. J., & SESCO, B. (1994). Distorting children's reports and interpretations of events through suggestion. *Journal of Applied Psychology*, *79*, 108-120. doi: 10.1037/0021-9010.79.1.108

*McCauley, M. R. & Fisher, R. P. (1995). Facilitating children's eyewitness recall with the revised cognitive interview. *Journal of Applied Psychology*, *80*, 510-516. doi: 10.1037/0021-9010.80.4.510

- McGinn, T., Taylor, B., McColgan, M., & McQuilkan, J. (2016). Social work literature searching: Current issues with databases and online search engines. *Research on Social Work Practice, 26*, 266-277. doi: 10.1177/1049731514549423
- Malloy, L. C., La Rooy, D. J., Lamb, M. E., & Katz, C. (2011). Developmentally sensitive interviewing for legal purposes. In M. E. Lamb, D. J. La Rooy, L. C. Malloy & C. Katz (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 1-15). Chichester, UK: Wiley-Blackwell.
- Malpass, R. S., Tredoux, C. G., Schreiber Compo, N., McQuiston-Surrett, D., Maclin, O. H., Zimmerman, L. A., & Topp, L. D. (2008). Study space analysis for policy development. *Applied Cognitive Psychology, 22*, 789-801. doi: 10.1002/acp.1483
- *Melinder, A., Alexander, K., Cho, Y. I., Goodman, G. S., Thoresen, C., Lonnum, K., & Magnussen, S. (2010). Children's eyewitness memory: A comparison of two interviewing strategies as realized by forensic professionals. *Journal of Experimental Child Psychology, 105*, 156-177. doi: 10.1016/j.jecp.2009.04.004
- *Melnyk, L., & Bruck, M. (2004). Timing moderates the effects of repeated suggestive interviewing on children's eyewitness memory. *Applied Cognitive Psychology, 18*, 613-631. doi: 10.1002/acp.1013
- Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*, 340-372. doi: 10.1037/a0020518

Mills, R., Alati, R., O'Callaghan, M., Najman, J. M., Williams, G. M., Bor, W., & Strathearn, L. (2011). Child abuse and neglect and cognitive function at 14 years of age: Findings from a birth cohort. *Pediatrics*, *127*, 4-10. doi: 10.1542/peds.2009-3479

Ministry of Justice (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Guidance on Using Special Measures*.

Retrieved from

https://www.cps.gov.uk/sites/default/files/documents/legal_guidance/best_evidence_in_criminal_proceedings.pdf.

National Society for the Prevention of Cruelty to Children. (2014). *The criminal justice response to child sexual abuse*. Retrieved from

<http://www.nspcc.org.uk/globalassets/documents/research-reports/how-safe-children-2014-criminal-justice-response-child-sexual-abuse.pdf>

New Zealand Law Commission. (1996). *The evidence of children and other vulnerable witnesses*. Retrieved from <http://www.lawcom.govt.nz/our-projects/evidence-law-evidence-children-and-other-vulnerable-witnesses?id=690>

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: A systematic review and meta-analysis. *Public Library of Science Med*, *9*(11), e1001349. doi:10.1371/journal.pmed.1001349

- *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. doi: 10.1037/0012-1649.42.2.332
- *Ornstein, P. A., Gordon, B. N., & Larus, D. M. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology, 6*, 49-60. doi: 10.1002/acp.2350060103
- *Otgaar, H., Candel, I., Smeets, T., & Merckelbach, H. (2010). 'You didn't take Lucy's skirt off': The effect of misleading information on omissions and commissions in children's memory reports. *Legal and Criminological Psychology, 15*, 229-241. doi: 10.1348/135532509X471951
- *Peterson, C. (1996). The preschool child witness: Errors in accounts of traumatic injury. *Canadian Journal of Behavioural Science, 8*, 36-42. doi: 10.1037/0008-400X.28.1.36
- *Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology, 35*, 1493-1506. doi: 10.1037/0012-1649.35.6.1493
- *Peterson, C. (2010). 'And I was very very crying': Children's self-descriptions of distress as predictors of recall. *Applied Cognitive Psychology, 24*, 909-924. doi: 10.1002/acp.1636
- *Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development, 67*, 3045-3070. doi: 10.2307/1131766

- *Peterson, C., Moores, L., & White, G. (2001). Recounting the same events again and again: Children's consistency across multiple interviews. *Applied Cognitive Psychology, 15*, 353-371. doi: 10.1002/acp.708
- *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. doi: 10.1007/s10979-005-6833-6
- *Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology, 15*, S7-S24. doi: 10.1002/acp.832
- *Pipe, M.-E., Gee, S., Wilson, J. C., & Egerton, J. M. (Experiment 1 and 2; 1999). Children's recall 1 or 2 years after an event. *Developmental Psychology, 35*, 781-789. doi: 10.1037/0012-1649.35.3.781
- *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. doi: 10.1002/acp.1053
- *Pipe, M.-E., & Wilson, J. C. (1994). Cues and secrets: Influences on children's event reports. *Developmental Psychology, 30*, 515-525. doi: 10.1037/0012-1649.30.4.515

Plotnikoff, J., & Woolfson, R. (2001). *An evaluation of child witness support*. The Scottish Executive Central Research Unit. Retrieved from <http://www.scotland.gov.uk/Publications/2001/>

*Powell, M. B., & Thomson, D. M. (1997). The effect of an intervening interview on children's ability to remember one occurrence of a repeated event. *Legal and Criminological Psychology, 2*, 247-262. doi: 10.1111/j.2044-8333.1997.tb00346.x

*Price, H. L., Connolly, D. A., & Gordon, H. M. (2016). Children who experienced a repeated event only appear less accurate in a second interview than those who experienced a unique event. *Law and Human Behavior, 40*, 362-373. doi: 10.1037/lhb0000194

*Principe, G. F., Ornstein, P. A., Baker-Ward, L., & Gordon, B. N. (2000). The effects of intervening experiences on children's memory for a physical examination. *Applied Cognitive Psychology, 14*, 59-80. doi: 10.1002/(SICI)1099-0720(200001)14:1%3C59::AID-ACP637%3E3.0.CO;2-4

*Quas, J. A., Bauer, A., & Boyce, W. T. (2004). Physiological reactivity, social support, and memory in early childhood. *Child Development, 75*, 797-814. doi: 10.1111/j.1467-8624.2004.00707.x

*Salmon, K., & Pipe, M.-E. (1997). Props and children's event reports: The impact of a 1-year delay. *Journal of Experimental Child Psychology, 65*, 261-292. doi: 10.1006/jecp.1996.2362

*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. doi: 10.1002/(SICI)1099-0720(200003/04)14:2<99::AID-ACP639>3.0.CO;2-5

The Scottish Executive. (2011). *Guidance on joint investigative interviewing of child witnesses in Scotland*. Retrieved from <https://www2.gov.scot/Publications/2011/12/16102728/0>

*Steward, M. S., & Steward, D. S. (1996). Interviewing young children about body touch and handling. *Monographs of the Society for Research in Child Development, 61* (Serial No. 248). doi: 10.2307/1166205

Sullivan, P. M., & Knutson, J. F. (2000). Maltreatment and disabilities: A population-based epidemiological study. *Child Abuse & Neglect, 24*, 1257-1273. doi: 10.1016/s0145-2134(00)00190-3

*Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive interviewing and the accuracy of children's accounts. *Law and Human Behavior, 21*, 405-426. doi: 10.1023/A:1024859219764

Tyler, K. A. (2002). Social and emotional outcomes of childhood sexual abuse: A review of recent research. *Aggression and Violent Behavior, 7*, 567-589. doi: 10.1016/S1359-1789(01)00047-7

Waterhouse, G. F. (2016). *Investigating the Forensic Interviewing of Children: Multiple Interviews and Social Support*. (Unpublished doctoral dissertation). London South Bank University, England.

Table 1

An extract of the Study Space Analysis showing number of studies with children in this age group (and percentage) against descriptors of the to-be-remembered event

Cross-Study Variable	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
Event Medium							
Life experience	7 (8.2)	14 (16.5)	9 (10.6)	3 (3.5)	3 (3.5)	0 (0)	0 (0)
Staged/live	4 (4.7)	12 (14.1)	21 (24.7)	5 (5.9)	2 (2.4)	0 (0)	1 (1.2)
Video	0 (0)	0 (0)	2 (2.4)	2 (2.4)	0 (0)	0 (0)	0 (0)
Length of Time of Event							
Less than 2 minutes	0 (0)	1 (1.2)	3 (3.5)	2 (2.4)	0 (0)	0 (0)	0 (0)
3-5 minutes	0 (0)	1 (1.2)	3 (3.5)	0 (0)	1 (1.2)	0 (0)	0 (0)
6-10 minutes	0 (0)	0 (0)	3 (3.5)	1 (1.2)	0 (0)	0 (0)	0 (0)
11-30 minutes	0 (0)	2 (2.4)	4 (4.7)	1 (1.2)	0 (0)	0 (0)	0 (0)
31 minutes – 1 hour	1 (1.2)	4 (4.7)	3 (3.5)	1 (1.2)	1 (1.2)	0 (0)	0 (0)
Over an hour	0 (0)	0 (0)	1 (1.2)	0 (0)	0 (0)	0 (0)	0 (0)
Varied	7 (8.2)	9 (10.6)	6 (7.1)	3 (3.5)	3 (3.5)	0 (0)	0 (0)
Unknown	3 (3.5)	9 (10.6)	9 (10.6)	2 (2.4)	0 (0)	0 (0)	1 (1.2)
Repetition of Event							
Single	11 (12.2)	25 (27.8)	31 (34.4)	8 (8.9)	5 (5.6)	0(0)	1 (1.1)
Repeated	1 (1.1)	4 (4.4)	3 (3.3)	0(0)	0(0)	0(0)	1 (1.1)
Type of Event							
Crime	0 (0)	1 (1.2)	2 (2.4)	1 (1.2)	0 (0)	0 (0)	0 (0)

Injury/Negative	7 (8.2)	13 (15.3)	9 (10.6)	4 (4.7)	3 (3.5)	0 (0)	0 (0)
Neutral	0 (0)	1(1.2)	1 (1.2)	1 (1.2)	1 (1.2)	0 (0)	0 (0)
Positive	4 (4.7)	11 (12.9)	20 (23.5)	4 (4.7)	1 (1.2)	0 (0)	1 (1.2)
Participation							
Participated	10 (10.4)	24 (25.0)	26 (27.1)	8 (8.3)	4 (4.2)	0 (0)	1 (1.0)
Little participation	1 (1.0)	1 (1.0)	1 (1.0)	0 (0)	0 (0)	0 (0)	0 (0)
Observed	1 (1.0)	2 (2.1)	9 (9.4)	5 (5.2)	1 (1.0)	0 (0)	0 (0)
Narrative	1 (1.0)	1 (1.0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Abuse-Related Aspects							
Touch	6 (6.7)	14 (15.7)	12 (13.5)	3 (3.4)	3 (3.4)	0 (0)	0 (0)
Photograph	2 (2.2)	3 (3.4)	2 (2.2)	0 (0)	0 (0)	0 (0)	0 (0)
Asked to keep secret	0 (0)	0 (0)	1 (1.1)	0 (0)	0 (0)	0 (0)	0 (0)
Removal of toy's clothes	0 (0)	1 (1.1)	1 (1.1)	1 (1.1)	0 (0)	0 (0)	0 (0)
None	4 (4.5)	10 (11.2)	18 (20.2)	5 (5.6)	2 (2.2)	0 (0)	1 (1.1)

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **variables and ages**.

Table 2

An excerpt of the Dependent Variable x Cross-Study Variable matrix showing number of studies with children in this age group (and percentage) against key dependent variables for the first, second and third interviews

Dependent Variables	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
First Interview							
Number of Details	3 (14.3)	6 (28.6)	7 (33.3)	3 (14.3)	2 (9.5)	0 (0)	0 (0)
Accuracy	7 (10.0)	22 (31.4)	27 (38.6)	7 (10.0)	5 (7.1)	0 (0)	2 (2.9)
Second Interview							
Number of Details	3 (14.3)	6 (28.6)	7 (33.3)	3 (14.3)	2 (9.5)	0 (0)	0 (0)
Accuracy	8 (11.3)	22 (31.0)	27 (38.0)	7 (9.9)	5 (7.0)	0 (0)	2 (2.8)
Misled Details	1 (20.0)	2 (40.0)	1 (20.0)	0 (0)	1 (20.0)	0 (0)	0 (0)
Unique Recall	1 (11.1)	3 (33.3)	3 (33.3)	1 (11.1)	1 (11.1)	0 (0)	0 (0)
Repeated Recall	2 (16.7)	4 (33.3)	4 (33.3)	1 (8.3)	1 (8.3)	0 (0)	0 (0)
Accuracy and Consistency	3 (9.4)	9 (28.1)	15 (46.9)	3 (9.4)	2 (6.3)	0 (0)	0 (0)
Omissions	0 (0)	3 (60.0)	2 (40.0)	0 (0)	0 (0)	0 (0)	0 (0)
Change in Answers	1 (50.0)	1 (50.0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Cumulative Recall	0 (0)	0 (0)	2 (100.0)	0 (0)	0 (0)	0 (0)	0 (0)

Third Interview							
Number of							
Details	1 (14.3)	1 (14.3)	2 (28.6)	2 (28.6)	1 (14.3)	0 (0)	0 (0)
Accuracy	3 (13.0)	7 (30.4)	8 (34.8)	2 (8.7)	3 (13.0)	0 (0)	0 (0)
Misled Details	0 (0)	1 (50.0)	0 (0)	0 (0)	1 (50.0)	0 (0)	0 (0)
Unique Recall	0 (0)	1 (25.0)	2 (50.0)	0 (0)	0 (0)	0 (0)	1 (25.0)
Repeated							
Recall	1 (25.0)	1 (25.0)	2 (50.0)	0 (0)	0 (0)	0 (0)	0 (0)
Accuracy and							
Consistency	2 (20.0)	2 (20.0)	4 (40.0)	1 (10.0)	1 (10.0)	0 (0)	0 (0)

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **ages**.

Supplementary Online Material – A Study Space Analysis for Multiple Interviewing of Child Witnesses

Additional Results and Discussion

1.1.1 Who should re-interview?

The question of who should conduct multiple interviews is less well answered by the current literature. Although it was not always clear whether the same interviewer conducted all of the interviews with a child (see Table 1), half used different interviewers for all interviews (50.0%) and slightly more than a quarter used the same interviewer (27.8%). In some studies, children were interviewed by a combination (7.8%), which means that they may have had two or more interviews with the same interviewer and other interviews with new interviewers. However, comparison of these two conditions is necessary to determine which might produce better recall and only one study actually made this comparison (Bjorklund *et al.*, 2000).

Only three of the 43 independent samples were interviewed by professional interviewers and these all comprised samples of children under eight years old (see Table 1). This limits the age generalisability of these studies' findings. In addition, the differences between how trained students and researchers with knowledge of the interviewing literature conduct interviews and how professional interviewers with vast experience of interviewing could be myriad. Thus, although the researchers may have experienced the same training and be able to conduct a good quality interview, research using professional interviewers is necessary to be more representative of real forensic interviews.

One study evaluated different types of interviewers, comparing police interviewers with clinicians (Melinder *et al.*, 2010). In some countries, non-police interviewers conduct forensic interviews of child victims/witnesses (mainly child practitioners, for example, the Children's Houses programme in Scandinavia, Korkman *et al.*, 2017). From the current

literature, we are unable to determine if this results in more successful interviews or not. This highlights two key areas for further research; (1) Research that examines whether children experiencing multiple interviews benefit from a change in interviewer across interviews, and (2) Research that examines the benefits and disadvantages of police vs. non-police interviewers.

[Table 1 placed here]

2.1.1 What delay should there be between interviews?

The timing of the child's first interview should not, theoretically, be affected by whether they are going to experience one or more than one interview. For example, according to ABE (Ministry of Justice, 2011), children should be interviewed as quickly as possible after the allegation is made. This is based on substantial research that shows children's (and adults') memory diminishes with time (Lamb *et al.*, 2008). Thus, studies that examine multiple interviews should involve an initial retention interval similar to those found in real interviews, as this is a realistic representation of both delays related to disclosure and those related to the investigation. The majority of first interviews in the present SSA were conducted within a month of children's exposure to the TBR event (91.5% of all groups; see Table 2). The limited research suggests that a high proportion of cases are not even reported to the police within one month (e.g., at least 32% of cases, Goodman *et al.*, 1992), and that the retention interval can be highly varied. Leander (2010) found 15 children had been interviewed within three days of the assault but a further ten up to a year later and two over a year later. Although the literature on delay is not substantial, it suggests that research using an initial retention interval of over a month would be more representative of some real forensic investigations.

The delays between the first and subsequent interviews varied widely. Second interviews were generally conducted within one month (40.2%), with a third between one and six months later (33.3%) and a quarter over six months later (26.5%). When children were interviewed a third time, they were often interviewed over six months after their second interview (37.8% of age groups). Three to eight year olds experienced a wider range of delays than older children.

Only four comparisons between different delays were conducted in the studies included in the current sample, and thus further research is crucial. Knowing the ideal length of time between multiple interviews for obtaining the most accurate, detailed and useful information from children could be helpful for those planning interviews, especially if policy change encourages pre-planned multiple interviewing schedules.

[Table 2 placed here]

3.1.1 How should multiple interviews be conducted?

For most of the age groups, children were only re-interviewed once (57.4%; see Table 3). However, some were interviewed three (17.6%), four (11.1%), or five times (12.0%). Only one study involved children completing six interviews (Bjorklund *et al.*, 1998). Thus, recommendations regarding the optimal number of follow-up interviews are limited by the under-representation of studies examining more than a first and second interview. Given that in some countries children are regularly interviewed more frequently than that (for example, some studies have found testifying children to be interviewed an average of four to five times, Goodman *et al.*, 1992; Malloy *et al.*, 2007), this area of research should be developed.

[Table 3 placed here]

In regards to the types of questions asked in the multiple interviews conducted, the majority of studies included interviews that correspond with best practice. This was predominantly evident for very young participants (i.e., 2 to 3 year olds), where studies involving question types that elicit less accurate or less complete details (e.g., misleading, forced choice and yes/no questions) were rare (see Table 4). However, approximately half of all the studies with 3 to 8 year old participants purposely included suggestive questions in their interviews (e.g., misleading questions). This is beneficial in that research that has looked at interviewers' use of question types shows that even after training, poor interviewing practice continues, including suggestive questions (e.g., Lamb *et al.*, 2009; Powell *et al.*, 2014). Therefore, looking at a range of interviewing styles is important for determining whether multiple interviewing would work in practice.

The vast majority of studies included a second interview in which interviewers used the same question types as in the first interview (see Table 4). These all involved multiple interviews that asked the children about the same events again, rather than extended interviews (e.g., asking for recall on a different topic within the same event, on which there is very little research). These interviews, therefore, do not replicate some current practice guidelines, such as the Extended Forensic Interview Protocol, which covers different phases of the interview process in each of the separate interviews (National Children's Advocacy Center, 2014). However, some field studies have found interviewing techniques in first and second child interviews using alternate protocols to be conducted using similar question types (Waterhouse *et al.*, 2016).

[Table 4 placed here]

References

See main article for full references for SSA

- Bjorklund, D. F., Bjorklund, B. R., Douglas Brown, R., & Cassel, W. S. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science, 2*, 99-111. doi: 10.1207/s1532480xads0202_4
- Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Douglas Brown, R., Park, C. L., Ernst, K., & Owen, F. A. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*, 421-433. doi: 10.1002/1099-0720(200009)14:5%3C421::AID-ACP659%3E3.0.CO;2-4
- Goodman, G. S., Taub, E. P., Jones, D. P. H., England, P., Port, L. K., Rudy, L., & Prado, L. (1992). Testifying in criminal court: Emotional effects on child sexual assault victims. *Monographs of the Society for Research in Child Development, 57* (5, Serial No. 229). doi: 10.2307/1166127
- Korkman, J., Pakkanen, T., & Laajasalo, T. (2017). Child Forensic Interviewing in Finland: Investigating Suspected Child Abuse at the Forensic Psychology Unit for Children and Adolescents. In S. Johansson, K. Stefansen, E. Bakketeig, & A. Kaldal (Eds.), *Collaborating against child abuse: Exploring the Nordic Barnahus model*. London, UK: Palgrave Macmillan.

- Lamb, M. E., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2008). Factors affecting the capacities and limitations of young witnesses. In *Tell Me What Happened: Structured Investigative Interviews of Child Victims and Witnesses* (pp. 19-61). Chichester, UK: Wiley.
- Lamb, M. E., Orbach, Y., Sternberg, K. J., Aldridge, J., Pearson, S., Stewart, H. L., Esplin, P. W., & Bowler, L. (2009). Use of a structured investigative protocol enhances the quality of investigative interviews with alleged victims of child sexual abuse in Britain. *Applied Cognitive Psychology, 23*, 449-467. doi: 10.1002/acp.1489
- Leander, L. (2010). Police interviews with child sexual abuse victims: Patterns of reporting, avoidance and denial. *Child Abuse & Neglect, 34*, 192-205. doi: 10.1016/j.chiabu.2009.09.011
- Malloy, L. C., Lyon, T. D., & Quas, J. A. (2007). Filial dependency and recantation of child sexual abuse allegations. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 162-170. doi: 10.1097/01.chi.0000246067.77953.f7
- Melinder, A., Alexander, K., Cho, Y. I., Goodman, G. S., Thoresen, C., Lonnum, K., & Magnussen, S. (2010). Children's eyewitness memory: A comparison of two interviewing strategies as realized by forensic professionals. *Journal of Experimental Child Psychology, 105*, 156-177. doi: 10.1016/j.jecp.2009.04.004
- Ministry of Justice (2011). *Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Guidance on Using Special Measures.*

Retrieved from

https://www.cps.gov.uk/sites/default/files/documents/legal_guidance/best_evidence_in_criminal_proceedings.pdf.

National Children's Advocacy Center. (2014). *National Children's Advocacy Center's Extended Forensic Interview Protocol*. Retrieved from <http://www.chicagocac.org/wp-content/uploads/2015/10/EFI-Protocol.pdf>

Powell, M. B., Hughes-Scholes, C. H., Smith, R., & Sharman, S. J. (2014). The relationship between investigative interviewing experience and open-ended question usage. *Police Practice and Research: An International Journal*, *15*, 283-292. doi: 10.1080/15614263.2012.704170

Waterhouse, G. F., Ridley, A., Bull, R., La Rooy, D. J., & Wilcock, R. (2016). Dynamics of repeated interviews with children. *Applied Cognitive Psychology*, *30*, 713-721. doi: 10.1002/acp.3246

Table 1

An extract of the Study Space Analysis showing number of studies with children in this age group (and percentage) against interviewer details

Cross-Study Variable	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
Interviewer Profession							
Experimenter / Researcher / Research Assistant	4 (4.6)	11 (12.6)	11 (12.6)	1 (1.1)	1 (1.1)	0 (0)	1 (1.1)
Student	0 (0)	1 (1.1)	4 (4.6)	2 (2.3)	1 (1.1)	0 (0)	0 (0)
Professional	1 (1.1)	1 (1.1)	2 (2.3)	0 (0)	0 (0)	0 (0)	0 (0)
Mixed (including professionals and researchers)	0 (0)	1 (1.1)	1 (1.1)	0 (0)	0 (0)	0 (0)	0 (0)
Unknown	6 (6.9)	12 (13.8)	15 (17.2)	7 (8.0)	3 (3.4)	0 (0)	1 (1.1)
Interviewer Consistency							
Same	3 (3.3)	7 (7.8)	9 (10.0)	3 (3.3)	2 (2.2)	0 (0)	1 (1.1)

Different	5 (5.6)	14 (15.6)	15 (16.7)	6 (6.7)	3 (3.3)	0 (0)	2 (2.2)
Mixed	1 (1.1)	3 (3.3)	3 (3.3)	0 (0)	0 (0)	0 (0)	0 (0)
Unknown	2 (2.2)	4 (4.4)	6 (6.7)	1 (1.1)	0 (0)	0 (0)	0 (0)

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **variables and ages**.

Table 2

An extract of the Study Space Analysis showing number of studies with children in this age group (and percentage) against interview timings

Cross-Study Variable	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
First Interview Delay (or Retention Interval)							
Immediate/Same day	4 (4.3)	10 (10.6)	15 (16.0)	5 (5.3)	0 (0)	0 (0)	1 (1.1)
1-7 days	7 (7.4)	11 (11.7)	9 (9.6)	3 (3.2)	5 (5.3)	0 (0)	1 (1.1)
8 days – 1 month	1 (1.1)	5 (5.3)	6 (6.4)	2 (2.1)	0 (0)	0 (0)	0 (0)
1-2 months	0 (0)	0 (0)	2 (2.1)	0 (0)	0 (0)	0 (0)	0 (0)
2-6 months	0 (0)	1 (1.1)	2 (2.1)	1 (1.1)	0 (0)	0 (0)	0 (0)
6 months – 1 year	0 (0)	0 (0)	1 (1.1)	0 (0)	0 (0)	0 (0)	0 (0)
Short (immediate - 1 week)	0 (0)	0 (0)	1 (1.1)	0 (0)	0 (0)	0 (0)	0 (0)
Long (1-6 months)	0 (0)	0 (0)	1 (1.1)	0 (0)	0 (0)	0 (0)	0 (0)
Second Interview Delay (from first interview)							
Immediate/Same day	0 (0)	1 (1.0)	2 (2.0)	1 (1.0)	0 (0)	0 (0)	0 (0)
1-7 days	3 (2.9)	6 (5.9)	7 (6.9)	1 (1.0)	0 (0)	0 (0)	1 (1.0)
8 days – 1 month	2 (2.0)	6 (5.9)	8 (7.8)	2 (2.0)	1 (1.0)	0 (0)	0 (0)
1-2 months	3 (2.9)	3 (2.9)	4 (3.9)	3 (2.9)	0 (0)	0 (0)	0 (0)
2-6 months	1 (1.0)	5 (4.9)	11 (10.8)	3 (2.9)	1 (1.0)	0 (0)	0 (0)
6 months – 1 year	3 (2.9)	7 (6.9)	8 (7.8)	2 (2.0)	3 (2.9)	0 (0)	0 (0)
Over 1 year	1 (1.0)	2 (2.0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.0)
Third Interview Delay (from second interview)							

Immediate/Same day	0 (0)	0 (0)	1 (2.7)	0 (0)	0 (0)	0 (0)	0 (0)
1-7 days	0 (0)	1 (2.7)	3 (8.1)	2 (5.4)	0 (0)	0 (0)	0 (0)
8 days – 1 month	0 (0)	3 (8.1)	2 (5.4)	0 (0)	0 (0)	0 (0)	1 (2.7)
1-2 months	1 (2.7)	1 (2.7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
2-6 months	1 (2.7)	3 (8.1)	3 (8.1)	1 (2.7)	0 (0)	0 (0)	0 (0)
6 months – 1 year	2 (5.4)	2 (5.4)	4 (10.8)	3 (8.1)	2 (5.4)	0 (0)	0 (0)
Over 1 year	0 (0)	0 (0)	1 (2.7)	0 (0)	0 (0)	0 (0)	0 (0)

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **variables and ages**.

Table 3

An extract of the Study Space Analysis depicting the number of interviews conducted with each age group (and percentage)

Number of Interviews	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
Two	6 (5.6)	21 (19.4)	24 (22.2)	5 (4.6)	4 (3.7)	0 (0)	2 (1.9)
Three	3 (2.8)	4 (3.7)	8 (7.4)	3 (2.8)	1 (0.9)	0 (0)	0 (0)
Four	1 (0.9)	2 (1.9)	5 (4.6)	3 (2.8)	1 (0.9)	0 (0)	0 (0)
Five	2 (1.9)	4 (3.7)	4 (3.7)	2 (1.9)	1 (0.9)	0 (0)	0 (0)
Six	0 (0)	1 (0.9)	1 (0.9)	0 (0)	0 (0)	0 (0)	0 (0)

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **variables and ages**.

Table 4

An extract of the Study Space Analysis showing number of studies with children in this age group (and percentage) against question types used in the first and second interviews

Question Types	Age						
	2-3	3-5	5-8	8-11	11-13	14-18	Adult
First Interview							
Free Recall	12 (3.9)	27 (8.7)	32 (10.4)	8 (2.6)	5 (1.6)	0 (0)	2 (0.6)
Open-ended	9 (2.9)	19 (6.1)	24 (7.8)	4 (1.3)	2 (0.6)	0 (0)	1 (0.3)
Wh-	7 (2.3)	14 (4.5)	19 (6.1)	6 (1.9)	5 (1.6)	0 (0)	0 (0)
Yes/No	6 (1.9)	14 (4.5)	15 (4.9)	3 (1.0)	2 (0.6)	0 (0)	0 (0)
Misleading	5 (1.6)	13 (4.2)	14 (4.5)	3 (1.0)	2 (0.6)	0 (0)	1 (0.3)
Leading	5 (1.6)	8 (2.6)	10 (3.2)	3 (1.0)	2 (0.6)	0 (0)	0 (0)
Forced-Choice	0 (0)	1 (0.3)	3 (1.0)	2 (0.6)	1 (0.3)	0 (0)	0 (0)
Second Interview							
Free Recall	12 (3.8)	27 (8.6)	32 (10.2)	8 (2.5)	5 (1.6)	0 (0)	2 (0.6)
Open-ended	9 (2.9)	20 (6.4)	25 (7.8)	4 (1.3)	2 (0.6)	0 (0)	1 (0.3)
Wh-	8 (2.5)	15 (4.8)	20 (6.4)	7 (2.2)	5 (1.6)	0 (0)	1 (0.3)
Yes/No	6 (1.9)	14 (4.5)	15 (4.8)	3 (1.0)	2 (0.6)	0 (0)	0 (0)
Misleading	4 (1.3)	12 (3.8)	14 (4.5)	3 (1.0)	2 (0.6)	0 (0)	1 (0.3)
Leading	4 (1.3)	7 (2.2)	10 (3.2)	3 (1.0)	2 (0.6)	0 (0)	0 (0)
Forced-Choice	0 (0)	2 (0.6)	3 (1.0)	2 (0.6)	1 (0.3)	0 (0)	1 (0.3)
First and Second Interview Question Type Similarity							
Same	10 (10.9)	22 (23.9)	28 (30.4)	7 (7.6)	5 (5.4)	0 (0)	0 (0)

Different	2 (2.2)	6 (6.5)	8 (8.7)	2 (2.2)	0 (0)	0 (0)	2 (2.2)
-----------	----------------	----------------	---------	----------------	--------------	--------------	----------------

Note. Cells in bold represent areas in which the number of experiments is lower than that expected if all studies were evenly distributed across **variables and ages**.